Machine Learning-Based Beamforming for Unmanned Aerial Vehicles Equipped with Reconfigurable Intelligent Surfaces

Ishtiaq Ahmad, Ramsha Narmeen, Zdenek Becvar, and Ismail Guvenc

Abstract

Unmanned aerial vehicles (UAVs) equipped with reconfigurable intelligent surfaces (RISs) have emerged as a promising technology for numerous applications involving aerial networks. However, the UAV-RIS concept faces challenges related to the deployment of the UAV-RIS, especially in cases, where UAV-RIS is combined with emerging technologies, such as beamforming, sensitive to propagation channel variation. In this article, we first overview various use-cases of UAV-RIS beamforming considering practical scenarios. Aiming to improve the performance of communication channels, we propose a machine learning-based beamforming policy for UAV-RIS by employing prioritized experience replay (PER) based deep Q-Network (DQN). Compared to traditional approaches, the proposed PER DQN-based beamforming for UAV-RIS communication provides significant enhancements in performance. Finally, we highlight some potential directions for future research.

Introduction

In recent years, major advancements in the fifth-generation mobile networks towards the sixth-generation (6G) took place. Future 6G networks should offer ultra-high data rates, 3D space global coverage and connectivity, extremely high reliability, and low latency. However, ultra-reliable and high-capacity wireless communication is challenged by the random and time-varying wireless propagation channels.

To deal with time and frequency selective wireless channels, numerous techniques have been adopted in the literature. For example, adaptive modulation and coding schemes [1], different space-time-frequency diversity techniques [2], dynamic power/rate control, and beamforming are used according to the channel conditions [3]. However, the above-mentioned techniques not only impose an additional overhead, but also allow only limited control over the wireless channel. Hence, the final hurdle to establishing a system with high-capacity, ultra-reliability, and low latency in time and frequency selective wireless channels persists.

Recently, achieving high-capacity, ultra-reliability, and low latency communications in the time/frequency selective channels is explored with the assistance of unmanned aerial vehicles (UAVs). Based on the local radio environment, raising the height of UAVs can help to avoid blockages [4]. The UAVs allow to adjust their position to create favorable communication channels with the ground terminals [5]. To further mitigate blockage problems, reconfigurable intelligent surface (RIS) [6] assisting the UAV communication is adopted to increase the overall network performance [4]. Using the RIS is a promising way to achieve high-capacity and ultra-reliable communication via low-cost passive/active reflecting elements integrated on a plane surface. The RIS elements individually reflect the incident signal while also changing either its phase or amplitude [7, 8].

The RIS can be deployed on fixed locations, for example, on buildings, to extend coverage even to blank spots, where the signal is blocked. However, in a highly congested environment, such as a metropolitan city, the signal from the transmitter to the receiver may need to be reflected by many fixed RISs to avoid the obstacles, resulting in a significant path loss and and significantly increased deployment cost. Besides, to cover all possible coverage holes, where users may potentially appear, many RISs are required even if some may remain unused when no users are active in exclusive coverage areas of these RISs. In contrast, the RIS mounted on the UAV (UAV-RIS) [4] provides panoramic full-angle reflection (360°) towards the user located on the ground and introduces flexibility to significantly extend coverage and improve network throughput [9].

In the UAV-RIS networks, the beamforming scheme can be adapted into various coverage areas making the management more challenging than in the networks, where RIS is deployed on fixed locations with local coverage only. Performance of the beamforming in UAV-RIS is deteriorated due to the fluctuation of the UAVs, thus giving a rise to the problem of beam misalignment [10]. Hence, dynamic beam alignment design necessitates improving reliability by continuously interacting with the environment and correct-

Ishtiaq Ahmad is with Gomal University, Pakistan, and also with Czech Technical University in Prague, Faculty of Electrical Engineering, Czech Republic; Ramsha Narmeen (corresponding author) and Zdenek Becvar are with Czech Technical University in Prague, Faculty of Electrical Engineering, Czech Republic; Ismail Guvenc is with North Carolina State University, USA.

Digital Object Identifier: 10.1109/MWC.004.2100694

ing the beam alignment. A reasonably appealing machine learning-based method to reduce beam misalignment is to manage beamforming based on reinforcement learning (RL) and combine it with a deep neural network to choose a reliable beamforming policy in real-time [11].

Although the UAV equipped with RIS is discussed in [12, 13], a detailed examination of current applications and corresponding benefits when both UAV and RIS are combined together in a wireless network, and utilizing various machine learning techniques for beamforming are still missing. This gap motivates our current work. In this article, we first overview promising practical applications of UAV-RIS in future 6G mobile networks and, subsequently, we give an overview of various machine learning-based solutions for effective beamforming in the UAV-RIS aided networks. Subsequently, we investigate how to meet communication reliability and quality-of-service (QoS) requirements under realistic time-varying channels. The contributions of our work are summarized as follows:

- we provide a comprehensive discussion of use-cases of the UAV-RIS concept and their deployment scenarios;
- a comparative overview of machine learning algorithms to determine the policies for beamforming in the UAV-RIS is presented;
- we introduce a machine learning-based beamforming approach that relies on prioritized experience replay (PER);
- we show an improvement in the performance of the UAV-RIS aided networks using proposed PER-based beamforming compared to state-of-the-art-works.

