2022 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-2159-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISIT50566.2022.9834463

QoE-Centric Multi-User mmWave Scheduling: A Beam Alignment and Buffer Predictive Approach

Babak Badnava, Sravan Reddy Chintareddy, Morteza Hashemi
Department of Electrical Engineering and Computer Science, University of Kansas

***Abstract*—In this paper, we consider the multi-user scheduling problem in millimeter wave (mmWave) video streaming networks, which comprises a streaming server and several users, each requesting a video stream with a different resolution. The main objective is to optimize the long-term average quality of experience (QoE) for all users. We tackle this problem by considering the physical layer characteristics of the mmWave network, including the beam alignment overhead due to pencil-beams. To develop an efficient scheduling policy, we leverage the contextual multi-armed bandit (MAB) models to propose a beam alignment overhead and buffer predictive streaming solution, dubbed *B2P-Stream*. The proposed *B2P-Stream* algorithm optimally balances the trade-off between the overhead and users' buffer levels, and improves the QoE by reducing the beam alignment overhead for users of higher resolutions. We also provide a theoretical guarantee for our proposed method and prove that it guarantees a sub-linear regret bound. Finally, we examine our proposed framework through extensive simulations. We provide a detailed comparison of the *B2P-Stream* against a uniformly random and Round-robin (RR) policies and show that it outperforms both of them in providing a better QoE and fairness. We also analyze the scalability and robustness of the *B2P-Stream* algorithm with different network configurations.**

***Index Terms*—Quality of Experience, mmWave Networking, Multi-user Streaming and Scheduling**

I. Introduction

3GPP broadband wireless standards such as LTE-Advanced and fifth generation (5G) technologies and IEEE wireless standards such as 802.11ad and 802.11ay have enabled high data transfer and data-intensive applications, and are moving towards all-connected small-cell networks. Now, the monthly global average data usage per smartphone is about 11.4GB and it is expected to reach 41GB by the end of 2027. Also, about 69% of the world's mobile data traffic corresponds to mobile video streaming, and it is expected to reach 79% by 2027 [1–3]. This deluge of data traffic, especially demands for high resolution video streaming on portable mobile devices, will pose significant challenges for the wireless and cellular network providers to meet the quality of experience (QoE) requirements. In contrast to the quality of service (QoS) that is usually quantified in terms of achieved rate and latency, video streaming QoE depends on several factors such as the resolution of video frames, number of re-buffering events, and frequency of resolution switches.

In terms of required infrastructure, millimeter wave (mmWave) networks are capable of providing multi-Gbps data rates, which makes them suitable to meet the ever-increasing demand for video streaming applications [4, 5]. However, unlike omni-directional communications in sub-6 GHz, high data rates in mmWave systems come at the price of large coordination overhead due to highly directional communications needed to compensate for large channel losses [6–8].

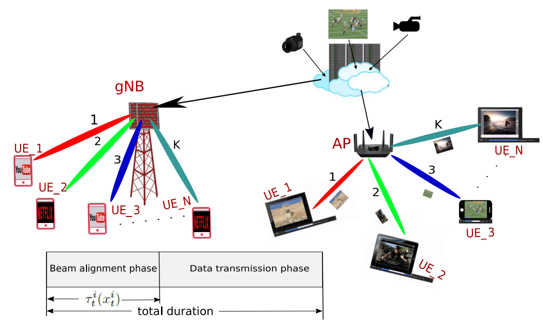


Fig. 1: System model depicting mmWave capable base station (gNB) and an Access Point (AP) serving K of N users simultaneously.

Although there are extensive works on providing more efficient beam alignment[1] solutions [9–12], this process still consumes resources that otherwise could have been utilized for high-bit-rate data transfer.

A mmWave base station, equipped with multiple RF chains (K), can serve up to K users at the same time. As such, multi-user management plays a central role to guarantee low-latency and high QoE for all users [11, 12]. The key point, however, is that due to beam alignment overhead, switching from one mobile user to another one incurs a *switching cost*, as denoted by $\tau(.)$ in Figure 1.

Assuming that the QoE is a function of the playback buffer level and resolution of the video frames stored in the buffer, the system needs to balance between users' buffer levels vs. beam alignment overhead to optimize the QoE across all the users in the network. For instance, in the extreme situation, the system could serve only a fixed subset of K users, with the objective of reducing the beam alignment overhead. While this minimizes the risk of zero playback buffer for those K users, the other under-served users would exhaust their playback buffers, which leads to significant QoE degradation. On the other hand, quickly switching between users results in significant beam alignment overheads that is a relatively very slow process vis-à-vis data transfer.

