Sensing Information Assisted Routing Scheme for UAV Networks

Wenrui Liu, Yumeng Liu, and Zhiqing Wei(✉)

Beijing University of Posts and Telecommunications, Beijing, China
weizhiqing@bupt.edu.cn

Abstract. Flying Ad-Hoc Network (FANET) is widely used in network communication services in military, emergency relief, and environmental monitoring. A good routing protocol can provide a guarantee for its reliable transmission in harsh communication conditions. Due to the unique high mobility and frequent topology changes of UAV nodes, traditional routing protocols cannot meet the needs of establishing efficient and reliable paths in the UAV networking process. Therefore, this paper proposes a sensing information-assisted routing algorithm for UAV networks. Utilizing the perception information obtained from the interaction between the UAV and the environment, the movement prediction method is introduced into the routing strategy. Through the prediction of the node movement state, the relative position between nodes can be judged, and the survival time of the link can be calculated. Moreover, in route maintenance, it is possible to determine whether a link is disconnected based on the relative position and movement status of nodes, and re-build routes before disconnection, reducing the packet loss rate during transmission and improving network performance. At the same time, the reinforcement learning method Q-Learning is used to assist routing decisions. Most existing Q-Learning-based protocols use fixed parameters. In the scheme proposed in this paper, Q-Learning parameters can be adaptively adjusted according to network conditions. Comprehensively measure multiple indicators such as transmission delay, channel conditions, and link quality, increase routing metrics, and perform multi-objective optimization to find the optimal path between the source and the target.

Keywords: Routing Protocol · UAV ad hoc network · Reinforcement Learning

1 Introduction

In recent years, UAVs (Unmanned Aerial Vehicles, UAVs) have been widely studied in different military and civilian fields due to their advantages of low

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant U21B2014, 62271081, and 92267202, and in part by the National Key Research and Development Program of China under Grant 2020YFA0711303.

J. Wang (Ed.): MobiMedia 2023, LNICST 576, pp. 309–320, 2025.
https://doi.org/10.1007/978-3-031-60347-1_25

cost, simple structure, flexible lifting. UAVs have many uses, such as performing surveillance and monitoring tasks, as an aerial data collection equipment, providing communication network services in areas of natural disasters and conflicts [1].

For a single UAV, there are many limitations such as low task execution efficiency, insufficient anti-interference ability, unreliable communication performance, so there is the concept of UAV cluster. Multiple UAVs can quickly build a communication network in a self-organizing manner without any predetermined infrastructure, the so-called multi-UAV network or Flying Ad-hoc Networks (FANET) [2].

In FANET, the rapid movement of UAV nodes leads to frequent topology changes, so there is a serious problem of poor communication quality. The high mobility of nodes makes it difficult to predict the state of the link, so the link quality parameter needs to be updated more frequently. FANET is a wireless network with dynamic and unsustainable topology, thus, routing in a network with easy-to-disconnect characteristics is one of the main challenges of FANET [3].

Many routing protocols have been designed for wireless ad hoc networks, such as proactive routing protocols, reactive routing protocols and geographic routing protocols [4]. Proactive routing protocols create routes before forwarding data packets, and all nodes periodically maintain routing table information, which will generate large control overhead. In contrast, reactive routing protocols create routes when packets are forwarded, but they introduce greater latency due to the need for path discovery. Due to the change of network topology, routing based on location information becomes one of the main options to improve routing performances. GPSR [5] (Greedy Perimeter Stateless Routing for Wireless Networks) is a typical routing protocol based on geographical information. It designs data transmission strategy according to the geographical information of network nodes. When the routing hole problem occurs frequently, the number of hops will increase, because the decision-making next hop is random, which leads to the degradation of routing performance. However, owing to the high dynamics of nodes and the instability of topology in FANET, an adaptive and highly autonomous routing protocol is required, which can find reliable neighbors to complete data transmission by adaptively sensing the changes of the environment.