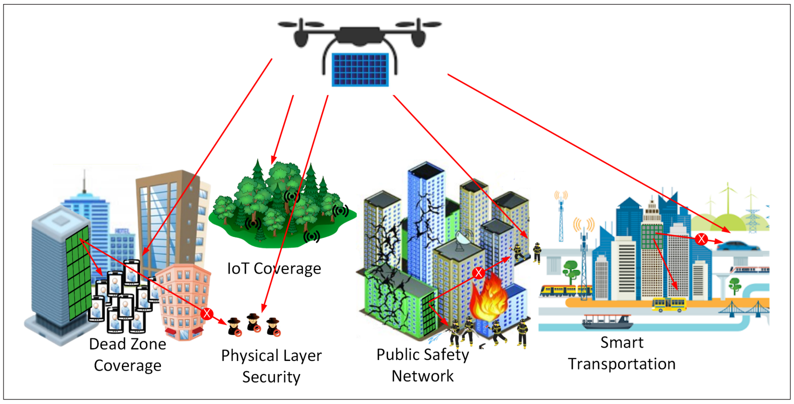

FIGURE 1. UAV-RIS use-cases for improved coverage and reliability.

Use-Cases for UAVs Equipped With RIS

In this section, we discuss numerous use-cases of the UAV-RIS-aided wireless networks. We also consider the fundamental design issues in the UAV-RIS wireless network and provide potential practical solutions.

The UAV-RIS concept and its significant use-cases in the future mobile communication network, such as UAV-RIS used for enhancing coverage, secure communication, traffic monitoring, public safety users, and IoT devices are illustrated in Fig. 1 and discussed in following subsections.

UAV-RIS Enhanced Coverage

A patch of small RIS attached to the UAVs can significantly improve the strength of the received signal and the UAVs at high altitudes can provide full angle coverage to the ground users. The UAV-RIS concept can efficiently manage the features of an incident signal, like the amplitude, phase, and frequency, to provide reliable coverage to a specific area. By considering the buildings blocking the users' reception of the signal, the position of the UAV-RIS is determined first and, then, the RIS controller determines the phase shift and direction of the incoming signal.

The UAV-RIS network is more likely to establish LoS links with the users when compared with ground base stations only, as shown in Fig. 2a. In particular, the UAV-RIS can evade the surrounding obstacles (e.g., buildings or trees) more effectively by enabling the LoS link between the blocked users on the ground and the terrestrial base stations by using the RIS. Hence, the UAV-RIS concept can improve the performance of ground users. To further boost the performance, the UAV-RIS beamforming exploits machine learning techniques such as RL and deep Q-Network (DQN). In this way, beamforming is optimized to minimize the communication delay, maximize throughput, or reduce energy consumption.

UAV-RIS Secure Communication

Physical layer security can efficiently minimize the leakage of important information and improve data security. The UAV-RIS concept can be used to enhance the security of the physical layer in the UAV to ground communication. The UAV-RIS systems are competent in providing strong LoS links by dynamically adjusting the transmission to improve security. Due to the ability to reshape wireless channels, the UAV-RIS system enables a more controllable communication environment to enhance physical layer security.

In addition, high mobility and flexible deployment of the UAV-RIS necessitate improved security. Merging the UAV-RIS concept and physical layer security introduces an appealing approach to provide pervasive secure wireless communication for next-generation wireless systems. In particular, as illustrated in Fig. 2b, the UAV-RIS concept can reduce signal leakage to ground eavesdroppers by transmitting cancellation and/or jamming signals, hence, providing security to legitimate users.

UAV-RIS Aided Road Traffic Monitoring

The UAV-RIS can detect numerous alarming events such as negligence of traffic rules or safe distance, car accidents, or over-speeding. The UAV-RIS concept can be deployed in different areas on the highway or motorway to increase the safety of vehicles, pedestrians, and other road users. In fully connected road vehicles, each vehicle can report events to other vehicles through vehicle-to-vehicle communication facilitated through the UAV-RIS.

Besides, the wireless channel quality can change rapidly over time due to the highly dynamic motion of vehicles and due to blockages introduced, for example, by buildings or large vehicles. In case of uncertainties, such as random increase/decrease in speed and exploiting the safe distance among vehicles, the vehicle can inform the traffic police and road vehicles through the UAV-RIS immediately to limit the number of accidents. The UAV-RIS concept can intelligently adjust the beamforming to enhance the quality of vehi-

Humans may also wear many IoT devices, such as health monitoring sensors, on them or in their surroundings, hence, it is convenient to deploy the UAV-RIS concept to provide high-speed communication, larger bandwidth, minimum latency, and energy consumption to ensure the prevalence of the IoT.