In this paper, we consider the interplay between beam alignment overhead (i.e., switching cost) and multi-user scheduling in order to enhance the QoE across all users. On one hand, the optimal scheduling should take the switching cost into account, and on the other hand, switching cost is a function of the scheduling algorithm that determines the beam quality. This is in contrast to the classical scheduling problems, where the switching overhead is traditionally assumed to be negligible compared to the service time [13].

[1] In this paper, beam alignment, collectively, refers to initial beam search, beam tracking, beam refinement, and beam switching.

Within this context, and given that beam alignment overhead is a function of the previous schedules, we develop a multi-user scheduling algorithm that works based on predicting the *beam alignment overhead and buffer*. We refer to this algorithm as *B2P-Stream* that is built upon the contextual multi-armed bandit (MAB) models to optimally balance the trade-offs between buffer levels and beam alignment overhead. To maximize the average QoE for all users, the streaming server estimates the beam alignment overhead as well as the playback buffer level at each user, and selects K users out of N users at each time slot.

There are several studies that have considered the scheduling task under different scenarios and using different tools. Jiang et al. [14] used a multi-task deep learning scheme, [15] uses Lyapunov optimization, [11, 12] consider the scheduling under user mobility. A different group of works studied QoE prediction and approximation using reference signal received power (RSRP) and throughput [16]; packet loss, jitter, and delay [17]; congestion indicators of a 5G network [18], and many other works in both categories [19–29]. However, our work aims to integrate the unique characteristics of the mmWave communication (i.e., beam alignment overhead) into a QoE-centric multi-user scheduling framework. In summary, the main contributions of this paper are as follows:

- We model the beam alignment overhead of individual users based on the last time we served that particular user, and we propose a dynamic model for the users' buffer level prediction.
- Given the playback buffer level, we model the QoE for each user and formulate an optimization problem to improve the long-term average QoE for all the users.
- We develop a MAB-based scheduling policy, called *B2P-Stream*, which provides a sub-linear regret bound, to solve the defined optimization problem. This algorithm incorporates estimated buffer level of each user and schedules users with the help of a heuristic trend function on the beam alignment overhead.

II. System Model and Problem Formulation

We consider a mmWave network that consists of N users and a single base station (BS) or access point (AP), referred to as the streaming server. We only focus on video streaming users and ignore other types of traffic. At each time slot, the streaming server selects K users out of N users to serve simultaneously. The K beams generated by the server are used to stream different video frames to each of the K users. Different resolution of each video is available at the server side, and each user may ask for a different resolution based on the channel condition.

Beam Alignment Model and Assumptions: As shown in Figure 1, each time slot is divided into two phases: beam alignment and data transmission. In this paper, we normalize the duration of each time slot to be equal to 1 unit of time. Thus, given that beam alignment takes $\tau(.)$ units, $1 - \tau(.)$ is the amount of time left for the data transfer phase. This overhead can occupy from 10% up to 43% of the time slot duration in the cellular networks [30, 31]. However, in a more general sense, the function $\tau(.)$ can be expressed in terms of the time interval between two consecutive schedules of a user, i.e., the beam alignment overhead for a user at a specific time depends on how long ago that user was served by the server. This model captures the "freshness" of the beam for the user. We can model this characteristic using a *non-decreasing function* of the last time a user has been served. Thus, the beam alignment overhead of user i at time t is denoted by $\tau^i(x^i_t)$, where x^i_t is the amount of time that has been passed since the last schedule of user i at time t.

Playback Buffer Dynamics: Each user i has a finite playback buffer of size s^i bytes to store video frames. Considering the resolution of the video v^i_{res} that the user is playing and its bit rate v^i_{rate}, the user has at most s^i/v^i_{rate} seconds of video to play. We denote b^i_t as the buffer level of user i at time t in seconds. When the proper beam has been created, the server can start streaming to the user at rate R. As such, the amount of data transferred to the user at a time t is obtained as follows:

$$d^i_t = (1 - \tau^i_t(x^i_t))R/v^i_{rate}, \tag{1}$$

where $\tau^i_t(x^i_t)$ is the beam alignment overhead, and d^i_t determines the amount of data, in seconds, that the server sends to the user. Due to blockage and other environmental issues, the user may not receive all the data that has been sent by the server. For the sake of exposition, we assume that the probability of successful reception is given by $\mathbb{P}^i_t$. Therefore, the amount of received data is given by $y^i_t = d^i_t \mathbb{P}^i_t$. The probability value $\mathbb{P}^i_t$ depends on the several factors such as user mobility, blockage, and propagation environment.

