Coordinated Machine Learning for Handover in Mobile Networks with Transparent Relaying UAVs

Petr Skaba, Zdenek Becvar, Pavel Mach
Faculty of Electrical Engineering
Czech Technical University in Prague
Prague, Czech Republic
zdenek.becvar@fel.cvut.cz

Ismail Guvenc
Department of Electrical and Computer Engineering
North Carolina State University
Raleigh, North Carolina, USA
iguvenc@ncsu.edu

Abstract*—The Unmanned Aerial Vehicles (UAVs) acting as relays in the mobile networks are usually energy constrained. To improve the energy efficiency of such networks, UAVs should operate in a transparent relaying mode. In such mode, however, the channel quality between users and UAVs cannot be measured, since the transparent relays do not transmit own signaling. A lack of information on the quality of channel between users and UAVs limits practical implementation and is serious restraint for mobility management. To overcome this limitation, we develop a novel concept of coordinated machine learning for handover of users and UAVs in the mobile networks with transparent relaying UAVs. First, we predict the channel quality from other known information in the network via deep neural network (DNN). Such predicted channel quality is then fed into deep reinforcement learning (DRL) for an adjustment of handover parameter – cell individual offset (CIO). Unfortunately, a simple concatenation of the DNN and the DRL leads to a notable performance degradation. Hence, we propose a coordination of the DNN for channel quality prediction and the DRL for CIO setting. The coordination consists in a mutual exchange of performance-related information and an update of DNN according to a reward of DRL. The proposal increases the sum capacity by up to 12.7% while reducing the number of user and UAV handovers by up to 12.9% and 16.4%, respectively, compared to related works.

***Index Terms*—6G, channel quality, coordination, handover, machine learning, mobility management, UAV**

I. Introduction

The Unmanned Aerial Vehicles (UAVs) acting as relays are seen as a promising solution to provide connectivity for user equipments (UEs) in the future mobile networks due to a high dynamicity and flexibility of the UAVs' deployment. An integration of the UAVs to the mobile networks, however, imposes new challenges related to mobility management. The reason is that not only the UEs, but also the UAVs themselves, are inherently mobile. Thus, the UEs' connection to conventional ground base stations (GBSs) should be also managed jointly with the connection of the UEs to the UAVs or to the GBSs [1], [2]. Besides, the trajectory of the UAVs serving the UEs is arbitrary and difficult to predict, as the UAVs' position depends on an unpredictable movement of the UEs. Consequently, the quality of the channels between GBSs and UAVs as well as between UEs and GBS or UAV can change rapidly. This leads to frequent handovers or even to a handover failures [2].

This work was supported by the project 23-05646S funded by the Czech Science Foundation and by the NSF project CNS-1910153.

In a conventional terrestrial networks, the handover of UE is initiated if the GBS to which the handover should be performed (target GBS) offers a higher channel quality than a current serving GBS. However, such approach can result in frequent handovers and, consequently, to high signaling overhead, connection interruption, or excessive energy consumption. Hence, the handover is postponed until the received signal strength from the target GBS is by an offset and/or hysteresis above the signal strength from the serving GBS for a period of time known as time-to-trigger (TTT) [3], [4]. Setting of the handover parameters is optimized, e.g., in [5]–[7] to balance between a mitigation of redundant handovers and a sum capacity offered to the UEs for the networks with static GBSs. However, the movement of the UAVs leads to unpredictable changes in the network topology and the existing solutions for UEs' handover among GBSs are not able to cope with these changes [1].

Although the handover of UAVs is addressed, e.g., in [9]–[11], these works assume the UAVs in a role of the UEs with known future trajectory. Thus, the challenges related to the unknown future trajectory of the UAVs serving the UEs are not addressed in [9]–[11]. The handover for the UAVs serving the UEs is tackled in our prior work [1], where we develop a framework for a dynamic setting of a cell individual offset (CIO) for the GBSs and the UAVs to increase the sum capacity of the served UEs while avoiding redundant handovers. However, the UAVs are in non-transparent mode carrying out all communication-related control and management functions as the traditional GBSs. Such comprehensive management leads to a high complexity, weight, and energy consumption [12], making the *non-transparent relaying impractical for* the energy-constrained *UAV relays*.

For the UAVs serving the UEs, a transparent relaying is seen as a suitable choice due to low complexity of such relays [13]. Thus, in this paper, we focus on the transparent UAV relays. Unfortunately, the transparent relays do not transmit own reference signals for channel quality measurement and these signals are sent to the UEs directly by the GBS [12]–[15]. As a result, the quality of the channels between the transparent relay and the UEs cannot be measured. This is a serious barrier for practical deployment of the transparent UAV relays [15].

