

# ARTIFICIAL INTELLIGENCE AIDED JOINT BIT RATE SELECTION AND RADIO RESOURCE ALLOCATION FOR ADAPTIVE VIDEO STREAMING OVER F-RANS

Jienan Chen, Zhongxiang Wei, Shuai Li, and Bin Cao

## ABSTRACT

Recently, fog-computing-based radio access networks (F-RANs) have been conceptualized to provide high quality of experience (QoE) for adaptive bit rate (ABR) streaming, where additional computing capacity is supplemented on fog nodes to facilitate complicated cross-layer optimization (i.e., joint bit rate selection and radio resource allocation). However, finding an optimal global solution with acceptable complexity is still infeasible by the conventional optimization methods. In this work, we propose an artificial intelligence (AI) aided joint bit rate selection and radio resource allocation scheme referred to as iABR, which provides a new vision for handling the over-complicated optimization in F-RANs. Based on multi-agent hierarchy deep reinforcement learning, the proposed iABR can dynamically allocate radio resource and select bit rate in a multiuser scenario, by perceiving the network environment and clients' player information. Moreover, long short-term memory (LSTM) is employed by the iABR algorithm, which enables accurate prediction of the change of channel quality by learning the history of the wireless channel. Hence, iABR is able to adjust the action policy in advance to accommodate the future channel quality for avoiding bit rate fluctuation. According to the experimental results, the iABR exhibits higher QoE in terms of high average bit rate, low rebuffering ratio, and average bit rate variance. Last but not least, the QoE performance of all the clients are fairly guaranteed by the iABR algorithm, enhancing the practicality of AI-driven F-RANs for multimedia service delivery.

## INTRODUCTION

Forecasts indicate that by 2020, multimedia will require 90 percent of the global IP data traffic, where video delivery will account for 50 percent of the overall data stream via wireless communications. To satisfy the rapidly growing demand, the fog-computing-based radio access network (F-RAN) [1] has emerged as a key technology for delivering video media services, where storage and computation tasks can be offloaded from a centralized cloud center to fog nodes. Also, the fronthaul load is released and response time is reduced by caching content at the edge. There-

fore, the F-RAN is envisioned as a critical enabler to support high transmission rate and low latency for mobile video service [2].

To provide a high quality of experience (QoE) for video stream services, high video bit rate, low rebuffering ratio, and low bit rate variance are the most critical metrics. However, the three inter-related metrics impose a great challenge for satisfying the stringent QoE requirement of video service, especially in wireless communications [7]. For instance, a high video bit rate is preferred to improve the QoE at the risk of a high rebuffering ratio. In the rebuffering process, video watching at the client side has to be paused until the buffer is reloaded, which brings degraded QoE for the clients. On the other hand, frequently adjusting the video quality may avoid the problem of rebuffering, while causing strong fluctuation in bit rate and video quality. The most popular tool for managing video streaming is adaptive bit rate (ABR), such as the widely used Media Player Classic (MPC) [4], and the related work is summarized in Table 1. To be specific, the ABR algorithms adaptively choose a bit rate for each video chunk, where the decision is based on the video content and network environments, such as the network throughput, playback buffer occupancy, and video chunk size [8]. For the conventional cloud-based network, the controller of the ABR is deployed on the client side (video player), and only selects a proper video bit rate without the ability of adaptive resource allocation [9]. In contrast, an F-RAN provides an additional action dimension, namely radio resource allocation on the physical layer, for ABR streaming service, which is radio resource allocation on the physical layer [1]. Benefiting from the dedicated infrastructure, the fog-server (F-Server) is enabled to dynamically allocate radio resources, such as bandwidth, to improve the QoE of video service for each client. To this end, the cross-layer optimization-based ABR — joint bit selection at the application layer and radio resource allocation at the physical layer — has attracted much attention. A joint cache and radio resource allocation design was proposed in [5, 10], whereas bit rate selection was ignored. Tian *et al.* [10] proposed a joint radio communication, caching, and computing design for mobile virtual reality delivery in F-RANs, while only trying to maximize the average

Work	Pensieve [3]	Robust MPC [4]	JCRA [5]	DQA [6]	Proposed iABR
Actions	Bit rate only	Bit rate only	Cache and bit rate	Radio resources and bit rate	Radio resources and bit rate
Algorithm	Deep reinforcement learning	Classic MPC algorithm	Stackelberg game-based optimization	Time-step-based Lyapunov optimization	Multi-agent hierarchy deep reinforcement learning
Advantages	AI-based environment perceiving throughput prediction for bit rate selection	Throughput prediction for the next chunks based on past five chunks	Improved cache hit ratio	Joint resource allocation and bit rate selection	AI aided scheme joint resource allocation and bit rate selection High QoE in multi-client scenario
Limitations	Without resource allocation	Without resource allocation	Without bit rate selection	No channel prediction local optimum solution	Relatively high training complexity

TABLE 1. Computation performance comparison of different search methods.

delay tolerance instead of QoE. Most relevant to the approaches considered in this article, joint bit rate selection and radio resource allocation were designed in [6]. However, by finding the maximal QoE for the multiple time slices, the employed greedy-like searching may lead to a local optimum solution. In summary, the aforementioned algorithms either fail to optimize bit rate selection and radio resource allocation jointly, or only lead to local optimum solutions, for the sake of tractability of the optimization problems. Moreover, considering the variance of wireless channel quality, guaranteeing high QoE in a multi-client scenario becomes more difficult, where each client requests a large amount of exclusive resource to get a high QoE service under the constraint of the limited radio resource.