More and more researchers have begun to use machine learning algorithms to solve communication problems. Compared with other algorithms in the field of machine learning, reinforcement learning is more suitable for solving routing problems in the network. First of all, reinforcement learning does not need to learn and test based on the existing training data sets, and the whole process of its learning is very similar to steps of continuous routing establishment due to data transmission requirements in the networking process. Therefore, many scholars have proposed routing algorithms based on reinforcement learning methods. Q-Learning is an adaptive reinforcement learning scheme [6], which receives input feedback from the environment and provides support for adaptive routing protocol scheme design. In Q-Learning, the agent can continuously adjust its action strategy according to the rewards of environmental feedback to better adapt to the dynamic topology. Based on the Q-Learning algorithm, this paper proposes a scheme that utilizes Sensing information to assist routing selection.

This paper is organized as follows. Section 2 reviews the relevant research content of UAV Ad Hoc routing protocols and relevant literatures in this field. Section 3 presents the system model of the proposed scheme. Section 4 introduces in detail the sensing information assisted routing scheme for UAV network proposed. Section 5 simulates and analyzes the performance of the routing scheme proposed in this paper. Finally, Sect. 6 summarizes all the contents of this paper.

2 Related Work

The flying ad hoc network FANET proposed based on the mobile ad hoc network can be regarded as a new form in which all nodes are UAVs. It can provide reliable real-time network communication between multiple UAVs, so that multiple UAVs can quickly perform various tasks collaboratively, and it becomes a solution to the communication problem between multiple UAVs. The main purpose of FANET routing protocol is to improve routing stability and network data transmission efficiency.

In recent years, the applications in UAV Ad Hoc Networks have become more and more diversified, and higher requirements have been put forward for the intelligence of routing protocols. Moreover, since the emergence of AI algorithms such as machine learning, many researchers have also applied various machine learning algorithms to routing protocols to solve problems such as routing selection in the network.

Boyan and Littman first used reinforcement learning to solve routing problems, and [7] proposed an adaptive algorithm called Q-routing. This algorithm considers the shortest path from the source node to the destination node, but in many cases the shortest path is not necessarily the best transmission path, and there may be certain network congestion. Jung et al. proposed a Q-learning-based geographic routing protocol for UAV networks called QGeo, which reduces the impact of high node mobility through position estimation [8]. Compared with traditional routing protocols, this scheme has obvious improvements in the performance of end-to-end delay and network overhead. Stefania et al. [9] proposed a Q-SQUARE algorithm, which models UAV path planning as a Markov decision process, uses information such as residual energy of nodes and transmission delay to select routes that can guarantee transmission quality to ensure good communication in the UAV network. [10] used the reinforcement learning algorithm to determine an optimal route by considering the remaining energy and node stability in the network, which prolongs the life of the network and reduces the network energy consumption and link interruption probability. In [11], a Q-learning-based routing protocol (QGrid) for vehicular ad hoc networks is proposed, which divides node locations into grids and makes routing decisions from both macro and micro aspects. This method improves the information delivery rate between mobile nodes.

On the basis of previous research, this paper comprehensively measures multiple indicators such as transmission delay, channel conditions, and link quality, and can dynamically adjust the parameters of the Q-Learning model. Multi-objective optimization is carried out during routing selection to seek the optimal path between the source and the target.

3 System Model

3.1 FANET Model

In this paper, we consider a FANET composed of multiple UAV nodes. In a multi-UAV network, the UAV sending data is regarded as the source node, the UAV receiving data is regarded as the target node, and the other UAVs are regarded as relay nodes for forwarding data.

In this paper, we consider a FANET composed of multiple UAV nodes. In a multi-UAV network, the UAV sending data is regarded as the source node, the UAV receiving data is regarded as the target node, and the other UAVs are regarded as relay nodes for forwarding data.

Assuming that there are n UAVs in the network, and all UAVs are deployed in 3D space. Each device is equipped with GPS positioning, inertial measurement unit (IMU), camera, sensor and wireless communication interface, etc. All UAVs can use GPS to know their own location information and moving speed, and has a communication range with a constant distance of R anywhere in the network (see Fig. 1).