To enable deployment of the transparent UAV relays, the channel quality between the UEs and the transparent UAVs for handover of UEs as well as UAVs can be predicted via deep neural network (DNN), as proposed in [16]. Such predicted channel quality can be fed into any solution for setting of CIO or other handover parameters. The optimization of the handover parameters at the current time impacts performance in the future. Hence, the optimal handover decision requires knowledge of the (long-term) *future* channel qualities. Since the future channel qualities are not known and cannot be determined in practice, deep reinforcement learning (DRL) is a suitable tool to learn the handover parameters for the GBSs and the UAVs serving the UEs.

To enable feasible handover of the transparent UAVs serving the UEs, both the DNN-based channel prediction and DRL-based handover optimization are mandatory and cannot be easily substituted by classical (non-machine learning) approaches. Nevertheless, even if both machine learning approaches can reach a high accuracy on their own, their independent deployment leads to a degradation in the overall performance. Such degradation is a result of a multiplication and an accumulation of the errors in DNN-based channel quality prediction and in DRL-based handover optimization.

Therefore, motivated by the practical challenges of handover decision in the networks with the energy efficient UAV relays in transparent mode, we propose a coordination of the machine learning-based solutions for the channel quality prediction (based on DNN) and handover optimization (based on DRL), to manage the handover of both the UEs and the transparent UAVs. To mitigate the performance degradation due to an independent deployment of both machine learning solutions, we develop an intelligent handover management framework facilitated via a coordinated DNN for channel quality prediction and DRL for CIO setting so that the overall accumulated learning error is minimized and, thus, the performance is improved. We demonstrate that the proposed coordination of the machine learning tools leads to increased sum capacity of the UEs (by up to 12.7%) and, at the same time, reduced number of handovers of UEs and UAVs (by up to 12.9% and 16.4%, respectively) compared to state-of-the-art works.

The rest of paper is structured as follows. First, we define system model and assumptions adopted in our paper. Then, we formulate the problem and discuss related challenges in Section III. Next, the proposed concept is described and compared with related works in Sections IV and V, respectively. The last section concludes the paper and outlines future directions.

II. System Model

We consider a set $\mathcal{N} = \{n_1, n_2, \dots, n_N\}$ of the UEs deployed in an area covered by the sets $\mathcal{G} = \{g_1, g_2, \dots, g_G\}$ of GBSs and by the set $\mathcal{F} = \{f_1, f_2, \dots, f_F\}$ of the UAVs acting as flying base stations (see Fig. 1). The UEs change their position over time and the positions of the UAVs are adjusted according to the actual position of the served UEs. Since our proposed work is independent of the UAVs' positioning, we assume the UAVs follow the center of gravity of the

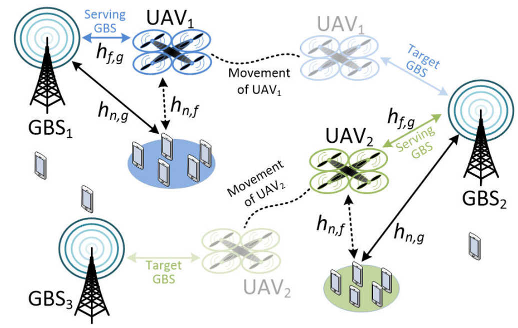

Fig. 1. System model with GBSs and transparent UAVs serving mobile UEs.

served UEs as, e.g., in [1]. Each UE can be served by any of $\mathcal{K} = \mathcal{G} + \mathcal{F} = \{k_1, \dots, k_G, k_{G+1}, \dots, k_K\}$ base stations (BSs) encompassing the GBSs and the UAVs, i.e., $K = F+G$. We define a binary variable $\beta^{UE}_{n,k}$, which indicates if the n-th UE is connected to the k-th BS ($\beta^{UE}_{n,k} = 1$) or not ($\beta^{UE}_{n,k} = 0$). Furthermore, the variable $\beta^{UAV}_{f,g}$ indicates if the f-th UAV is connected to the g-th GBS ($\beta^{UAV}_{f,g} = 1$) or not ($\beta^{UAV}_{f,g} = 0$).