Recently, the booming of artificial intelligence (AI) provides a new vision for network design and optimization. AI-based deep reinforcement learning (DRL) exhibits outstanding capabilities on perceiving the network environment and making the decision for resource management [11]. Since the F-RAN provides additional ability in algorithm computing, AI algorithms can be efficiently executed at the F-RAN to alleviate the computation burden at the clients [12]. As a result, it is believed that AI can be integrated with F-RAN systems for handling complicated optimization problems.

To the best of our knowledge, this is the first work that applies an AI-based DRL algorithm to solve the joint bit rate selection and radio resource allocation optimization for ABR streaming service in a multi-client scenario referred to as iABR. The contributions of this work are summarized as follows:

- We propose a two-stage hierarchy AI-based DRL to jointly optimize the bit rate selection and radio resource allocation for ABR streaming service. By training the modules of the AI design (i.e., actor network, critic network, long short-term memory [LSTM] channel quality predictor, and environment simulator), the original over-complicated joint design problem can be effectively handled with acceptable complexity, and a near-optimal joint bit rate selection and radio resource allocation policy is readily endorsed. Hence, higher average bit rate, lower rebuffering ratio, lower bit rate variance, and superior overall QoE performance are demonstrated compared to the existing methods.
- A more practical multi-client scenario is addressed, where a multi-agent-based coop-

erative DRL algorithm is proposed for simultaneously delivering high QoE video services for all the clients. Each actor executes a specific action according to its local information, while the critic network evaluates the overall action according to the collected global results from all the actors, and then guides the behavior of each actor in a cooperative manner. As a result, the algorithm achieves a balanced QoE for all the clients, enhancing the practicality of delivering video over the F-RAN structure.

- To track the rapid change of wireless channel and seek the ability of instantaneous manner resource allocation, LSTM is employed to learn from the history of wireless channel quality and predict the future gain of wireless channel quality. Benefiting from LSTM prediction, the predicted channel quality well matches the real channel [13]. Hence, iABR can adjust the action policy in advance to accommodate the future of wireless channel quality, which facilitates an instantaneous manner of resource allocation for each client.

## SYSTEM MODEL

The considered F-RAN system is illustrated in Fig. 1, and consists of multiple mobile video players, a fog server (F-server), a fog radio access point (F-RAP), and a cloud center [14]. When mobile clients access the F-RAP and require video stream service, the F-server will send the corresponding video data to the F-RAP. If the required video is available in the client connected F-server, it will check if the neighborhood F-servers have the required video and access the content from the neighborhood F-server. At last, if the video content is unavailable in the F-server layer, the F-server first downloads the required video stream from the cloud center. Since the video content has a large probability of being stored in the F-server layer, and the achievable rate of optical fiber channel between F-servers is much higher than that of the wireless channel (between F-RAP and clients), it is assumed that the required video stream can always be cached on the F-server for problem simplification.

As illustrated by Fig. 2, the joint radio resource allocation and bit rate selection design is considered in this work for multi-client-based F-RAN. In particular, the video stream is divided into several video chunks, where each chunk corresponds to around 1–4 s playtime. The mobile clients store video chunks on a local buffer and output the data stream for video playback after receiving

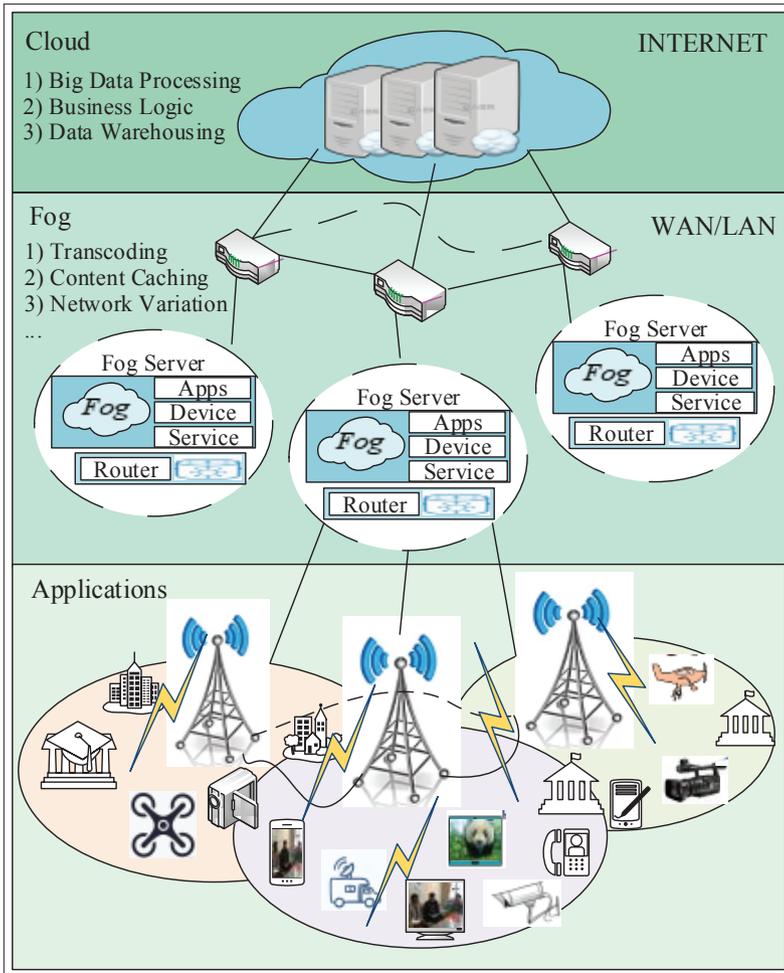


FIGURE 1. The system architecture of fog-computing-based radio access networks.

a certain number of video chunks. If the chunk downloading speed is slower than the speed of video playback, the buffer may become empty after a while, which leads to the rebuffering process.