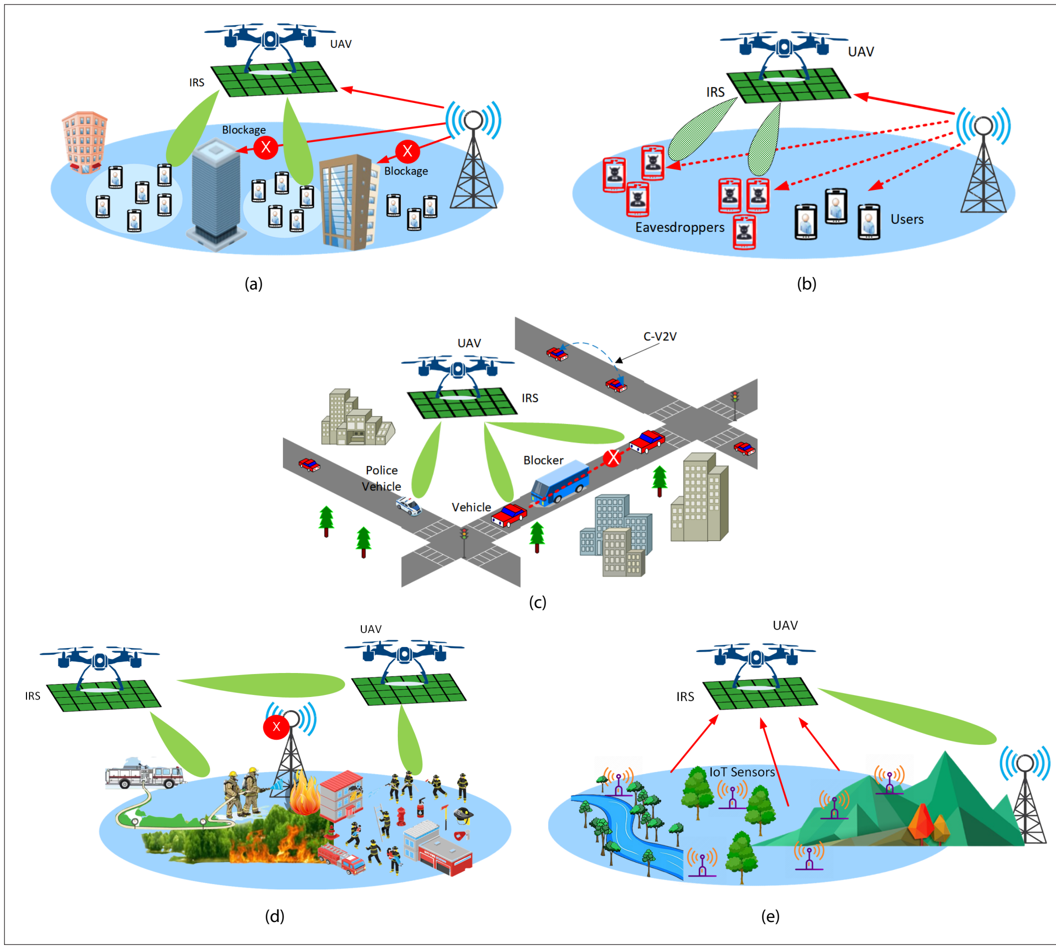

FIGURE 2. Use-cases of UAV-RIS: a) UAV-RIS enhanced coverage; b) UAV-RIS secure communication; c) UAV-RIS aided road traffic monitoring; d) UAV-RIS aided public safety networks; e) UAV-RIS aided IoT devices.

cle-to-vehicle links and increase the coverage of signal with its high energy efficiency and low cost. The UAV-RIS can also be deployed in a specific area for an event reporting to the traffic police. An example illustration for UAV-RIS aiding vehicle-to-vehicle communication is shown in Fig. 2c.

UAV-RIS Aided Public Safety Networks

The UAV-RIS concept can be deployed in case of common communication infrastructure being destructed to provide flexible services, quick deployment, low latency, coverage guarantees, and extended battery lifetime for public safety users. As shown in Fig. 2d, the UAV-RIS provides coverage to mission-critical high priority public safety users to protect against threats such as an act of terrorism, natural disasters, and technological accidents. This coverage helps to exchange information for situational awareness (e.g., voice, data, or video.) and improves the cooperation among public safety users.

The base stations could be destroyed or crash in malicious attacks or natural disasters and the remaining operational network can be inadequate to provide coverage to high-priority public safety users. To cope with this issue, the UAV-RIS can facilitate a direct or multi-hop communication scenario [14]. The UAV-RIS concept with public safety networks aids in minimizing network congestion and coverage gaps. The outage probability of public safety users is significantly reduced by adding flexible UAV-RIS relays into the disaster regions as shown in Fig. 2d. Using RIS reflection angle and phase shifts, the UAV-RISs can simultaneously provide coverage to nearby public safety users and act as a relay to cover public safety users located at large distances.