The connection of the UEs to the BSs and the connection of the UAVs to the GBSs change over time as the handovers of UEs and UAVs are performed. For the handover triggering, we follow A3 event (defined by 3GPP [17]), thus, the handover of the x-th device (UE or UAV, i.e., $x \in \{\text{“}n\text{”}, \text{“}f\text{”}\}$) is initiated if the following condition is satisfied for a period of TTT:

$$s_{t,x} + CIO_t - \Delta_H > s_{s,x} + CIO_s \tag{1}$$

where $s_{t,x}$ and $s_{s,x}$ represent the level of signal received by the x-th device from the target and serving BSs, respectively; CIO_t and CIO_s are the CIOs of the target and serving BSs, respectively; and Δ_H is the hysteresis. To avoid cluttering and to focus on key novelty laying in the proposed coordination of machine learning, we target the optimization of CIO while the optimizations of TTT and Δ_H are left for the future work.

The channel quality is defined as signal to interference plus noise ratio at the j-th receiver from the i-th transmitter so that:

$$\gamma_{i,j} = \frac{P_i h_{i,j}}{\sigma^2 + \sum_{\forall k \in \mathcal{K}, k \neq i} P_k h_{k,j}} \tag{2}$$

where P_i stands for the transmitting power of the transmitter (represented by the BSs), $h_{i,j}$ and $h_{k,j}$ are the channel qualities from the i-th and k-th transmitters, respectively, to the j-th receiver, and σ^2 stands for the noise. The summation $\sum_{\forall k \in \mathcal{K}, k \neq i} P_k h_{k,j}$ represents the interference from other BSs, as all BSs reuse the same communication resources.

Since we target downlink communication, the transmitter is represented by either the g-th GBS or by the f-th UAV (i.e., $i \in \{\text{“}g\text{”}, \text{“}f\text{”}\}$ and the receiver is either the f-th UAV or th n-th UE (i.e., $j \in \{\text{“}f\text{”}, \text{“}n\text{”}\}$). Then, the communication capacity of the n-th UE served directly by the g-th GBS and via the f-th UAV are, respectively, defined as:

$$c_{n,g} = B_n \log_2(1 + \gamma_{n,g}) \tag{3}$$

$$c_{n,f,g} = \frac{B_n}{2} \min\big(\log_2(1 + \gamma_{n,f}, \log_2(1 + \gamma_{f,g})\big) \tag{4}$$

where B_n is the channel bandwidth allocated to the n-th UE. Note that we target optimization of handover while the optimization of bandwidth is not directly related to such problem. Thus, we allocate the bandwidth according to the required capacity of the n-th UEs $c_{req,n}$ as in [1] so that:

$$B_n = \frac{c_{req,n}}{\log_2(1+\gamma_{k,n})} \tag{5}$$

where $\gamma_{k,n}$ stands for the channel quality between the serving k-th BS and the n-th UE. If B would not be sufficient for all served UEs, the bandwidth is sequentially allocated starting from the UE with the highest channel quality until any bandwidth is available or all UEs are served with $c_{n,req}$.

The communication capacity of the n-th UE is generally defined as:

$$c_n = c_{n,g}\beta^{UE}_{n,g} + c_{n,f,g}\beta^{UE}_{n,f}\beta^{UAV}_{f,g} \tag{6}$$

where $\beta^{UE}_{n,g} = 1$ and $\beta^{UE}_{n,f} = 1$ indicate connection of the n-th UE to the g-th GBS and the f-th UAV, respectively.

The UEs served by the same k-th BS impose the total load $\rho_k = \sum_{\forall n \in \mathcal{N}} \beta^{UE}_{n,k} B_n / B_k$ to k-th BS. The load $\rho_k \in \langle 0, 1\rangle$ represents the proportion of bandwidth allocated by the k-th BS to its served UEs out of the all bandwidth of the k-th BS.

III. Problem Formulation

Our goal is to optimize the handover of UEs and UAVs in the scenario with unknown channels $\boldsymbol{h^*_{n,f}}$ among the UEs and the transparent UAV relays so that the sum capacity of the UEs served by the UAVs is maximized while avoiding redundant handovers. To this end, we define the problem of CIO setting for all BSs, i.e., determining $\boldsymbol{CIO^*} = \{CIO_1, CIO_2, \ldots, CIO_K\}$, as:

$$\boldsymbol{h^*_{n,f}}, \boldsymbol{CIO^*} = \underset{\boldsymbol{CIO} \in \mathcal{O}, \boldsymbol{h_{n,f}} \in \mathbb{R}}{\operatorname{argmax}} \sum_{\forall n \in \mathcal{N}_f} c_n - \mu_n \tag{7}$$

where $\mathcal{N}_f \subset \mathcal{N}$ is the set of UEs connected to the UAVs, $\mathcal{O}$ is the set of possible CIO values from CIO_{min} to CIO_{max}, and μ_n stands for the cost of handovers represented by signaling (dozens or hundreds of kbits per handover per UE, see [17]).