The joint bit rate selection and radio resource allocation for the ABR streaming service are formulated as a Markov decision process (MDP). We first introduce the concept used in the MDP problem (i.e., state, action, and reward), and then the work flow of the whole MDP is demonstrated.

- *State*. The state at time stamp  $t$ , written as  $s_t$ , consists of instantaneous wireless channel quality  $q_k$ , buffer status  $p_k$ , video quality chunk size  $\chi_k$ , and untransmitted chunk  $\zeta_k$ . Hence, the state at time step  $t$  is expressed as  $s_t \leftarrow \{(q_1, p_1, \chi_1, \zeta_1), \dots, (q_K, p_K, \chi_K, \zeta_K)\}$ .
- *Action*. The action at time step  $t$ , written as  $a_t$ , instructs the policy of bit rate selection  $b_k$  and resource allocation  $c_k$  for clients  $k = \{1, 2, \dots, K\}$  as  $a_t \leftarrow \{(b_1, c_1), \dots, (b_K, c_K)\}$ . To be specific, the action of resource allocation  $c_k$  for client  $k$  (bandwidth) decides the communication rate of the client  $k$  as  $R_k = c_k \times \log_2(1 + q_k)$  with a constraint  $\sum_K c_k \leq B$ , where  $B$  denotes total bandwidth and  $q_k$  is channel quality of client  $k$ . Hence, together with the action of bit rate selection  $b_k$ , the buffer length of the client  $k$  is calculated as  $p_k \leftarrow p_k - (b_k - R_k) \cdot \Delta t$ . That is, if the transmission rate is lower than the video

playback rate ( $R_k < b_k$ ), the buffer becomes empty after a certain time and we have  $p_k \leq 0$ . In that case, the video player has to be paused for rebuffering, leading to a degraded QoE.

- *Reward*. The reward at time step  $t$ , written as  $r_t$ , is used for guiding and correcting the action policy, and hence leads to an improved QoE for ABR streaming service. The well-known reward value [11] is employed to evaluate the QoE, written as

$$r_t = \frac{1}{K} \sum \left( \begin{array}{l} \alpha \times \text{average\_bitrate} - \beta \times \text{rebuffering\_ratio} \\ - \lambda \times \text{quality\_variance} \end{array} \right).$$

In particular, the first item denotes the average bit rate by an expectation function

$$\frac{1}{T} \sum_{t=T}^t b_k(t)$$

over a time period  $T$ . The second item represents the rebuffering ratio, which gives a penalty on QoE performance if the rebuffering is required. The third item denotes the average bit rate variation, which tracks the magnitude of the changes in the quality form as

$$\frac{1}{T} \sum_{t=T}^t |b_k - b_{k-1}|.$$

The contribution or the optimization priority of each item is weighted by the weight factors  $\alpha$ ,  $\beta$ , and  $\lambda$ , where  $\sum(\alpha + \beta + \lambda) = 1$ , ( $\alpha, \beta, \lambda \in [0,1]$ ).

At each time step  $t$ , the agent observes the overall state  $s_t$  to choose an action  $a_t$ , and then a reward  $r_t$  is generated with the action  $a_t$ . Afterward, the state  $s_t$  at time step  $t$  is updated from  $s_t$  to  $s_{t+1}$  for the operation at the next time step  $t + 1$ .

## iABR: MULTI-AGENT HIERARCHY LEARNING ABR ALGORITHMS

In the iABR, we employ a multi-agent actor-critic DRL for the ABR streaming service in a multi-client scenario, which consists of multiple actor networks, a critic network, an environment simulator, and multiple LSTM channel predictors, as shown in Fig. 3. Now, we first introduce the functionalities of each module, and then the flow of the whole algorithm is demonstrated.

**Actor Network:** The actor network aims to generate the decision of resource allocation and bit rate selection by extracting the feature of the network environment and video play status at the client side. The actor network consists of a multi-agent network with a two-stage hierarchy deep neural network (DNN) and  $K$  agents. In the two-stage hierarchy DNN, the first DNN is employed to make a decision for resource allocation, and the second DNN is responsible for selecting bit rate. To be specific, the action of resource allocation exported from the first DNN is used as the input of the second stage DNN, and the second DNN outputs the set  $(b_k, c_k)$  as joint resource allocation and bit rate selection for client  $k$ ,  $k \in K$ . After gathering all actions from  $K$  clients, the action results  $\{(b_1, c_1), \dots, (b_K, c_K)\}$  are imported to the environment simulator to update the state (i.e., resource allocation of each client and client video playing status) and calculate the reward