UAV-RIS Aided IoT Communication

The Internet of Things (IoT) plays a significant role in the area where human beings are not inhabited. Therefore, IoT devices are deployed in regions such as forests, deserts, or over the water to monitor and collect surrounding information [15]. Additionally, humans may also wear many IoT devices, such as health monitoring sensors, on them or in their surroundings, hence, it is convenient to deploy the UAV-RIS concept to provide high-speed communication, larger bandwidth, minimum latency, and energy consumption to ensure the prevalence of the IoT.

The UAV-RIS concept can be used to support the real-time transfer of data from the distributed IoT devices using reflection angle and phase shifts of RIS as shown in Fig. 2e. Precisely, by deploying the UAV-RIS near the IoT devices with properly outlining beamforming policy, the data rate is improved with a reduced transmitting power of IoT devices for a particular data collection rate. The UAV-RIS concept is more likely to provide LoS links with the IoT devices due to their high elevation. Moreover, the UAV-RIS can change their locations to provide on-demand coverage to IoT devices and transmits collected data to base stations.

Furthermore, there may be sensors buried under the soil or in water leading to signal attenuation, especially when the foliage is high and when the seasonal crops grow. In such a case, the UAV-RIS can help in collecting the data periodically (e.g., every day or every week, depending on the delay sensitivity of the data), with the UAV acting similar to a data mule.

Machine Learning Approaches for Reliable Beamforming with UAV-RIS

In this section, we discuss machine learning-based beamforming approaches for the UAV-RIS to improve the system reliability and rates of various users, such as blocked users, public safety users, legitimate users (security purposes), vulnerable vehicles, and IoT sensors under practical dynamic channels. We also aim to ensure the QoS requirements of all users. In particular, we consider three different approaches that rely on RL, DQN, and PER DQN-based beamforming to optimize the reflection phases of RIS in time-varying environments.

RL-Based Beamforming

First, we consider a model-free RL approach in order to resolve the issues of decision-making (beamforming) in a dynamic, rapidly changing wireless environment. Hence, we treat this problem through RL, where the UAV-RIS is considered as a part of the *environment*, and the central controller located at the base station is treated as a learning agent. The state space includes channel information, transmission data rate, and QoS satisfaction level. The action space contains the beamforming vector (flatting and broadening) chosen by the central controller and the transition probability shows the probability of transitions into a new state. Moreover, the reward function estimates whether the beamforming policy is able to maintain a high data rate (considering the users' QoS satisfaction level) or not. When the reward function satisfies the aimed objectives, the overall system performance is improved and the beamforming policy is adopted at the UAV-RIS.

DQN-Based Beamforming

Q-learning is an effective way to generate an outcome for the learning agent so that the learning agent benefits from figuring out precisely which action to execute. In DQN, we employ a neural network to approximate the function of the Q-value. The state space is inserted as an input to the DQN, which then generates an output considering the Q-values of all possible actions. There are three steps involved in DQNs. In the first step, all the events from the users are stored in memory. Secondly, the system determines the future action by calculating the maximum output. In the final step, the loss function is found by calculating the mean square error of the predicted value (Q) and the target value (Q*). When we utilize the RL-based beamforming policy, the target or actual values are not considered. In DQN-based beamforming, in contrast, the system updates the equation of Q-value derived from the Bellman equation. The reward is the substantive outcome achieved in RL. Using the backpropagation for convergence from a reward function, the network updates its policy gradient and attains a beamforming strategy.

Drawbacks of RL and DQN based Beamforming

Despite the tremendous advantages of the RL, the RL suffers from a slow speed of convergence, and the RL is also not well tailored for the problem of UAV-RIS beamforming, which includes continuous state space. The policy gradient method converges to a suboptimal solution and it has the ability to deal with consecutive state-action space. Moreover, it is unmanageable for policy gradient algorithms and RL to determine the optimization problem under the input state space with high dimensions (large data). Contrary, DQN works well in learning policies with high-dimensional state space, but the non-linear Q-function possibly leads to an unpredictable learning process.

Despite the tremendous advantages of the RL, the RL suffers from a slow speed of convergence, and the RL is also not well tailored for the problem of UAV-RIS beamforming, which includes continuous state space.

PER DQN-Based Beamforming

Even though DQN is able to perform efficiently in learning with a high-dimensional state space, the deep neural network utilized in DQN may lead to divergence in the dataset causing correlations between input and output. For this purpose, we propose to exploit experience replay to avoid the unpredictable learning process of the RL algorithm. The experience replay is deployed as a circular buffer, where the oldest transition in the buffer is discarded to create space for recent transformations. In particular, the experience replay is implemented as a fixed-size buffer (Fig. 3) that stores current transitions gathered by the system. It enhances the efficiency of DQN by allowing data to be reused many times for training instead of discarding data. We optimize the beamforming and provide a strong LoS connection by using PER DQN. However, by placing eavesdroppers or jammers outside the relatively narrow radiation region of the RIS beam, a network can mitigate the risk of security attacks. The RISs are able to control the radio propagation environment and act as key enablers for improving the physical layer security of wireless communication systems in an economical and energy-efficient manner. Therefore, the optimized beamforming in the UAV-IRS helps to enhance the security of the physical layer in UAV-to-ground communication. It also improves the stability of the system during the training process.

