Deep Reinforcement Learning-based User-to-Access Points Association in Sub-6 GHz/mmWave Beyond 5G Networks

Thi Ha Ly DINH, Megumi KANEKO (NII), Keisuke WAKAO, Kenichi KAWAMURA, Takatsune MORIYAMA, Hirantha ABEYSEKERA, Yasushi TAKATORI (NTT)

Introduction

- **Integration of mmWave and sub-6GHz interfaces** is envisioned to fulfill the stringent requirements of Beyond 5G applications
- Many issues are yet to be solved for seamless integration
- We investigate the problem of **distributed user-to-access points (AP) association and beamforming**
- We propose the **Deep Q-Network-based joint user association and beamforming** (**DQN-UABF**) method

Problem definition

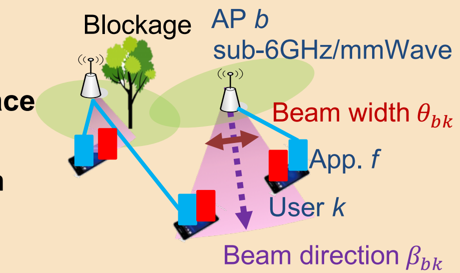
Goal: **Maximize average sum-rate**

subject to

C1. Each app. of each user is **served by at most 1 AP/interface**

C2. Allocated rate of each app. for each user is **larger than min required rate R_{kf}**

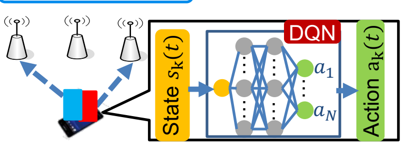
C3. APs are **not overloaded**



Proposed method

- User k decides to request its desired (AP, interface) for each app., through learning by **Deep Q-Network (DQN)**
  - State $s_k(t)$: actual association (user k, app. f, interface v, AP b) at time t
  - Action $a_k(t)$: desired future association (user k, app. f, interface v, AP b) at time t

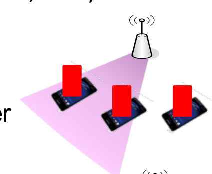
Step 1: Each user **ϵ -greedy method**



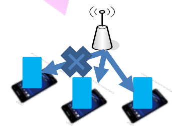
- Select the action $a_k(t)$ for its current state $s_k(t)$ from its **DQN** Q

Step 2: Each AP

- mmWave interface:
  - Select best user cluster
  - Set θ_{bk}, β_{bk}



- Sub-6GHz interface: as [1]
  - Select best users/apps.
  - If overload: remove users/apps. with max load



- Drop non-selected apps., send association decision to users through feedback

Step 3: Each user

- Calculate the reward from AP feedback
- Update DQN weights, move to its new state s'_k

Results

- 15 users, 50 random positions
  - App. 1 with $R_{k1} = \mathbf{100\ Mbps}$
  - App. 2 with $R_{k2} = \mathbf{1\ Mbps}$

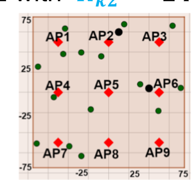


Fig. 1: Simulation scenario

- #hidden layers: 2, #neurons per hidden layer: 16, epoch size: 10, memory size: 100

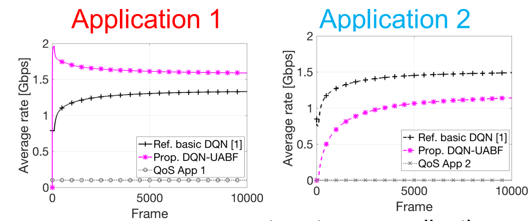


Fig.2: Average data rate per application

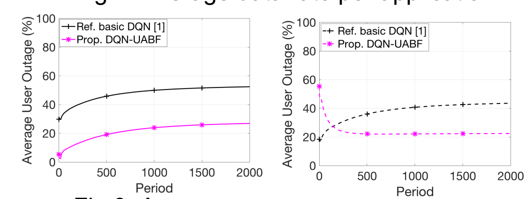


Fig.3: Average user outage per application

Proposed DQN-UABF vs ***Reference basic DQN*** [1]:

✓**Better adaptation of served rates** to QoS demands: higher rate for app. 1 with higher R_{k1}, lower rate for app. 2 with lower R_{k2},

✓**Reduction of outage** events,

✓**Enhanced outage fairness** among applications.

[1] T. H. L. Dinh, M. Kaneko et al., "Reinforcement Learning-aided Distributed User-to-Access Points Association in Interfering Networks," in *IEEE Global Communications Conference*, pp.1–6, 2019.