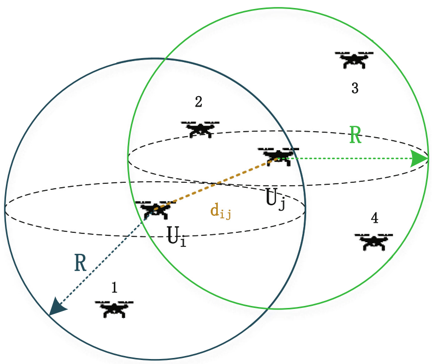

Fig. 1. FANET model.

3.2 Mobile Prediction Model

In a 3D scene, the state vector of the UAV node motion can be defined as

$$X = [x, y, z, v_x, v_y, v_{z,} a_x, a_y, a_z]^T \tag{1}$$

and the observation vector is $Z = [x, y, z]^T$. Supposing that the state vector at time t is X_t, so [12]

$$X_{t+1} = AX_t + w_t \tag{2}$$

$$Z_t = HX_t + v_t \tag{3}$$

where X_{t+1} represents the predicted state for the next moment of X_t .$A$is state transition matrix, and w_t is process noise and it follows Gaussian distribution $w_t \sim N(0, Q)$. Z_t is observations of position vectors and the measurement noise v_t is assumed to be a Gaussian distribution $v_t \sim N(0, R)$ The measurement matrix H is as follows:

$$H = \begin{bmatrix} 1\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,1\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,1\,0\,0\,0\,0\,0\,0 \end{bmatrix} \tag{4}$$

In this paper, UAV node movement adopts Gauss-Markov Mobility (GMM) model, and node acceleration is as follows [13]:

$$a_{t+1} = \lambda a_t + (1 - \lambda)\overline{a} + \sqrt{(1 - \lambda^2)}a_t \tag{5}$$

According to the physical meaning, X_{t+1} is updated by:

$$X_{t+1} = AX_t + \mu + w_t \tag{6}$$

And A and μ can be defined as:

$$A = \begin{bmatrix} 1\,0\,0 & \Delta T & 0 & 0 & \Delta T^2/2 & 0 & 0 \\ & 1\,0 & 0 & \Delta T & 0 & \Delta T^2/2 & 0 \\ & & 1 & 0 & 0 & \Delta T & 0 & \Delta T^2/2 \\ & & & 1 & 0 & 0 & \Delta T & 0 & 0 \\ & & & & 1 & 0 & 0 & \Delta T & 0 \\ & & & & & 1 & 0 & 0 & \Delta T \\ & & & & & & \lambda_x + \sqrt{1-\lambda_x^2} & & \\ & & & & & & & \lambda_y + \sqrt{1-\lambda_y^2} & \\ & & & & & & & & \lambda_z + \sqrt{1-\lambda_z^2} \end{bmatrix} \tag{7}$$

$$\mu = \begin{bmatrix} 0\,0\,0\,0\,0\,0\,(1 - \lambda_x)\,\bar{a_x}\;(1 - \lambda_y)\,\bar{a_y}\;(1 - \lambda_z)\,\bar{a_z} \end{bmatrix}^T \tag{8}$$

where λ_x , λ_y and λ_z are random parameters with sizes ranging from 0 to 1. By adjusting the size of the random parameters, different degrees of randomness can be simulated.

In the mobile prediction stage [12]:

$$\widehat{X_t'} = A\widehat{X_{t-1}} + \mu \tag{9}$$

$$P_t' = AP_{t-1}A^T + Q \tag{10}$$

where P_t is the covariance matrix.

The update formula is [12]:

$$\widehat{X_t} = \widehat{X_t{}'} + K_t(Z_t - H\widehat{X_t{}'}) \tag{11}$$

$$K_t = P_t{}'H^T(HP_t{}'H^T + R)^{-1} \tag{12}$$

$$P_t = (I - K_tH)P_t{}' \tag{13}$$

where K_t is the Kalman filter gain.