All changes in DQN are sampled from the buffer at fixed intervals and utilized in the process of training. The commonly used sampling approach is PER, where new transitions are initialized to the ultimate priority values seen so far and these are only updated once they are sampled. The priority is updated after the loss function is received in the neural network. At the end of the training, the system avoids overfitting by considering constantly prioritized experiences.

Due to the highly dynamic and high-dimensional properties and the unpredictable channel state information of the UAV-RIS communication, we adopt a deep PER learning-based beamforming. The PER algorithm is employed to support the learning agent in adopting faster learning in dynamic wireless environments. In particular, the

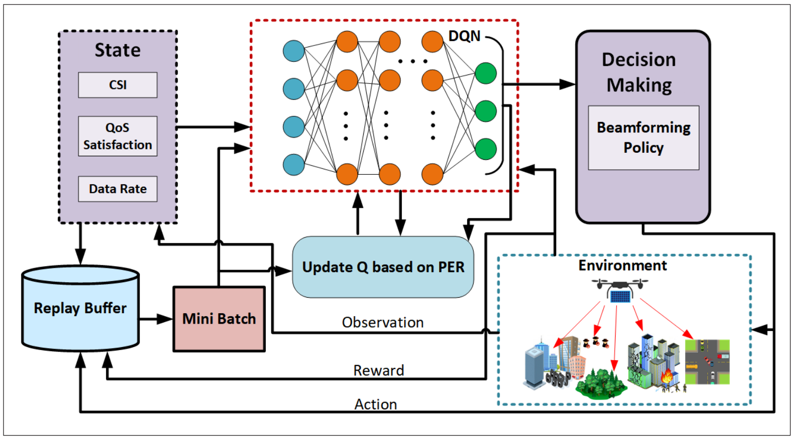

FIGURE 3. Deep PER-based beamforming for the UAV-RIS systems.

learning agent injects the already stored data (i.e., channel state information, transmission data rate, and QoS satisfaction level), the reward (feedback) obtained after interacting with the environment, updated Q based on PER, and the data from the replay buffer (historical experience) to the DQN for training the learning model. Consequently, the learning agent applies the trained model to obtain the decision (reflecting beamforming matrix) based on its learned strategy as shown in Fig. 3.

The PER DQN algorithm is completed in two phases: the training phase and the implementation phase. In the training phase, the central controller at the base station gathers surrounding information from the environment and makes a decision on the beamforming setting. The controller initializes system parameters and stores the current state including transmission data rate and channel state information. In addition, the state space acts as an input to DQN that trains the deep neural network model. The greedy scheme is adopted to adjust the exploitation and exploration (the actions having maximum reward are chosen, considering the current information). After that, when the selected action is performed, the agent gets a reward function from the surrounding environment. After sufficient training, the learning model initializes the implementation phase. During this phase, the RIS controller employs the trained model to determine its selected action space after passing through the DQN with the observed state-space from the UAV-RIS communication. More specifically, the RIS controller determines the action with the highest value based on the trained deep PER-based learning design.

Subsequently, the beamforming matrix (reflecting beamforming) in the UAV-RIS concept is set according to the chosen action. In the end, the environment again inputs an instantaneous reward and a new state to the learning agent. The comparison among machine learning-based RIS beamforming systems (RL, DQN, PER DQN) is discussed in Table 1.

Performance Comparison of UAV-RIS Beamforming Techniques

In this section, we compare the performance of the three different UAV-RIS beamforming techniques described in the previous sections using simulations. The RIS element count ranges from 10 to 90. The pathloss is modeled as $PL_0 - 10\phi\log10(d/d_0)$ (dB), where PL_0 = 30 dB is the path loss at the reference distance, ϕ is the path loss exponent and we set it as ϕ = 3, and d represents the distance between the transmitter and the receiver. The learning model consists of three hidden layers, each of which has 500, 250, and 200 neurons. As our data is of large dimensions or features, to get an optimum solution, we use 3 hidden layers. When we use a larger number of hidden layers (more than 3), the complexity of the model is increased and this leads to overfitting.

In Fig. 4a, we observe that the validation loss is slightly higher than the training loss and lower than the testing loss. It shows that weights of the deep neural networks provide an appropriate mapping between input and output samples and the effectiveness of the designed neural network weights. In the case of higher validation, training, and testing losses, the model suffers from underfitting, which would require altering the number of neurons.