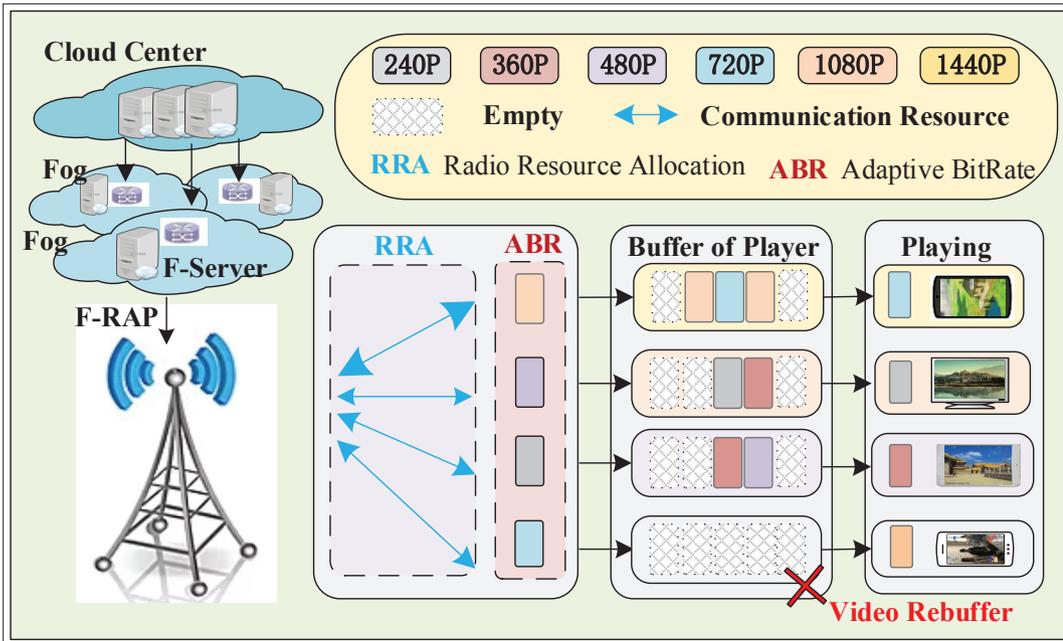


FIGURE 2. The joint radio resource allocation and bit rate selection design for multiple clients over F-RAN.

$r_t$ . In this design, we employ temporal-difference (TD) learning to learn from experiences with a TD error function for each agent. The TD error,  $TD_{loss}$ , is represented as  $TD_{loss} = reward_{env} - reward_c$ , where  $reward_{env}$  and  $reward_c$  are the reward value evaluated from the environment simulator and critic network, respectively. A higher TD error indicates that the probability of selecting action  $a_t$  is high in the future, and vice versa. Thus, we can get feedback on the TD error to each agent actor network to guide network parameter updating. The first stage actor network has five layers with six inputs, including current channel quality, predicted channel quality, video chunk information, untransmitted chunk size, and buffer status. The number of neurons in each DNN layer is [6, 128, 64, 32, 1]. The hidden layers have a tanh activation function, while the output layer has a softmax function. The number of neurons for the second stage network in each layer is [7, 256, 128, 64, 32, 1], where the output is the bit rate selection probability.

**Critic Network:** As mentioned, each actor network and an agent is an independent reinforcement learner and hence learns their action policy independently. However, in a multi-client scenario, the learning process of all the clients needs to be effectively cooperative to obtain desirable reward value for all the clients. To this end, we introduce a global reward value that is evaluated by a common critic network. In particular, the actions from all the clients, together with the state information, are used as the input of the critic network to calculate the expected reward value. The goal of a critic network is to calculate the reward value of the instantaneous state of the network environment and video playing status and action of the actor network. Physically, the critic network acts as an evaluator to estimate the performance of the current actor policy. We employ multiple-input multiple-output DNN for evaluating the multiple reward values for  $K$  clients. The state information and the action of

the clients generated from the actor network are concatenated together and imported as the input of the critic network. With the expected reward value  $reward_c$  calculated by the critic network, the evolution of the instantaneous state and action are exported to indicate the advantage of the current policy by TD loss. The input data of critic network is a matrix with dimensions  $K \times I$ , where  $K$  is the maximal client number and  $I$  is the number of inputs for each client. In this design, each client includes nine inputs. Apart from the same six inputs of state information in the actor network, it also includes action for radio resource allocation and bit rate selection, and an indicator to denote the current active video playing clients. The critic network consists of a three-stage neural network. The first stage is a group of one-dimension convolution neural networks (CNNs) with 512 channels and nine depth. After the CNN stage, a four-hidden-layers-based DNN with number of neurons [512, 256, 128, 64] is connected to the CNN stage. The third stage consists of  $K$  sub-layer DNNs with dimension [64, 16, 1]. Hence, the total output size of a critic network is  $K$ .

**LSTM Channel Quality Predictor:** To fully realize the functionalities of the actor and critic networks for joint bit rate selection and radio resource allocation design, the acquisition of instantaneous wireless channel quality is of importance. One possible solution is frequently exchanging pilot signals for channel estimation, which causes significant overhead and may not be suitable for the AI-driven F-RAN systems. Hence, we employ an LSTM-based neural network to predict the change of wireless channel quality, which enables an accurate prediction of the real wireless channel between the F-RAN and multiple clients. LSTM is formulated from recurrent neural networks with improved capacities for storing and retrieving channel information over a long period, and hence the history of wireless channel quality can be analyzed by the LSTM predictor. In other words, we import the history of wireless

Each actor network and an agent is an independent reinforcement learner and hence learns their action policy independently. However, in a multi-client scenario, the learning process of all the clients needs to be effectively cooperative to obtain desirable reward value for all the clients. To this end, we introduce a global reward value that is evaluated by a common critic network.

The actor networks generate the actions of resource allocation and bitrate selection for each client with the input of the environment state. Since each actor network or agent independent reinforcement learners learn their policy independently, and a global reward value is evaluated by a shared critic network cooperating with other agents.