A key challenge and difference with respect to related works is that a proper $\boldsymbol{CIO^*}$ setting in our targeted practical scenario with the energy efficient transparent UAVs is directly dependent on the prediction of the channel quality $\boldsymbol{h^*_{n,f}}$ among the transparent UAV relays and the UEs. Such channel quality is not commonly available, since the transparent UAVs do not transmit any own reference signals to measure the channel quality [12]–[15]. Hence, we propose a completely novel coordination between the DNN-based channel quality prediction and the DRL-based CIO setting handling both sub-problems jointly with a mutual awareness of each other.

IV. Proposed oordinated Machine Learning for Mobility Management

In this section, we first outline related works on channel quality prediction and CIO setting, which are a basement for our work. Then, we describe the proposed concept of the coordinated machine learning for the UEs' ad UAVs' handover optimization in the scenario with the transparent UAV relays.

A. Background on Channel Quality Prediction and CIO

The proposed concept builds on our prior works on DNN-based channel quality prediction for device-to-device (D2D) communication [16] and DRL-based adjustment of CIO [1].

The DNN-based D2D channel quality prediction exploits known channel quality of two UEs to the serving GBS and few neighboring GBSs. Information on the channel quality from the two UEs to the few GBSs is fed into DNN to predict the quality of the direct channel between the two UEs. This concept is beneficial in scenarios with massive number of UEs, since it reduces overhead related to channel quality measurement for radio resource management purposes. As shown in [16], the concept yields high correlation (about 95%) between true and predicted D2D channel qualities even in the urban scenario with buildings resulting in a non-light-of-sight (NLoS) communication. The idea can be applied also to predict the channel between the UE and the serving transparent UAV, which is in such case seen as one UE communicating with the other UE.

The CIO setting is the problem, where a decision at the current time step impacts performance in all subsequent time steps in the environment with a high randomness and unpredictable future evolution. Hence, the problem is NP-hard and requires a prior knowledge of the future channel qualities, which is not available and cannot be predicted for a long-term period of the network operation. Such problem can be solved by DRL. As shown in [1], the actor-critic DRL leads to a high performance (in terms of sum capacity and handover mitigation) while it converges sufficiently fast for practical applications.

In line with [1], we define the states, actions and rewards for the actor-critic DRL in the following way. The set of states $S(t)$ at the time t represents the network status by means of the load of BSs ρ_k. To reflect potential changes due to handover of UE or UAV, the state space includes also an additional load $\rho_h(t)$ that would be added if the handover of UEs or UAVs is performed. Hence, the state space is defined as $S(t) = \{\rho_1(t), \ldots, \rho_K(t), \rho_h(t)\}$. The action corresponds to the CIO setting for individual BSs, hence, the action space is $A(t) = \{CIO_1(t), \ldots, CIO_K(t)\}$. The reward, inspired by the problem formulation from the CIO optimization perspective, is defined as:

$$R = \frac{\sum_{\forall n \in \mathcal{N}_f} c_n}{|\mathcal{N}_f| c_{req,n}} - \left(\sum_{\forall n \in \mathcal{N}_h} \rho^*_n \frac{\rho_{t,n}}{\rho_{s,n}} + |\mathcal{N}_h|\mu_n \right) + 1 \tag{8}$$

where $\mathcal{N}_h \subset \mathcal{N}$ is the set of UEs performing handover in the given time step (if the UAV performs handover, then all UEs served by this particular UAV are accounted for in $\mathcal{N}_h$), ρ^*_n is the load implied by the handover of the n-th UE, $\rho_{t,n}$ and $\rho_{s,i}$ are the loads of the target and serving BSs of the n-th UE, μ_n is the cost of handover of the n-th UE, and "+1" is the constant added for purposes of a stability of the system and it is determined experimentally.