The discount factor is set to 0.95, while the learning rate is set at 0.001. With an extremely high learning rate, a longer time to reach convergence would be required, and its reward performance is significantly lower than that of a learning rate of 0.001. Furthermore, if we use a too-low learning rate it will take a longer time to reach convergence. With a learning rate of 0.001, the model is capable of learning the problem as shown in Fig. 4a. Therefore, we choose the learning rate with a value of 0.001 in our tests. We focus on the long-term reward and we use a lower number of episodes. Hence, the discount factor value is set to 0.95. Over the first 100 episodes, the exploration rate is linearly annealed from 0.8 to 0.1 and, then, the exploration rate remains constant and the reliability requirement ranges from 99.1 percent to 99.9 percent. The threshold for QoS satisfaction is set to a minimum transmission data rate of 5 b/s/Hz. When the transmission data rate is lower than the 5 b/s/Hz, the service is considered as unsuccessful. The performance of the deep PER-based beamforming in the UAV-RIS communication network and the beamforming policy (broadening and flattening) is compared with the following approaches:

- The traditional RL-based beamforming scheme, where RL is employed to calculate the beamforming policy.
- The traditional DQN-based beamforming scheme, where a deep neural network is applied to evaluate the Q-value when utilizing the beamforming policy with the largest Q-value.
- The classical UAV-RIS communication, where no beamforming is adopted by UAV-RIS.

In Fig. 4b, the QoS satisfaction probability of the proposed scheme is examined and compared with RL and DQN-based RIS-beamforming and without RIS-beamforming. For all algorithms, with an increasing number of reflecting elements, the QoS satisfaction probability improves as well. The increase is because more reflecting features can provide additional signal paths and directionality. The proposed PER-DQN UAV-RIS beamforming outperforms notably all other approaches. For only 10 reflecting elements, the proposed PER-DQN reaches already 95 percent QoS satisfaction, while only 90 percent, 85 percent, and 78 percent are reached by DQN, RL, and without RIS beamforming approach-

Technology	Characters	Drawbacks
RL-based beamforming	• Agent precisely figures out the action to execute. • Policy gradient method converges to a suboptimal solution. • Ability to deal with continuous state-action spaces.	• Slow speed of convergence. • Problem in handling large state-action space.
DQN-based beamforming	• More accurate than RL. • Works well in learning with high-dimensional state space. • Able to handle high dimensional state-action space.	• Q-function (an estimator) possibly leads to an unpredictable learning process. • Sometimes causes the correlations between input and target values.
PER DQN-based beamforming	• Avoids the divergence caused in RL-based algorithms. • Improves the sample efficiency by allowing data to be reused many times for the process of training. • Enhances the system stability. • Improve data utilization.	• Difference between storage and real priority distributions sometimes causes bias into the gradients.

TABLE 1. Comparison of machine learning-based RIS beamforming.

es, respectively. At the same time, the proposed PER-DQN reaches 100 percent QoS satisfaction already for 50 reflecting elements, while 70 and 90 are required by DQN and RL-based approaches. This excellent performance of the proposed PER-DQN is due to the QoS-aware reward that is used in PER-DQN. Within a reasonable change in reliability thresholds, the machine learning-based scheme achieves a high successful transmission probability.

The PER-based beamforming policy outperforms the RL and DQN-based RIS-beamforming and the case without RIS-beamforming also in terms of the transmission probability under various reliability requirements as shown in Fig. 4c. As we get more strict for the reliability threshold, the successful transmission probability decreases rapidly for RL, DQN, and without RIS beamforming to 0.80, 0.70, and 0.45, respectively. In contrast, the proposed DQN-PER leads to only negligible degradation and the transmission probability remains at 0.92 even for 99.9 percent.

Research Challenges and Future Directions

To cover the future era of the "Internet of Drones with RIS" potential research directions are highlighted in this section.

Millimeter-Wave UAV-RIS Beamforming

Millimeterwave (mmWave) communication uses a wide bandwidth at frequencies of 28 GHz and above, and it can be coupled with UAV-RIS for high data rate communications. It is well known that mmWave signals suffer from high path loss, penetration loss, and blockages, and these aspects become even more apparent at sub-terahertz and terahertz frequencies likely to be supported by 6G networks. Generally, the LoS UAV-RIS and base station channels are favorable and practical in harsh channel conditions and can help tremendously in improving coverage at mmWave frequencies and beyond. Efficient mmWave beamforming is required to be developed due to the high mobility and altitude of UAVs in 3D mmWave UAV-RIS concept.

UAV-RIS Swarm Communications

Multiple UAVs equipped with RIS can be connected to each other to create a group of highly coordinated UAV-RISs to achieve a joint mission cooperatively. It is considered challenging for the UAV-RIS to link directly with the base station because of the large number of UAV-RIS. Alternatively, the base stations can provide connectivity between the UAV-RIS network and the core network. Cooperative communication is considered one of the potential strategies for providing assisted UAV-RIS to UAV-RIS communications. Moreover, the massive multiple input multiple output techniques for communications with multiple connected UAV-RIS are seen as leading research in the coming years. For the efficient UAV-RIS network topology and an impeccable integration of the UAV-RIS concept with the mobile networks, further investigations towards ultra-reliable and low latency communication are required.

Autonomous UAV-RIS Placement and Navigation

Machine learning techniques for the UAV-RIS adjusting locations, motion states, and trajectories while considering the dynamic changes in the environment like flight environments, control strategies, and/or any other bursting-out risks. In the autonomous UAV-RIS placement and navigation, it is required to model a framework based on double deep neural networks as this can help the UAVs-RIS to navigate and avoid obstacles successfully.