4 Q-learning in Routing Scheme

For the UAV Ad Hoc Network, based on the Q-learning algorithm, this paper proposes an efficient routing scheme SIRS (Sensing Information Assisted Routing Scheme).

4.1 Q-Learning Model

The Q-learning process consists of 4-tuples (S, A, P, R), where S represents the state set, A represents the possible behavior set, P is the state transition probability, and R is a reward function. The basic iterative formula for the Q-value is as follows [14]:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \tag{14}$$

During the routing decision process, data packets originated from UAVs and are then transmitted to target nodes through a multi-hop network, and the entire network is identified as an environment and learning system. Each data packet in the network is regarded as an Agent, and the state of the Agent indicates the node where the data packet is located. When the data packet is at node i, the state associated with the data packet is s_i. a_{ij} indicates that the action to be taken is to forward data from node i to neighbor node j. After the action is completed, the state of the Agent changes from s_i to s_j, and the reward value r for this action is obtained.

4.2 Sensing Performance Parameters

Delay. In this section, the network delay mainly focuses on queuing delay and transmission delay, and other delays are ignored. Assuming that each node in the network has a transmission queue, the time interval for receiving data packets on each node obeys Poisson distribution, and the node receives λ data packets per second. When the node performs data forwarding processing, the subsequent received data packets are queued in the queue, and the processing and forwarding time of each data packet is as follows:

$$T_s = \frac{L}{B} \tag{15}$$

where L is the size of the packet, and B is the link bandwidth. The average delay of data packets transmitted on the link $i-j$ is defined as:

$$W_{ij} = \frac{1}{\frac{1}{T_s} - \lambda_{ij}} = \frac{1}{\frac{B}{L} - \lambda_{ij}} \tag{16}$$

Node Load. Assuming that node i forwards m data packets within time interval T, the size of data packet j is Q_j, and the maximum cache length of node is $Q_{\max}$, so the load degree of node i is as follows:

$$M_i = \frac{\sum_{j=1}^{m} Q_j}{Q_{\max}} \tag{17}$$

Link Expiration Time. The link expiration time is the remaining time of the communication link between two nodes. Assuming that the single-hop link expiration time between node i and its neighbor node j at time $T = t$ is LET, it means that neighbor node j will leave the communication transmission range of node i at time $T = t + LET$. Assuming that nodes i and j are neighbors, and the communication range of a single UAV node is R. And the current position of node i is (x_1, y_1, z_1), the moving speed is (v_{1x}, v_{1y}, v_{1z}), the position of node j is (x_2, y_2, z_2), and the moving speed is (v_{2x}, v_{2y}, v_{2z}). It can be deduced that [15]:

$$\begin{aligned} &[(x_2 - x_1) + (v_{2x} - v_{1x})t]^2 + \\ &[(y_2 - y_1) + (v_{2y} - v_{1y})t]^2 + \\ &[(z_2 - z_1) + (v_{2z} - v_{1z})t]^2 = R^2 \end{aligned} \tag{18}$$

So $LET = t$ is as follows:

$$LET = \frac{\sqrt{(b^2+d^2+f^2)R^2 - [(ad-bc)^2 + (af-be)^2 + (ed-cf)^2]} - (ab+cd+ef)}{(b^2+d^2+f^2)} \tag{19}$$

$$\begin{aligned} a = x_2 - x_1, b = v_{2x} - v_{1x} \\ c = y_2 - y_1, d = v_{2y} - v_{1y} \\ e = z_2 - z_1, f = v_{2z} - v_{1z} \end{aligned} \tag{20}$$

Link Comprehensive Quality. This paper proposes a comprehensive evaluation method for the link quality. The link quality P can be comprehensively measured by using the load degree of the node and the link expiration time:

$$P = \lambda_1 LET - \lambda_2 M \tag{21}$$

where $\lambda_i (i = 1, 2)$ is positive number, representing the weighting coefficient. For different network scenarios, the size of the coefficient can be adjusted to change the weight.