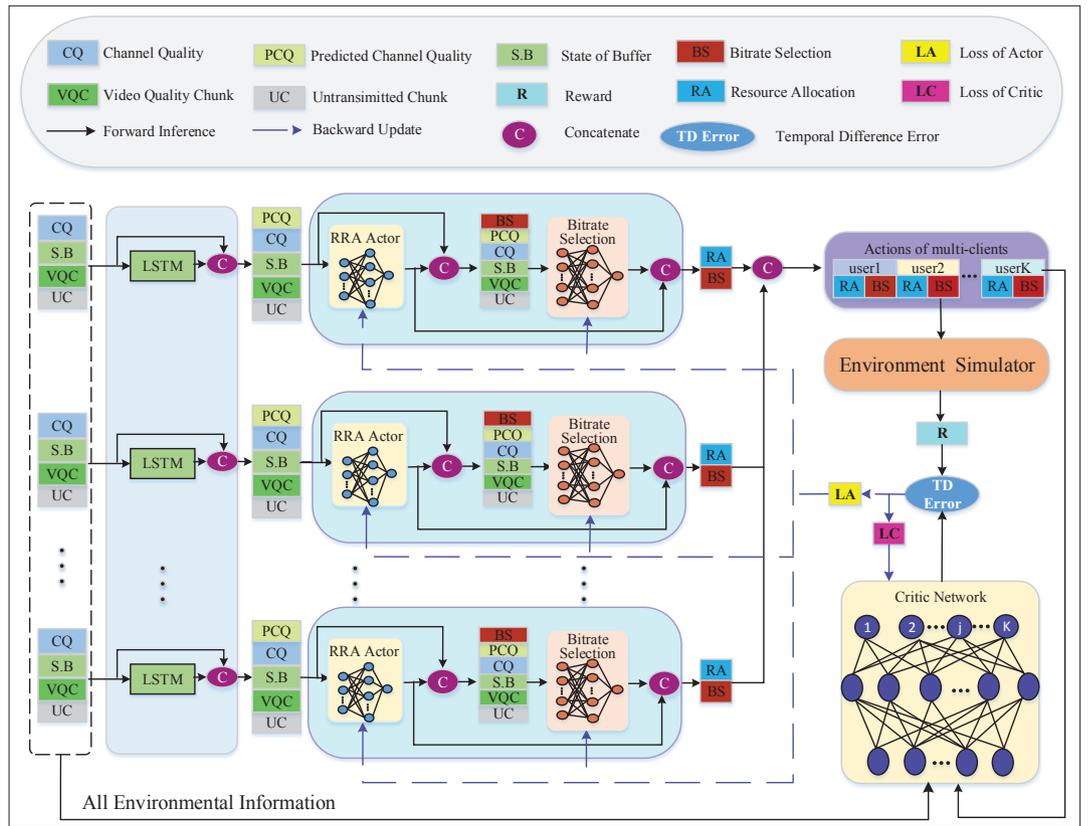


FIGURE 3. The multi-agent hierarchy-based A3C network scheme for joint bit rate selection and radio resource allocation. The actor network is responsible for resource allocation and bit rate selection; the critic network is used to evaluate reward value; LSTM predicts the change of wireless channel; and the environment simulator is employed to simulate video service for training purposes.

channel quality to LSTM. After analyzing the historical channel quality changing the trend, the LSTM predicts the future of wireless channel quality and exports it to the actor networks, where the action networks are connected sequentially with the LSTM. The LSTM network consists of a cell, an input gate, an output gate, and a forget gate. We collect channel quality values in the past 50 time steps as the input of the LSTM and output the channel quality prediction for the subsequent 10 time steps. If the time steps of LSTM are too long, the prediction will be more stable but may have a convergence problem. On the other hand, LSTM with short steps can be converged quickly, but with poor prediction performance. To achieve a balance point, we set the input time steps as 50 in our design. The hidden layer in each gate is of size [256, 128, 64, 10].

**Environment Simulator:** The environment simulator is built based on the framework of LTE network platforms such as openAirInterface [15] with Python to guarantee the fast data exchange between the simulator and actor-critic network. At the initial stage, since each neural network is untrained, the iABR may generate random action with a low reward value, which leads to a low QoE video service for the clients. To improve action accuracy, a training phase is needed to train each network by exploring the video streaming environment. Ideally, the perfect training phase can be performed on a real wireless radio network with an actual video streaming client. However, the emulation of such a real environment is time-consuming. It is because the train-

ing algorithm requires a completely downloaded video to obtain a reward. To address this issue, we build a simulation platform to evaluate the real wireless network and client applications. The environment simulator contains two layers, the physical layer and application layer. The physical layer updates the wireless channel quality for each client and calculates the transmission rate according to the channel quality and the allocated radio resource. On the other hand, the application layer updates the buffer state according to the transmission rate from the physical layer and the instantaneous video playback bit rate. The simulator keeps tracking average bit rate, rebuffering events, and bit rate variance to evaluate the QoE for critic network.

The actor networks generate the actions of resource allocation and bit rate selection for each client with the input of the environment state. Since each actor network or agent independent reinforcement learners learn their policy independently, a global reward value is evaluated by a shared critic network cooperating with other agents. To this end, the actions from all clients together with the state information are imported to the critic network to calculate the expected reward value, and a high reward value can be achieved for all of the clients. In the meantime, the actions are fed into the environment simulator to update the state and obtain a real reward  $r_t$  for all clients. The reward value from the environment simulator and critic network are used to calculate the loss value for critical network and actor network training. Since the reward value consists of

bit rate, rebuffering ratio, and bit rate variance, by maximizing the expected cumulative discounted reward

$$\mathbb{E} \left[ \sum_{t=0}^{\infty} \gamma^t \cdot r_t \right]$$

with a discount factor  $\gamma$ , it can guide the iABR algorithm to strike a good trade-off among the three metrics, and hence leads to improved QoE performance.