B. Proposed Coordination of DNN and DRL

In this section, we describe the proposed coordination of the DNN-based channel quality prediction and the DRL-based

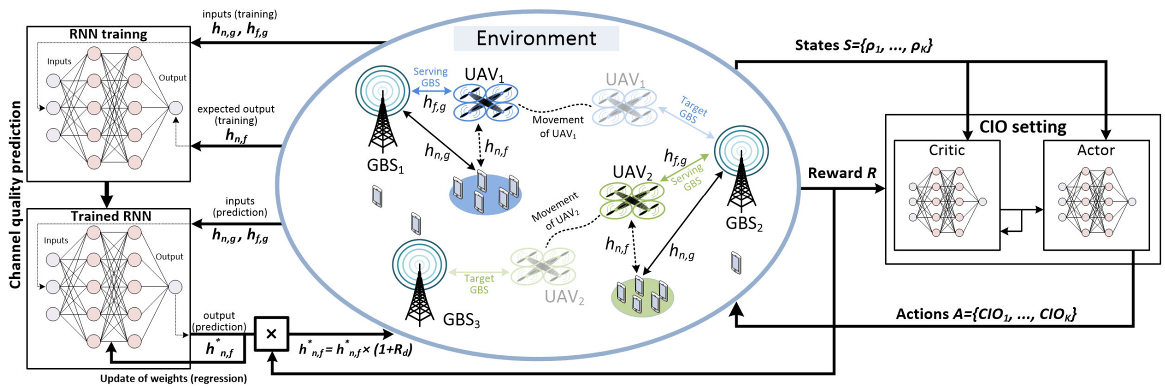

Fig. 2. Coordination of machine learning for RNN-based channel quality prediction between UEs and transparent UAV relays and DRL-based CIO setting.

CIO setting towards the sum capacity maximization while reducing the number of UEs' and UAVs' handovers. The proposed concept of coordinated machine learning predicting the channel quality between UEs and transparent UAV relays and setting CIO for all BSs is depicted in Fig. 2. The environment, where the proposal is applied, represents the mobile network with UEs, GBSs, and UAVs. The channel quality from the UEs and the UAVs to the GBSs is commonly known to the network and is measured in a traditional way via reference signals [18]. Such information represents inputs for the training of channel quality prediction. The targets for supervised learning are represented by true channel qualities between the UAVs and the UEs. The CIO is continuously adjusted over time according to the loads of BSs (representing the state of the network) based on outcomes of the DRL.

Now, lets us explain details of the principle of the proposed coordination of both machine learning tools towards the same objective. First, motivated by an experimental observation that any error in the channel quality prediction is propagated more significantly and degrades performance notably, the DNN for channel quality prediction considered in related work [16] should be updated to adjust the prediction itself. To this end, we replace a general DNN with a recursive neural network (RNN). This allows to consider the actual output of the RNN at the time t (predicted channel quality $\boldsymbol{h^*_{n,k}}$) for an update of the internal state (weights) of the RNN. Since the actual inputs by means of the channels of the UEs to the serving and neighboring BSs are known to the RNN, using the output of the RNN is enough to adjust future predictions. The RNN weights are adjusted using regression considering originally trained and the new predicted channel qualities between the UE and the transparent UAVs. This allows us to avoid deviation of the channel quality prediction from a DNN training phase in case the DRL-based CIO setting would lead to impractical values resulting from an error in the channel quality prediction. In other words, the loop in the RNN guarantees that a potentially wrong or sub-optimal setting of CIO would not steer the channel quality prediction towards wrong values as well.

Second, since the channel quality prediction and CIO setting are related to each other and mutually dependent (handover decision depends on predicted channel quality), the reward R normally intended only for DRL creates a virtual interconnection of both RNN and DRL. Thus, the reward is not fed only to the DRL, but also impacts output of the RRN. Of course, one could think about inserting the reward to RNN directly as an input. Unfortunately, such approach makes the training impractical, as an uncertainty in DRL decision leads to a slow convergence. Therefore, we exploit the reward for an adjustment of the predicted channel quality. To this end, we determine the difference in rewards (R_d) in two consecutive steps t and $t-1$ (corresponding to two consecutive handover events by any UE or UAV) so that:

$$R_d = \begin{cases} 1 - \big(R(t-1)/R(t)\big) & \text{if } R(t-1) \geq R(t) \\ 0 & \text{if } R(t-1) < R(t) \end{cases} \quad (9)$$

This variable indicates a change in the reward encapsulating both new channel conditions due to new positions of UEs and UAVs as well as changes in CIOs of individual BSs.