4.3 Reward Function

In the Q-learning algorithm, the reward function is the system feedback when the goal is completed, and it is the value measure for the choice of this action. The higher the reward value, the more meaningful the action is.

In the SIRS routing scheme, in order to better select effective nodes for data transmission, the reward here is defined as a comprehensive measure of delay and link quality. The reward value is defined as follows:

$$r = a * P - b * W \tag{22}$$

where W is the transmission delay, and P is the link comprehensive quality, whose specific definitions are in Sect. 4.2. Both a and b are weight coefficients, and $a + b = 1$ is satisfied.

The expression of the reward function is as follows [14]:

$$f_R = \begin{cases} r_{\max} & when\ s_{t+1}\, is\ destination \\ -r_{\max} & when\ s_t\, is\ localmaximum \\ r & otherwise \end{cases} \tag{23}$$

4.4 Adaptive Q-Learning Parameters

The learning rate α is a number between 0 and 1, which determines the extent to which the newly acquired information covers the old information. The higher the learning rate, the faster the Q value update. For the UAV network, its nodes move relatively faster and the link stability is weaker, so the update speed of Q value should be faster. An adaptive learning rate allocation method is introduced here, and the learning rate is adjusted according to the average transmission delay of the link. The learning rate α is defined as:

$$\alpha_{i,j} = \begin{cases} 1 - e^{-W_{i,j}} & , W_{i,j} \neq 0 \\ 0.3 & , W_{i,j} = 0 \end{cases} \tag{24}$$

where $W_{i,j}$ represents the average delay from node i to node j.

The discount factor γ represents future stability, and the higher the value, the more stable the expectation for the future Q value. In FANET, it is necessary to seek neighbor nodes that meet the conditions for data forwarding, so changes in the neighbors of nodes will have an important impact. We define the discount factor as a number related to the change of neighbor nodes. For node i, its discount factor can be defined as:

$$\gamma_i = \sqrt{1 - \frac{\left|\left(N_i(t) \cap \overline{N_i(t-1)}\right) \cup \left(N_i(t-1) \cap \overline{N_i(t)}\right)\right|}{\left(N_i(t) \cup N_i(t-1)\right)}} \tag{25}$$

where $N_i(t)$ and $N_i(t-1)$ represent the set of neighbors around node i at time t and *t-1*.

5 Performance Evaluation

According to the model designed above, this section conducts network simulation for the proposed SIRS routing scheme. By setting different network parameters, such as the number of nodes, the maximum moving speed of nodes and the amount of transmitted data, etc., the relevant performance of the SIRS routing scheme is verified.

5.1 Mobile prediction

In the same time interval (1 s), the position information and motion information of the UAV node during the movement process are predicted, and the relative error between the predicted information and the actual information is calculated, which can verify the performance of the mobile prediction model.

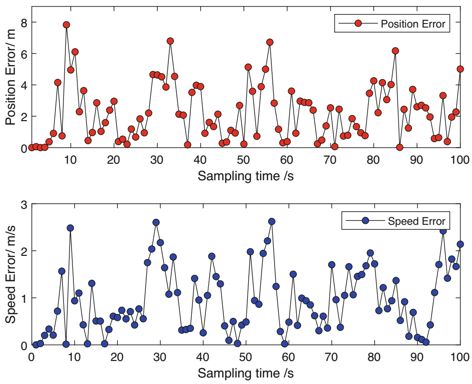

Fig. 2. Mobile prediction error.

In the 100 sampling process, after using the Kalman filter method, the position error of the node is within 0–8 m, the average error is 2.25 m, and the maximum error is 7.83 m. The speed error of the nodes is in the range of 0–3 m/s, the average speed error is 0.95 m/s, and the maximum error is 2.62 m/s(see Fig. 2).