## TRAINING METHOD

In this section, we describe the training methodology of the iABR. In the actor-critic training algorithm, both the actor and critic networks' parameters are updated sequentially and simultaneously. To be specific, the parameters of the actor network are updated toward the gradient direction indicated by the evaluated reward value from the critic network. In the meantime, the critic network parameters are updated by the reward value from the perceiving environment. In each step, the DRL agent generates an action according to the current actor network policy and observes the environment state and the instantaneous reward. The critic network then estimates value functions with the compatible function approximation and calculates TD errors, which are imported to the actor network for computing the advantage function and estimating the policy gradient. With the gradient, the actor network update its parameters toward the gradient. The optimum parameters of both the actor part and critic part would be found when the loss function converges to a low loss value. In particular, we store the historical experience in an experience pool and calculate the discounted reward return as an advantage. The training for the critic network is performed with the loss function to minimize the  $TD_{loss}$  or value loss, defined as  $Loss_C = \sum (TD_{loss})^2$ . The parameter used for updating the actor network is trained to minimize the policy loss as  $Loss_P = -\log(\pi) \times TD_{loss} - \varepsilon \times H(\pi)$ , where  $H(\pi)$  measures the action entropy to encourage the action exploration. Here,  $\pi$  denotes the policy of action and  $\varepsilon$  is a normalization factor to adjust the contribution ratio of  $H(\pi)$ .

For traditional DRL, there is only one single agent with one neural network interacting with a single environment simulation per iteration, which leads to a low convergence speed. To accelerate the training speed, asynchronous advantage actor critic (A3C) reinforcement learning is employed, which consists of a global network and multiple parallel worker agents with a set of network parameters. Each worker agent consists of an independent critic and actor network, which interacts with the environment in a parallel form to achieve higher training efficiency. At the beginning of the training iteration, the parameters of each worker agent network are set to be the same as those of the global network. Then each worker agent interacts with the environment with its own network parameters and calculates the gradients according to the loss function. After that, each work agent updates its own network parameter with the calculated gradient and uploads the gradient to the global network. The parameters of the global network are updated by the collected gradients from all worker agents, and then

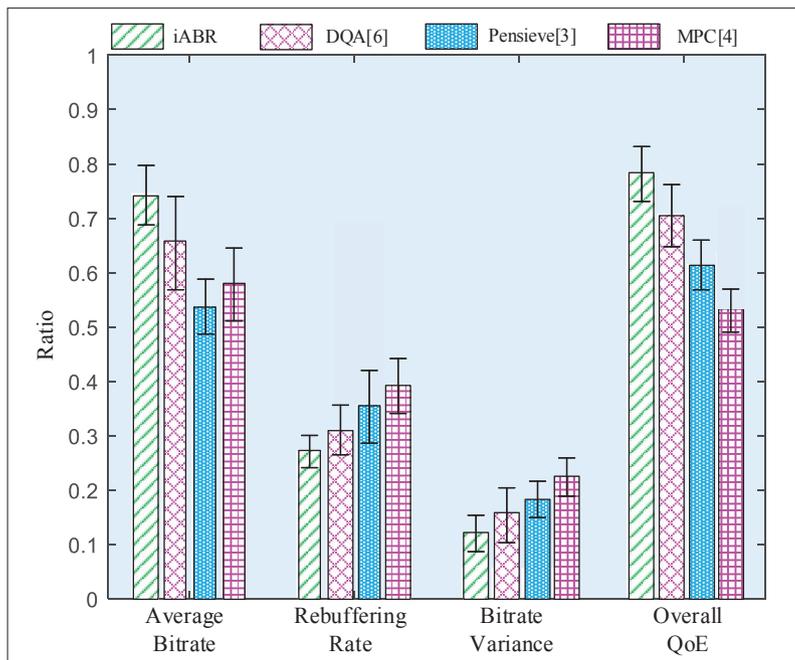


FIGURE 4. The performance comparison on the average bitrate, rebuffering ratio, bitrate variance and overall QoE.

the global network synchronizes the global network's parameter to all worker agents as a new iteration. The iteration process is performed until the network is converged (i.e., the loss function is reduced below a given level).

On the other hand, the training of LSTM is performed in a supervised learning manner by feeding training channel quality and the predicted quality label from the environment simulator.

## PERFORMANCE EVALUATION

In this section, we evaluate the QoE performance of iABR and its ability on channel prediction. Also, action distribution is analyzed and demonstrated. The experiment is built on an F-RAN environment, where the transmission delay between F-RAP and mobile clients is 10 ms, and the delay between F-server and F-RAP is 1 ms [14]. Without loss of generality, we assume there are four clients simultaneously requesting the video service. The total bandwidth of the wireless channel is 20 MHz. The video is encoded by the H.264/MPEG-4 codec at resolution ratio {240, 360, 480, 720, 1080, 1440}. We divide the video into 32 chunks with a total length of 96 s, where each chunk represents 3 s of video playback. The action execution time for the iABR is about 48 ms with GPU1080Ti and CPU platform.

The following schemes are selected as benchmarks.

- Pensieve [3], where only bit rate selection is involved
- Robust MPC [4], which is based on the classic MPC algorithm
- DQA [6], which jointly allocates radio resource and selects bit rate, and converts MDP to a time-sliced optimization problem by Lyapunov optimization