Now that we know how much data a user receives, the dynamics of the playback buffer level is given as follows:

$$b^i_{t+1} = \max\{b^i_t - 1, 0\} + u^i_t y^i_t, \tag{2}$$

where u^i_t is a binary control variable that determines whether the user i is scheduled at time t or not. Therefore, $u^i_t y^i_t$ determines the amount of seconds of the video that would be successfully transmitted to the user, if it is scheduled. We consider that each time slot is one second, and thus we subtract one second from the previous buffer level of user i, and then add the amounts of seconds that the user would receive in case of selection and reception.

Quality of Experience: The QoE for each user depends on the playback buffer level and the resolution of the video frames stored in the buffer. To characterize the QoE, we consider three factors. **(1)** Any interruption in the streaming is undesirable, and it happens whenever the playback buffer becomes empty. We call this event "zero-hit". **(2)** The QoE increases as the buffer level increases, but it has a diminishing return modeled as a logarithmic function. **(3)** The resolution of the video frames impacts the QoE. For two different users with the same amount of data in their playback buffers, the QoE of the user who plays a higher resolution is higher. We denote the QoE of user i at time t by q^i_t, and putting together these factors, the overall QoE can be expressed as:

$$q^i_t = (1 - \mathbb{1}_{\mathbf{0}}(b^t_i))\lambda(v^i_{res}) + \alpha \log(1 + b^i_t) - \gamma \mathbb{1}_{\mathbf{0}}(b^t_i). \tag{3}$$

The first term is an offset, which only depends on the resolution of the video the user is playing. The second term captures the diminishing return of the playback buffer, and the third term accounts for the zero-hit events that penalizes the QoE by a factor of γ.

Algorithm 1 *B2P-Stream*

Inputs:

T: Total number of timestamps

η_0: Initial learning rate

Algorithm:

1: **for** $t = 0$ to T **do**
2: $\quad A_t = \emptyset$
3: $\quad$ **for** $k = 1$ to K **do**
4: $\quad\quad$ Select arm $i_k = \arg\max_{i \notin A_t}(\tilde{\mu}_t^i + c_t^i + f^i(\tilde{b}_t))$
5: $\quad\quad A_t = A_t \cup \{i_k\}$
6: $\quad$ **end for**
7: $\quad \mathbf{u_t} = one\_hot(A_t, N)$
8: $\quad$ Perform $\mathbf{u_t}$ and observe QoE vector $\mathbf{r_t}$
9: $\quad$ **for** $i = 1$ to N **do**
10: $\quad\quad d_t^i = (1 - \tau_t^i(x_t^i))R$
11: $\quad\quad \tilde{b}_{t+1}^i = \max\{\tilde{b}_t^i - 1, 0\} + u_t^i d_t^i$
12: $\quad\quad \tilde{\mu}_{t+1}^i = \tilde{\mu}_t^i + \eta_t(r_t^i - \tilde{\mu}_t^i)$
13: $\quad\quad$ **if** $u_t^i == 1$ **then**
14: $\quad\quad\quad x_{t+1}^i = 0$
15: $\quad\quad$ **else**
16: $\quad\quad\quad x_{t+1}^i = x_t^i + 1$
17: $\quad\quad$ **end if**
18: $\quad$ **end for**
19: $\quad \eta_{t+1} = \eta_0 e^{-t/T}$
20: **end for**

QoE-Centric Optimization Problem: The objective of the server is to maximize the long-term average QoE for all users, given that switching to a new user (a user that was not scheduled in the previous time slot) incurs a beam alignment overhead $\tau(.)$. The decision variable is $\mathbf{u} \in \{0,1\}^N$, which is a binary vector of size N. At each time step, only K elements of $\mathbf{u}$ can be active, and the rest of them are zero. Therefore, we can formulate the following optimization problem:

$$\begin{cases} \max\limits_{\mathbf{u}} & \lim\limits_{T\to\infty} \frac{1}{T} \sum\limits_{t=1}^{T} \sum\limits_{i=1}^{N} q_t^i \\ \text{s.t.} & \sum_{i=1}^{N} u_t^i \leq K \ \forall_{t=1..T} \\ & b_{t+1}^i = \max\{b_t^i - 1, 0\} + u_t^i y_t^i \ \forall_{i=1..N, t=1..T} \end{cases} \tag{4}$$

The first constraint addresses the hardware limitations in terms of the number of RF chains, and the second constraints captures the playback buffer dynamics. The size of the decision space in Eq. 4 scales with the number of users in the network. In the next section, we establish an efficient scheduling framework based on contextual multi-armed bandits.