5.2 SIRS Performance

In the simulation environment, the moving speed and the number of nodes are changed respectively to verify the influence of the moving speed and the number of nodes on the routing delay and reliability, and further verify the overall performance of the routing scheme proposed in this paper.

As shown in Fig. 3, as the node's moving speed increases, the transmission delay and packet loss rate of the network also increase. In the simulation, the node moving speed increases from the lowest speed of 10 m/s to 60 m/s, the

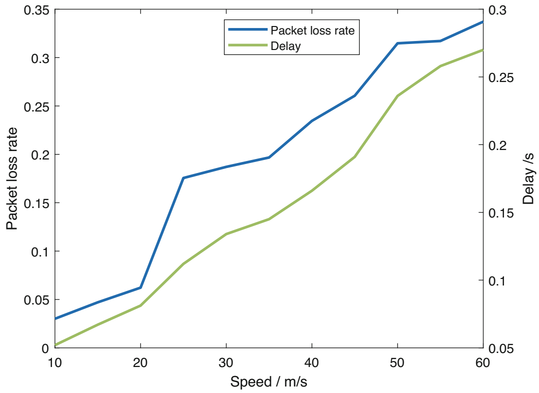

Fig. 3. Mobile prediction error.

packet loss rate increases from about 3% to 33.7%, and the transmission delay increases from 52 ms to about 270 ms. In the process of node acceleration, the packet loss rate and transmission delay are increasing. It can be seen that the high-speed mobility of UAV nodes is one of the important factors affecting the overall network data transmission. Therefore, when considering the communication protocol of UAV Ad Hoc Network, it is necessary to fully consider the impact of high mobility of nodes.

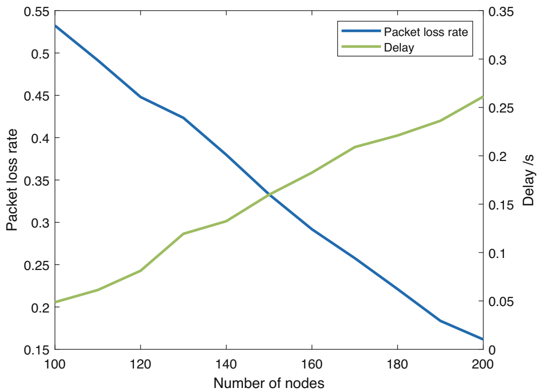

Fig. 4. Mobile prediction error.

In Fig. 4, as the number of nodes increases, the transmission delay increases and the packet loss rate decreases. The number of nodes continued to increase from 100 to 200, the transmission delay increased from about 50 ms to about 260 ms, and the packet loss rate decreased from 53% to about 15%. In the network, as the number of nodes increases, the number of neighbors of each node will also increase, and the probability of successful path construction will also increase, so the packet loss rate will decrease as the number of nodes in the network increases. On the contrary, the transmission delay increases with the increase of the number of nodes, because in the route construction process, the transmission hops of the transmission path will increase with the number of nodes, and more nodes in the network will also cause certain interference .

5.3 Performance Comparison

In order to verify the performance of SIRS proposed in this paper, the network overhead of SIRS is compared with traditional routing schemes.

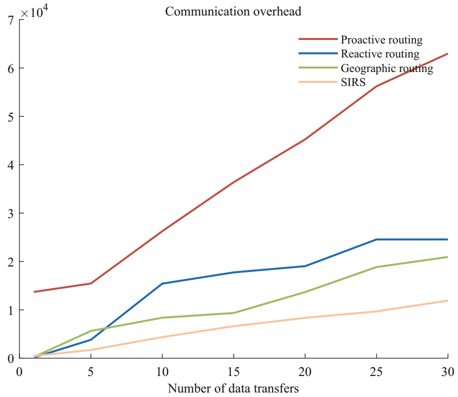

Fig. 5. Mobile prediction error.