The R_d adjusts the predicted channel quality so that $h^*_{n,f} = h^*_{n,f} \times (1+|R_d|)$. Of course, considering a metric representing solely the impact of the channel quality prediction accuracy while keeping the same CIO would be a more suitable and straightforward than the R_d. Unfortunately, such metric cannot be derived in real-world networks during the network operation, since the channel between the UE and the transparent UAV cannot be measured via traditional approaches and values estimated via simulation might be inaccurate. Thus, we adopt R_d encompassing both CIO and channel quality prediction accuracy. Still, note that if the update of channel quality prediction by R_d would lead to even a larger error in the prediction, the error is reflected in the CIO setting via DRL and would result in a lower reward. Then, the lower reward leads to an adjustment of the channel quality prediction in the next step. Hence the problem of non-isolated channel quality prediction and CIO setting is automatically and continuously monitored and resolved via the proposed coordination. Note also that the load of BSs is influenced by handovers, which are performed according to the predicted channel quality between the UEs and the transparent UAV relays. Thus, the inputs to the DRL for CIO setting already reflects the predicted channel quality and no additional input from the RNN is required.

V. Performance Evaluation

In this section, we define scenario, models, competitive algorithms, and metric for performance evaluation. Then, simulation results are presented and discussed.

A. Simulation Scenario and Models

We consider an area of 2×2 km with three GBSs deployed at random positions, but with the minimum distance among GBSs of 750 meters. Up to eight UAVs are also deployed in this area. Since optimization of the UAV's position is not relevant for our problem, the UAVs are placed to the center of gravity of their served UEs. The UEs move according to the random walk mobility model with the speed of each UE randomly selected from $\{1.5, 3, 4.5\}$ m/s. Like in the related works on handover management, we model path loss among all entities as $92.45 + 20 \times \log_{10}(d) + 20 \times \log_{10}(fr)$, where d is the distance (in km) between the BS and the UE and fr is the carrier frequency (in GHz). The results are averaged out over 25 simulation runs, each consisting of 18000 steps, each with a duration of 10 ms. The simulation parameters are summarized in Table I. The simulation environment is used also for creation of data set to train DNN, since no real-world data set for our purposes is available.

The RNN is composed of an input layer, six hidden layers out of which three fully connected layers (with 5 neurons) are interleaved with three LSTM layers (with 64 hidden units) allowing to implement the internal loop for the RNN. Sigmoid activation function is used. The output layer is based on regression and returns the predicted channel quality. For the training, the number of epochs is set to 150 and the batch size is 20 samples. These settings are determined experimentally.

The DRL is inspired by [1] and is represented by the actor and the critic created by two fully connected neural networks. The actor neural network has four hidden layers (120 neurons in each) and the activation is by means of ReLU. The output layer of the actor has $5^{|\mathcal{K}|}$ neurons, where 5 is the number of options for CIO setting, see Table I and [20]. We use softmax function in the output layer of the actor neural network, since our action space is discrete. The critic neural network has three hidden layers (120 neurons in each) and ReLU activation. The output layer of the critic has one neuron. The learning rate of the actor as well as critic is set to 0.01. The actor and critic settings are determined experimentally.

TABLE I
SIMULATION PARAMETERS AND SETTING

Number of UEs	45 per GBS and 15 per UAV
UAV altitude	80 m [1]
Number of GBSs/UAVs	3 / from 1 to 8
Carrier frequency	1.8 GHz
Bandwidth (all BSs)	100 MHz
Tx Power of GBS/UAVs	23 / 15 dBm
Hysteresis Δ_H / TTT	3 dB / 0.05 s
Handover cost per UE μ_n	100 kb [17]
Set of possible CIO values	{-6, -3, 0, 3, 6} dB [20]

B. Competitive Algorithms and Performance Metrics

We compare the proposed concept with following state-of-the-art approaches and benchmarks:

- *Association* – the UEs are always connected to the BS providing the highest channel quality leading to maximized capacity at the cost of a high number of handovers.
- *Static CIO and DNN channel quality prediction* – state of the art channel quality prediction [16] combined with traditional setting of CIO to a static value.
- *Static CIO with perfectly known channels* – static CIO as in 3GPP [17] with an impractical assumption on perfectly known channels among the all UAVs and UEs.

The performance is assessed using following metrics:

- *Capacity* – the sum capacity of UEs averaged over time and simulations.
- *Number of handovers* – total number of handovers during simulations. We investigate separately the number of handovers performed by the UEs and by the UAVs.