III. Beam and Buffer Predictive Streaming: B2P-Stream

Multi-Armed Bandit Models: A MAB problem is an interactive game between a learner and an environment [32]. The game repeats for a finite number of times. In each round of the game, the learner chooses an action (i.e., plays an arm) $\mathbf{u}$, and receives a reward $\mathbf{r}$, that is revealed by the environment. The reward can come from a stochastic distribution or chosen by the environment itself. The learner tries to find an optimal policy using the history of played actions and received rewards. To this end, the Upper Confidence Bound (UCB) method [33, 34] handles the *exploration and exploitation trade-off* by providing an upper bound for the estimation of the expected reward of each arm. The upper bound decreases as the number of reward samples from one arm increases, which means that we are more certain about the estimation of the expected value. There are other classes of MAB algorithms that are specified to different cases, such as the case that the learner can choose more than one arm at a time [35, 36], called combinatorial bandit problem; or another case where there are some contextual information available [33].

B2P-Stream Policy: In order to solve the optimization problem defined in Eq. 4, we model this problem as an instance of the *contextual* multi-armed bandit formulation. We designate r_t^i as the measurement of the QoE at time t for user i, and $\tilde{\mu}_t^i$ denotes the average of these measurements. In addition, the action set in this model is $\mathbb{U} \subseteq \{\mathbf{u} \in \{0,1\}^N : \|\mathbf{u}\|_1 \leq K\}$, which tells us that we have an N dimensional binary action vector that has at most K active elements.

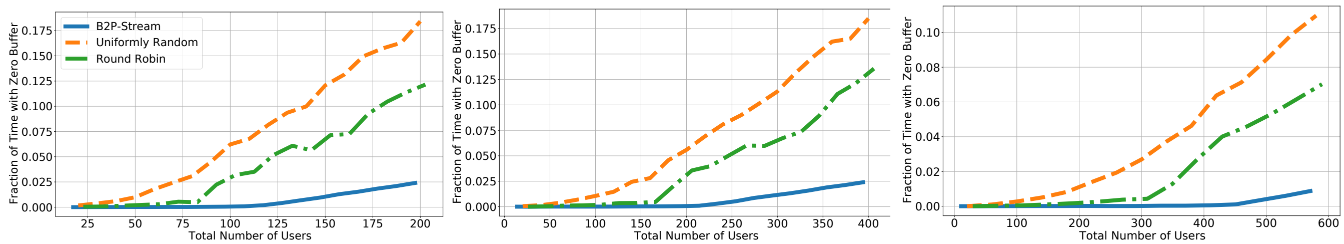
The contextual bandit model stems from the fact that the scheduler can estimate users' playback buffer level as follows:

$$\tilde{b}_{t+1}^i = \max\{\tilde{b}_t^i - 1, 0\} + u_t^i d_t^i. \tag{5}$$

Note that the state of each user changes over time according to Eq. 2, but the scheduler can only estimate the buffer level since there are unknown parameters such as the probability of successful frame reception by the user. This estimated playback buffer level along with the knowledge on beam alignment overhead function $\tau^i(.)$, which is a *non-decreasing function* as a function of the last time served, provide contextual information for the server. For the sake of presentation, we combine these two factors into a single *trend function* $f(\tilde{b}^i)$ that captures the estimated QoE for a user i. The trend function is then added to the average reward measurements received by the algorithm. In fact, MAB models with trend functions are finding applications in different domains [34].

The complete process is shown in Algorithm 1 in which first we select K users that provide the maximum outcome, and add them to a set A_t (lines 2 to 6). Then, in line 7, we create a N dimensional binary vector using A_t, and based on this vector, we create K beams and stream to the selected users, and measure the QoE $\mathbf{r_t}$. Then using Eq. 5, we update the playback buffer level estimation in line 11, for all the users. Next, we update the vector $\tilde{\mu}$ using the new measurements in line 12. Finally, from line 13 to 17, we either set the last time served to zero if we scheduled the user in the current time stamp or increase it if we did not schedule the user. Finally, the learning rate η is decreasing exponentially at each iteration.