In Fig. 5, in terms of communication overhead, SIRS is 81% lower than the proactive routing scheme, 51% lower than the reactive routing scheme, and 43% lower than the geographic routing scheme, and its performance has been greatly improved.

6 Conclusion

In this paper, we propose a sensing information-assisted routing scheme for UAV networks. Firstly, the UAV network system model and the movement prediction model based on Kalman filter are constructed. Through the movement prediction, the position information and motion of the UAV nodes can be known, which can reduce the interaction of motion information between nodes and reduce network overhead. Then a series of perceptual performance index parameters are defined, such as transmission delay, node load, link connectivity, etc. Based on the Q-Learning algorithm, SIRS (Sensing Information Assisted Routing Scheme) is designed. By setting adaptive parameters related to network performance, it can continuously adjust the path selection scheme and improve the adaptive capability of routing.

References

1. Micheletto, M., Petrucci, V., Santos, R., et al.: Flying real-time network to coordinate disaster relief activities in urban areas. Sensors **18**(5), 1662 (2018)
2. Bekmezci, I., Sahingoz, O.K., Temel, S.: Flying ad-hoc networks (FANETs): a survey. Ad Hoc Netw. **11**(3), 1254–1270 (2013)
3. Zhang, Q., Jiang, M., Feng, Z., et al.: IoT enabled UAV: network architecture and routing algorithm. IEEE Internet Things J. **6**(2), 3727–3742 (2019)
4. Fan, X., Cai, W., Lin, J.: A survey of routing protocols for highly dynamic mobile ad hoc networks. In: 2017 IEEE 17th International Conference on Communication Technology (ICCT), pp. 1412–1417. IEEE(2017)
5. Karp, B., Kung, H.T.: GPSR: Greedy perimeter stateless routing for wireless networks. In: Proceedings of the 6th Annual International Conference on Mobile Computing and Networking, pp. 243–254. ACM(2000)
6. Jung, W.S., Yim, J., et al.: QGeo: Q-learning-based geographic ad hoc routing protocol for unmanned robotic networks. IEEE Commun. Lett. **21**(10), 2258–2261 (2017)
7. Littman, M., Boyan, J.: A distributed reinforcement learning scheme for network routing. In: International Workshop on Applications of Neural Networks to Telecommunications, pp. 55–61. Psychology Press (2013)
8. Jung, W.S., Yim, J., Ko, Y.B.: QGeo: Q-Learning based geographic ad-hoc routing protocol for unmanned robotic networks. IEEE Commun. Lett. **21**(10), 2258–2261 (2017)
9. Colonnese, S., Cuomo, F., Pagliari, G., et al.: Q-SQUARE: a Q-learning approach to provide a QoE aware UAV flight path in cellular networks. Ad Hoc Netw. **91**, 101872 (2019)
10. Khan, M.F., Yau, K.L.A.: Route selection in 5G-based flying ad-hoc networks using reinforcement learning. In: 2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), pp. 23–28. IEEE (2020)
11. Li, R., Li, F., Li, X., et al.: QGrid: Q-learning based routing protocol for vehicular ad hoc networks. In: 2014 IEEE 33rd International Performance Computing and Communications Conference (IPCCC), pp. 1–8. IEEE (2014)
12. Xu, H., Cho, J., Lee, S.: A predictive and synchronized neighborhood tracking scheme for mobile ad hoc networks. In: 7th IEEE International Conference on Computer and Information Technology (CIT 2007) , pp. 429–434. IEEE (2007)
13. Liang, B., Haas, Z.J.: Predictive distance-based mobility management for multidimensional PCS networks. IEEE ACM Trans. Netw. **11**(5), 718–732 (2003)
14. Watkins, C., Dayan, P.: Q-learning. Mach. Learn. **8**(3–4), 279–292 (1992)
15. Hua, E.Y., Haas, Z.J.: Mobile-projected trajectory algorithm with velocity-change detection for predicting residual link lifetime in MANET. IEEE Trans. Veh. Technol. **64**(3), 1065–1078 (2015)