C. Simulation Results

First, we investigate the sum capacity of UEs in Fig. 3. As expected, the highest capacity is observed for the association, since the UEs are always connected to the BS with the highest channel quality. The proposed concept outperforms the static CIO with DNN predicted channels by up to 12.7%. This gain is introduced by a smart setting of CIO reflecting a potential error in the DNN channel quality prediction via the coordination of the DNN and the DRL. Figure 3 also demonstrates that even if we could know perfectly all channels among the UAVs and UEs, which is not realistic and feasible for the transparent UAVs, still, the proposed coordination of the DNN and the DRL introduces a gain in capacity even more than 5%. The gain results from the fact that the coordination between the DNN and the DRL allows to mitigate a negative impact of the errors in channel quality prediction on CIO setting.

Furthermore, we investigate the number of handovers performed by the UEs (Fig. 4) and by the UAVs (Fig. 5). For both cases, the proposed solution reaches the lowest number of handovers compared to all other approaches. The highest number of handovers is reached by the association (almost

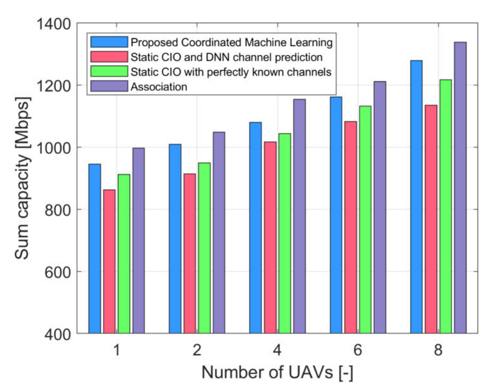

Fig. 3. Impact of the number of UAVs on the sum capacity of UEs.

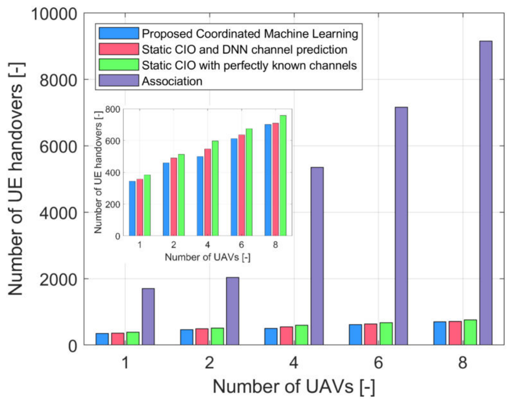

Fig. 4. Number of handovers performed by UEs during the simulations. Note that the zoomed part includes all algorithms but association, to illustrate differences among algorithms reaching similar performance.

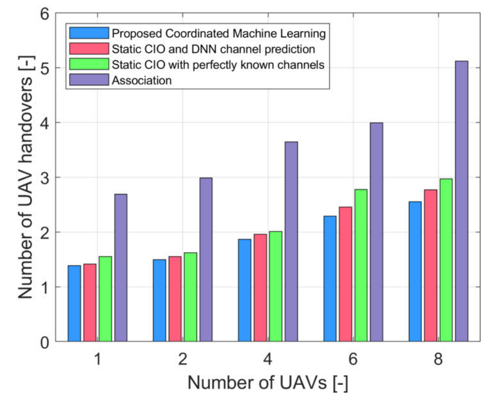

Fig. 5. Number of handovers performed by UAVs during the simulations. Note that the number of handovers performed by the UAVs is about two order of magnitude lower than the number of handovers by UEs, since, the number of UEs is higher (15 UEs/UAV and 45/GBS) and each UAV handover is accounted also for handovers of served UEs, because these UEs in fact switch among GBSs. Besides, UEs movement is more random leading to more handovers.

twice for UAVs and almost 10-times for UEs with respect to the proposal). This huge number of handovers is a cost paid for a small increase in the capacity. Comparing the proposal with static CIO setting with DNN-based channel prediction shows that the proposed coordinated approach reduces the number of handovers by 6.7% and 7.8% for the UEs and UAVs, respectively. The reduction in the number of handovers by the proposal is even more notable (12.9% and 16.4% for UEs and UAVs, respectively) if compared to the static CIO with all the channels among UEs and UAVs being perfectly known. The reason is that the perfectly known channels allow to associate the UEs so that the capacity is increased (see Fig. 3) compared to the solution with the DNN predicted channels, but the capacity increase is at the cost of additional handovers. This trade-off observed for the related works is eliminated by the proposed coordination of both CIO setting and DNN channel quality prediction enabling to reduce the number of handovers while increasing the capacity compared to the related works.

VI. Conclusions

We have proposed a novel coordination of the DNN for the prediction of the channel quality between UEs and energy efficient transparent UAV relays with the DRL for CIO setting to mitigate the accumulation of errors resulting from individual machine learning tools. The concept is based on an internal feedback in the neural network and an integration of the DRL reward to an adjustment of the output of the neural network. This allows both to increase the sum capacity and reduce the number of handovers compared to the related works.