The performance of the *B2P-Stream* can be theoretically measured in terms of regret that quantifies the gap with respect to the optimal solution. Let μ^i be the expected value of the rewards achieved by playing arm i (i.e., $\mu^i = \mathbb{E}(r^i)$), and $\mu^{i^*} = \max_i \mu^i$ be the expected value of the reward of the optimal arm. In this case, $\Delta_i = \mu^{i^*} - \mu^i$ is the immediate regret, and the accumulated stochastic regret is defined over T rounds of playing the game [32]: $R(T) = \sum_i \Delta_i \mathbb{E}(n_i(T))$, in which $n_i(T)$ is the number of times that arm i has been played over the time interval T. Then in Theorem 1, we show that using an L_f-Lipschitz trend function, we achieve a sub-linear regret bound for the *B2P-Stream* algorithm.



(a) 4 RF chains (b) 8 RF chains (c) 16 RF chains

Fig. 2: Fraction of simulation time that users experienced zero-hits per the total number of connected users to the network for *B2P-Stream*, Uniform and RR policy. The users scheduled with *B2P-Stream* experience far less number of zero-hits.

Theorem 1: Given an L_f-Lipschitz trend function, b_{max} as the maximum playback buffer level, and an $\alpha > 0$, the regret for *B2P-Stream* algorithm is upper-bounded by

$$R(T) \leq \sum_{i \neq i^*} \frac{2\alpha \log(T)}{\Delta_i - L_f b_{max}} + \frac{2\alpha}{\alpha - 1}(\Delta_i + L_f b_{max}),$$

where $\Delta_i = \mu^{i^*} - \mu^i$.

Proof: During the learning process, we either underestimate the value of all the sub-optimal actions, event G_t, overestimated the value of the optimal action, event H_t, or complement of these two events.

$$G_t) \quad \tilde{\mu}^i_{n_i} + f(\tilde{b}^i) \leq \mu^i + f(b^i) + c^i;$$
$$H_t) \quad \tilde{\mu}^{i^*}_{n_{i^*}} + f(\tilde{b}^{i^*}) \geq \mu^{i^*} + f(b^{i^*}) - c^{i^*},$$

where $c^i = \sqrt{\frac{\alpha \log(t)}{2n_i}}$. By using Hoeffding's inequality, we bound the probability of each of these two events taking place by $t^{-\alpha}$. Now, assuming that both G_t and H_t hold, we bound the number of sub-optimal arm pulls due to insufficient sampling up to this point, which means:

$$\tilde{\mu}^i_{n_i} + f(\tilde{b}^i) + \sqrt{\frac{\alpha \log(t)}{2n_i}} \geq \tilde{\mu}^{i^*}_{n_{i^*}} + f(b^{i^*}) + \sqrt{\frac{\alpha \log(t)}{2n_{i^*}}}. \quad (6)$$

Since G_t and H_t are assumed to be true, by adding c^i and c^{i^*} to both sides of G_t and H_t, respectively, we have:

$$\mu^i + f(b^i) + 2\sqrt{\frac{\alpha \log(t)}{2n_i}} \geq \tilde{\mu}^i_{n_i} + f(\tilde{b}^i) + \sqrt{\frac{\alpha \log(t)}{2n_i}} \quad (7)$$

$$\tilde{\mu}^{i^*}_{n_{i^*}} + f(\tilde{b}^{i^*}) + \sqrt{\frac{\alpha \log(t)}{2n_{i^*}}} \geq \mu^{i^*} + f(b^{i^*}) \quad (8)$$

Now, by chaining equations 6, 7, 8, and then arranging the results, we have:

$$n_i \leq \frac{2\alpha \log(t)}{(\mu^{i^*} - \mu^i + f(b^{i^*}) - f(b^i))^2} = \frac{2\alpha \log(t)}{(\Delta_i + \delta_i)^2}.$$

where $\Delta_i = \mu^{i^*} - \mu^i$ and $\delta_i = f(b^{i^*}) - f(b^i)$. Then, the expected number of times that an arm has been played is given by:

$$\mathbb{E}[n_i] = \sum_{t=1}^{T} \mathbb{E}[\mathbb{1}(I_t = i)] \leq \frac{2\alpha \log(T)}{(\Delta_i + \delta_i)^2} + \sum_{t=1}^{T} \mathbb{E}[\mathbb{1}\{G_t^c \cup H_t^c\}]$$
$$\leq \frac{2\alpha \log(T)}{(\Delta_i + \delta_i)^2} + \frac{2\alpha}{\alpha - 1}.$$