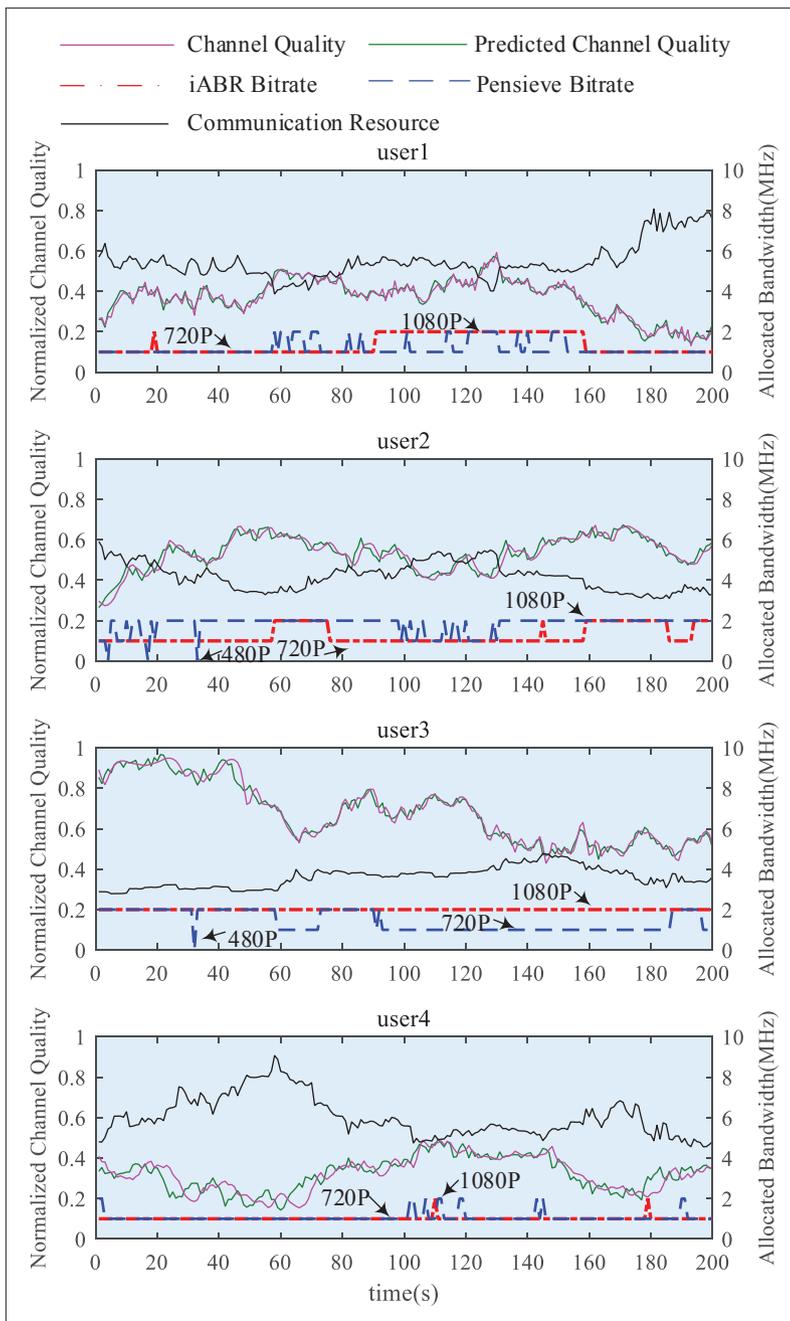


FIGURE 5. Based on the ability of predicting change of wireless channel (pink solid line), the iABR algorithm dynamically allocates resource (black solid line), and a stable bit rate (red dotted line) is guaranteed for all the clients.

### PERFORMANCE ANALYSIS ON AVERAGE BIT RATE, REBUFFERING RATIO, BIT RATE VARIANCE, AND OVERALL QOE

In Fig. 4, we evaluate the average bit rate, rebuffering ratio, bit rate variance, and overall QoE. First, as can be seen, the proposed iABR achieves higher average bit rate, lower rebuffering ratio, lower bit rate variance, and hence superior QoE performance. Compared to DQA, iABR demonstrates 13, 12, and 14 percent performance gain on the three metrics. This is because DQA simplifies the original MDP problem into a time-sliced optimization problem by Lyapunov optimization, and only a local optimum is achieved. Also, since only bit rate selection is considered by the pensieve, the pensieve presents 29 percent lower aver-

age bit rate, 19 percent higher rebuffering ratio, and 20 percent lower bit rate variance compared to our algorithm. Second, since bit rate, rebuffering ratio, and bit rate variance are the most important metrics affecting QoE of video service, it is observed that the proposed iABR algorithm readily achieves the highest QoE over the benchmarks, guaranteeing superior performance for ABR streaming service.

### PERFORMANCE ANALYSIS ON THE ABILITY OF CHANNEL PREDICTION

Figure 5 shows that the proposed algorithm presents a strong ability in predicting changes of wireless channels, and hence possesses the advantage of avoiding bit rate variance. We choose the closely related work, Pensieve [3], as the benchmark, which is also a DRL-based intelligent ARB with an equivalent amount of radio resource for each client. As can be seen, the proposed iABR exhibits a smooth bit rate (the red dotted line) for all the clients, where iABR has less bit rate variance. That is because the proposed iABR also considers the radio resource allocation to compensate for the changing of channel quality. Furthermore, iABR has a strong ability to predict the change of wireless channel by applying the LSTM channel predictor. As a result, the predicted channel quality (the solid green line) is well matched to the real wireless channel (the solid pink line), and hence the proposed iABR is able to dynamically allocate radio resource (solid black line) to guarantee a stable bit rate for all of the clients. On the other hand, an interesting observation shows that the allocated bandwidth of each client is directly determined by the instantaneous channel quality. That is, for a client with poor channel quality (solid pink line), higher bandwidth (solid black line) is allocated by the proposed algorithm to maintain a stable bit rate. This proves that in a multi-client scenario, the algorithm guarantees fairness among the clients in terms of the obtained radio resource and the QoE performance.

### CONCLUSIONS

In this work, an AI aided joint bit rate selection and radio resource allocation design is proposed for ABR streaming service in multi-client F-RAN systems. By learning from the network environment and client-side video playing status, the proposed iABR is capable of handling the complicated optimization problem: joint bit rate selection and radio resource allocation design in a multi-client scenario. By adaptively allocating radio resources and selecting preferred bit rate for multiple clients, the proposed iABR exhibits superior performance on average bit rate, rebuffering ratio, bit rate variance, and overall QoE compared to the existing methods. The success of applying AI for ABR streaming service motivates our further research on AI-driven F-RAN, such as intelligent cell handover for video delivery service and AI-aided joint computation and communication resource allocation.