Conclusions

The UAV-RIS concept has emerged as a promising technology to improve the coverage and reliability of next-generation wireless networks. However, the UAV-RIS concept faces various challenges for reliable communication. In this article, we first present use-cases to deploy UAV-RIS beamforming in various practical scenarios. After that, we discuss promising technologies for the UAV-RIS concept based on machine learning techniques such as RL, DQN, and PER DQN-based beamforming policies. To improve the overall performance, we propose a beamforming policy by employing the PER-based DQN to enhance the learning performance. The simulation results show that our PER-based DQN approach outperforms other machine learning-based beamforming techniques that can be used with UAV-RIS.

Acknowledgment

This work was supported by the Ministry of Education, Youth and Sport of the Czech Republic under Grant No. LTT20004 and in part by the NSF award CNS-1910153.

References

[1] M. Jian and R. Liu, "Baseband Signal Processing for Terahertz: Waveform Design, Modulation and Coding," *Proc.*

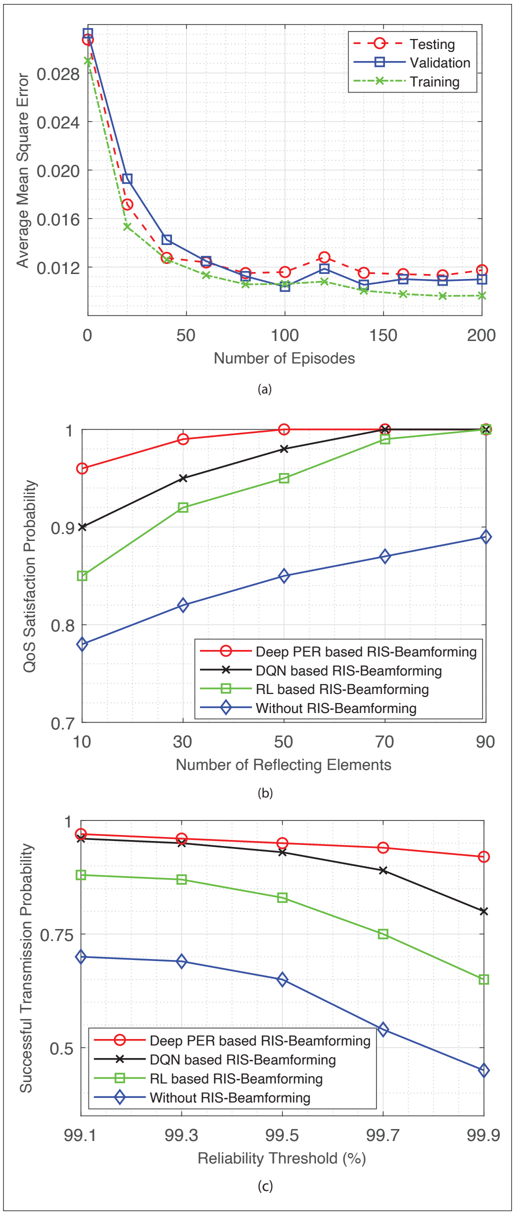

FIGURE 4. Performance of the proposed PER-based DQN approach for UAV-RIS beamforming in terms of: a) Average mean square error; b) QoS satisfaction probability; c) Successful transmission probability.

IEEE Int'l. Wireless Commun. and Mobile Computing Conf. (IWCMC), 2021, pp. 1710–15.

[2] Z. Huang and X. Cheng, "A 3-D Non-Stationary Model for Beyond 5G and 6G Vehicle-to-Vehicle mmWave Massive MIMO Channels," *IEEE Trans. Intelligent Transport. Syst.*, 2021, pp. 1–17.

[3] Q. Wu *et al.*, "Intelligent Reflecting Surface-Aided Wireless Communications: A Tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, 2021, pp. 3313–51.

[4] M. Mozaffari *et al.*, "A Tutorial on UAVs for Wireless Networks: Applications, Challenges, and Open Problems," *IEEE Commun. Surveys and Tutorials*, vol. 21, no. 3, 2019, pp. 2334–60.

[5] S. Li *et al.*, "Reconfigurable Intelligent Surface Assisted UAV Communication: Joint Trajectory Design and Passive Beamforming," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, 2020, pp. 716–20.

[6] C. Huang *et al.*, "Holographic MIMO Surfaces for 6G Wireless Networks: Opportunities, Challenges, and Trends," *IEEE Wireless Commun.*, vol. 27, no. 5, 2020, pp. 118–25.

[7] X. Cao *et al.*, "AI-Assisted MAC for Reconfigurable Intelligent-Surface-Aided Wireless Networks: Challenges and Opportunities," *IEEE Commun. Mag.*, vol. 59, no. 6, 2021, pp. 21–27.

[8] Y. Cheng et al., "RIS-Aided Wireless Communications: Extra Degrees of Freedom via Rotation and Location Optimization," *IEEE Trans. on Wireless Commun.*, 2022, pp. 1–1.