In the future, joint optimization of multiple handover decision parameters should be investigated and the interaction among machine learning tools should be further optimized.

References

[1] A. Madelkhanova, Z. Becvar and T. Spyropoulos, "Optimization of Cell Individual Offset for Handover of Flying Base Stations and Users," *IEEE Transactions on Wireless Communications*, vol. 22, no. 5, 2023.

[2] J. Angjo, I. Shayea, M. Ergen, H. Mohamad, A. Alhammadi and Y. I. Daradkeh, "Handover Management of Drones in Future Mobile Networks: 6G Technologies," *IEEE Access*, vol. 9, 2021.

[3] D. Lopez-Perez, I. Guvenc, X. Chu, "Mobility management challenges in 3GPP heterogeneous networks," *IEEE Communications Magazine*, vol. 50, no. 12, Dec. 2012.

[4] 3GPP, "NR; Radio Resource Control (RRC) protocol specification (Release 17)," 3GPP, TS 38.331 v17.6.0, Sept. 2023.

[5] G. Alsuhli, K. Banawan, K. Seddik and A. Elezabi, "Optimized Power and Cell Individual Offset for Cellular Load Balancing via Reinforcement Learning," *IEEE WCNC* 2021.

[6] K. da Costa Silva, Z. Becvar, C. R. Frances, "Adaptive Hysteresis Margin Based on Fuzzy Logic for Handover in Mobile Networks with Dense Small Cells", *IEEE Access*, vol. 6, 2018.

[7] J. Shodamola, U. Masood, M. Manalastas and A. Imran, "A Machine Learning based Framework for KPI Maximization in Emerging Networks using Mobility Parameters," *IEEE BlackSeaCom* 2020.

[8] R. Kwan, R. Arnott, R. Paterson, R. Trivisonno and M. Kubota, "On mobility load balancing for LTE systems," *IEEE VTC-Fall* 2010.

[9] M. M. U. Chowdhury, W. Saad and I. Guvenc "Mobility Management for Cellular-Connected UAVs: A Learning-Based Approach," *IEEE ICC workshops,* 2020.

[10] Y. Chen, X. Lin, T. Khan and M. Mozaffari, "Efficient drone mobility support using reinforcement learning," *IEEE WCNC* 2020.

[11] Y. Jang, S. M. Raza, H. Choo and M. Kim, "UAVs Handover Decision using Deep Reinforcement Learning," *IMCOM* 2022.

[12] 3GPP, "Understanding on Type 1 and Type 2 Relay," TSG RAN WG1 Meeting # 57bis, R1-092370, Huawei, LA, CA, USA, June 2009.

[13] M. Najla, Z. Becvar, P. Mach and D. Gesbert, "Positioning and Association Rules for Transparent Flying Relay Stations," *IEEE Wireless Communications Letters*, vol. 10, no. 6, 2021.

[14] R. N. Braithwaite, "Improving data throughput for cell-edge users in a LTE network using up-link HARQ relays," *IEEE VTC-Fall* 2011.

[15] 3GPP, "Type 2 Relay Summary", RAN1 Meeting #60, R1-100951, ALU, ALU Shanghai Bell, CHTTL, San Francisco, CA, USA, Feb. 2010.

[16] M. Najla, Z. Becvar, P. Mach and D. Gesbert, "Predicting Device-to-Device Channels from Cellular Channel Measurements: A Learning Approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, 2020.

[17] 3GPP, "Handover Procedures," 3GPP, TS 23.009 v16.0.0, Jul. 2020.

[18] 3GPP "NR; Physical Layer Measurements," TS 38.215, v17.1.0, 2022

[19] 3GPP, "Mobility Enhancements in Heterogeneous Networks (Release 11)," TR 36.839, v11.1.0, Dec. 2012.

[20] K. Attiah et al., "Load Balancing in Cellular Networks: A Reinforcement Learning Approach," *IEEE CCNC* 2020.

[21] A. Al-Hourani, S. Kandeepan and S. Lardner, "Optimal LAP Altitude for Maximum Coverage," IEEE Wireless Communications Letters, vol. 3, no. 6, pp. 569-572, Dec. 2014.

[22] R. I. Bor-Yaliniz, A. El-Keyi and H. Yanikomeroglu, "Efficient 3-D placement of an aerial base station in next generation cellular networks," *IEEE ICC* 2016.