### ACKNOWLEDGMENTS

This article is supported by the National Natural Science Foundation of China (Nos. 61971107, 61701059, and 61941114) and

## REFERENCES

- [1] D. You *et al.*, "Fog Computing as an Enabler for Immersive Media: Service Scenarios and Research Opportunities," *IEEE Access*, vol. 7, May 2019, pp. 65,797–65,810.
- [2] M. Peng *et al.*, "Fog-Computing-Based Radio Access Networks: Issues and Challenges," *IEEE Network*, vol. 30, no. 4, July 2016, pp. 46–53.
- [3] H. Mao, R. Netravali, and M. Alizadeh, "Neural Adaptive Video Streaming With Pensieve," *Proc. Conf. ACM Special Interest Group on Data Commun.*, Aug. 2017, pp. 197–210.
- [4] X. Yin *et al.*, "A Control-Theoretic Approach for Dynamic Adaptive Video Streaming Over HTTP," *SIGCOMM Comp. Commun. Rev.*, vol. 45, no. 4, Aug. 2015, pp. 325–38.
- [5] X. Xu, J. Liu, and X. Tao, "Mobile Edge Computing Enhanced Adaptive Bitrate Video Delivery with Joint Cache and Radio Resource Allocation," *IEEE Access*, vol. 5, Aug. 2017, pp. 16,406–15.
- [6] Y. Guo *et al.*, "Dynamic Quality Adaptation and Bandwidth Allocation for Adaptive Streaming Over Time-Varying Wireless Networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 12, Dec. 2017, pp. 8077–91.
- [7] D. Bethanabhotla, G. Caire, and M. J. Neely, "Adaptive Video Streaming for Wireless Networks with Multiple Users and Helpers," *IEEE Trans. Commun.*, vol. 63, no. 1, Jan. 2015, pp. 268–85.
- [8] Z. Akhtar *et al.*, "Oboe: Auto-Tuning Video ABR Algorithms to Network Conditions," *Proc. 2018 Conf. ACM Special Interest Group on Data Commun.*, Aug. 2018, pp. 44–58.
- [9] E. B. T. K. M. Schiller and R. Rothlisberger, *Improving Video Delivery with Fourier Analysis of Traffic in Multi-Access Edge Computing*, Springer, 2019, vol. 11618.
- [10] T. Dang and M. Peng, "Joint Radio Communication, Caching, and Computing Design for Mobile Virtual Reality Delivery in Fog Radio Access Networks," *IEEE JSAC*, vol. 37, no. 7, May 2019, pp. 1594–1607.
- [11] B. Cao *et al.*, "Intelligent Offloading in Multi-Access Edge Computing: A State-of-the-Art Review and Framework," *IEEE Commun. Mag.*, vol. 57, no. 3, Mar. 2019, pp. 56–62.
- [12] Y. Jiang *et al.*, "User Preference Learning-Based Edge Caching for Fog Radio Access Network," *IEEE Trans. Commun.*, vol. 67, no. 2, Feb. 2019, pp. 1268–83.
- [13] C. Luo *et al.*, "Channel State Information Prediction for 5G Wireless Communications: A Deep Learning Approach," *IEEE Trans. Network Science and Engineering*, vol. 1, no. 1, 2018.
- [14] M. Peng and K. Zhang, "Recent Advances in Fog Radio Access Networks: Performance Analysis and Radio Resource Allocation," *IEEE Access*, vol. 4, Aug. 2016, pp. 5003–09.
- [15] X. Foukas *et al.*, "Demo: Flexran: A Software-Defined RAN Platform," *Proc. 23rd Annual Int'l. Conf. Mobile Computing and Networking*, 2017, pp. 465–67.

## BIOGRAPHIES

JIANAN CHEN [S'10, M'14] received his Ph.D. degrees from the University of Electronic Science and Technology of China (UESTC) in 2014. He also worked at the University of Minnesota, Minneapolis, as a visiting scholar in 2012 and 2014, and then a postdoctoral scholar at the University of North Texas. He is currently an associate professor with the National Key Laboratory of Science and Technology on Communications at UESTC. He also served as Symposium Chair for Globalsip 2019 and TPC member for IEEE GLOBECOM 2019. His research interests are on machine-learning-based signal processing and artificial intelligence for networking.

SHUAI LI received his B.Sc. degree in electronic information engineering from Tianjing University in 2017. Currently, he is studying for a Master's degree in communication and information systems at UESTC. His research interests include communication algorithms, especially in 5G massive MIMO and application of deep learning in the communication field.

ZHONGXIANG WEI [M'17] (zhongxiang.wei@ucl.ac.uk) received his Ph.D. degree in electrical and electronics engineering from the University of Liverpool, United Kingdom, in 2017. From March 2016 to March 2017, he was with the Institution for Info-comm Research, Agency for Science, Technology, and Research (A\*STAR), Singapore, as a research assistant. He is currently a research associate in electrical and electronics engineering at University College London, United Kingdom. His research interests include constructive interference design, green communications, full-duplex, millimeter-wave communications, and algorithm design.

BIN CAO (bincao@uic.edu.cn) is an associate professor in the State Key Laboratory of Network and Switching Technology at Beijing University of Posts and Telecommunications. He received his Ph.D. degree (Honors) in communication and information systems from the National Key Laboratory of Science and Technology on Communications, UESTC in 2014. He also served as Symposium Co-Chair for IEEE ICNC 2018, Blockchain Workshop Co-Chair for CyberC 2019, and TPC member for numerous conferences. His research interests include blockchain systems, the Internet of Things, and mobile edge computing.