[9] S. K. Moorthy *et al.*, "Flybeam: Echo State Learning for Joint Flight and Beamforming Control in Wireless UAV Networks," *Proc. IEEE Int'l. Conf. Commun. (ICC)*, 2021, pp. 1–6.

[10] M. Banagar, H. S. Dhillon, and A. F. Molisch, "Impact of UAV Wobbling on the Air-to-Ground Wireless Channel," *IEEE Trans. Veh. Tech.*, vol. 69, no. 11, 2020, pp. 14 025–30.

[11] F. B. Mismar, B. L. Evans, and A. Alkhateeb, "Deep Reinforcement Learning for 5G Networks: Joint Beamforming, Power Control, and Interference Coordination," *IEEE Trans. Commun.*, vol. 68, no. 3, 2020, pp. 1581–92.

[12] S. Alfattani *et al.*, "Aerial Platforms With Reconfigurable Smart Surfaces for 5G and Beyond," *IEEE Commun. Mag.*, vol. 59, no. 1, 2021, pp. 96–102.

[13] X. Cao *et al.*, "Reconfigurable Intelligent Surface-Assisted Aerial-Terrestrial Communications via Multi-Task Learning," *IEEE JSAC*, vol. 30, no. 10, 2021, pp. 3035–50.

[14] P. Vamvakas, E. E. Tsiropoulou, and S. Papavassiliou, "On the Prospect of UAV-Assisted Communications Paradigm in Public Safety Networks," *Proc. IEEE INFOCOM Workshops*, 2019, pp. 762–67.

[15] S. Lhazmir *et al.*, "Performance Analysis of UAV-Assisted Ferrying for the Internet of things," *Proc. IEEE Symposium on Computers and Commun. (ISCC)*, 2019, pp. 1–6.

Biographies

Ishtiaq Ahmad received a B.S. degree in electrical engineering from the University of Engineering and Technology, Peshawar, Pakistan, in 2007, and the M.S. and Ph.D. degrees in electronic engineering from Inha University, South Korea, in 2014 and 2019, respectively. Since 2009, he has been a Lecturer with the Faculty of Engineering and Technology, Gomal University, Pakistan. His research interests include interference management in 3GPP LTE-A and 5G/6G systems, public safety and mobile ad-hoc networks (especially for UAV), cellular-V2X technology, maritime and underwater communications, and applications of artificial intelligence technologies. He was a recipient of the Jungseok International Scholarship for his M.S. and Ph.D. degrees at Inha University due to his excellent academic career. He received the Outstanding Research Award and the Excellent Student Award from Inha University, in 2019, for his excellence in journal publication and outstanding research achievements.

Ramsha Narmeen received a B.S. degree in computer system engineering from the University of Engineering and Technology, Peshawar, Pakistan, in 2015, and a master's degree in electronic engineering from Inha University, Korea. She has authored several international journals and IEEE Conf. papers. Her research interest includes underwater acoustic communication, interference management 3GPP LTE-A and 5G systems, cellular-V2X technology, mobile ad-hoc networks (especially for UAV), and applications of artificial intelligence technologies. She was a recipient of the Jungseok International Scholarship for her M.S. degree at Inha University due to her excellent academic career.

Zdenek Becvar received M.Sc. and Ph.D. in Telecommunication Engineering from the Czech Technical University (CTU) in Prague, Czech Republic in 2005 and 2010, respectively. He is currently an Associate Professor with the Department of Telecommunication Engineering, Faculty of Electrical Engineering, CTU in Prague. From 2006 to 2007, he was with Sitronics R and D Center in Prague, focusing on speech quality in VoIP. In 2009, he was involved in research activities at Vodafone R&D Center, CTU in Prague. He was on internships with Budapest Polytechnic, Budapest, Hungary, in 2007, CEA-Leti, France, in 2013, and EURECOM, France, in 2016 and 2019. From 2013 to 2017, he was a representative of CTU in Prague in ETSI and 3GPP standardization organizations. He leads the laboratory of mobile networks - 6Gmobile Research Lab targeting radio resource and mobility management in future mobile networks, edge computing, C-RAN, and UAVs/vehicular/IoT communications.

Ismail Guvenc is a Professor at the Department of Electrical and Computer Engineering at NC State University. His recent research interests include 5G/6G wireless networks, UAV communications, millimeter/terahertz communications, and heterogeneous networks. He has published more than 300 conference and journal papers and book chapters, several standardization contributions, four books, and over 30 U.S. patents. He is the PI and the director for the NSF AERPAW project and a site director for the NSF BWAC I/UCRC center. He is an IEEE Fellow, a senior member of the National Academy of Inventors, and recipient of several awards, including the NC State University Faculty Scholar Award (2021), R. Ray Bennett Faculty Fellow Award (2019), FIU COE Faculty Research Award (2016), NSF CAREER Award (2015), Ralph E. Powe Junior Faculty Award (2014), and USF Outstanding Dissertation Award (2006).