Thus, we can bound the regret by

$$R(T) = \sum_{i \neq i^*} (\Delta_i + \delta_i)\mathbb{E}[n_i] \leq \sum_{i \neq i^*} \frac{2\alpha \log(T)}{\Delta_i + \delta_i} + \frac{2\alpha}{\alpha - 1}(\Delta_i + \delta_i)$$
$$\leq \sum_{i \neq i^*} \frac{2\alpha \log(T)}{\Delta_i - L_f b_{max}} + \frac{2\alpha}{\alpha - 1}(\Delta_i + L_f b_{max}).$$

Let f be a L_f-Lipschitz function then we can upper bound $|\delta_i| = |f(b^{i^*}) - f(b^i)| \leq L_f |b^{i^*} - b^i| \leq L_f b_{max}$, where b_{max} is the maximum level of buffer. ■

IV. Numerical Results

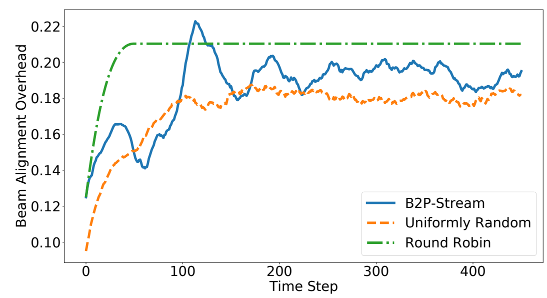


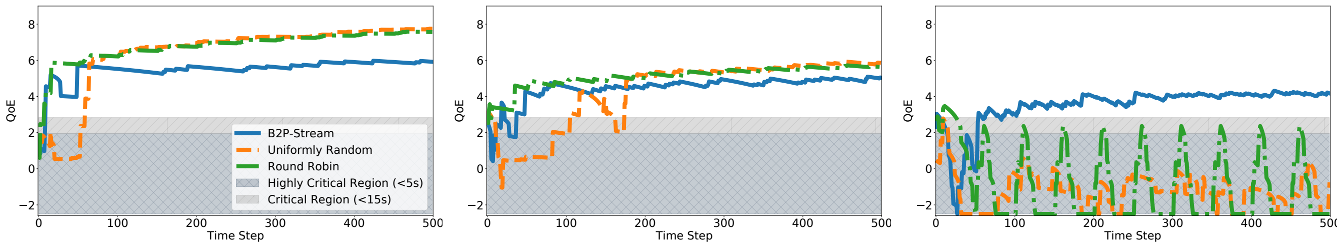
Fig. 3: Average beam alignment overhead when 4 users out of 200 are allowed to be scheduled at a time.

Simulation Setting: We evaluate the *B2P-Stream*'s performance under different conditions to make sure about the robustness of the method. We compare the *B2P-Stream* with two different baselines, namely the Uniform and the RR scheduling algorithms. In our simulations, we assumed that the path loss probability is negligible. Also, all of the users are initialized with zero buffer level.

The two main contributing factors in QoE, as shown in Eq. 3, are number of zero-hit experienced and buffer level. The zero buffer level is considered a highly unsatisfying situation for all the users, thus the method that provides a lower number of zero-hits is desirable. Also, users experience lower QoE as their playback buffer level approach zero. Thus, we define two critical situations to be able to compare the performance of different algorithms. We call the first critical situation ***"critical region,"*** that corresponds to the case when a user has less than fifteen seconds of data in the playback buffer. The other one is named ***"highly critical region,"*** which corresponds to the case when the user has less than five seconds of data in the playback buffer. Since we are initializing all the users with zero buffer levels, it is desirable that the scheduling algorithm avoid these two critical regions.

Each experiment has been run 10 times and for 500 time steps. At the beginning of each run, the video resolution and bit-rates are taken from [37], and the portion of users with specific resolution is inspired from [38].

Zero-hit Performance: Figure 2 compares the performance of *B2P-Stream* with respect to other baselines, as the number of RF chains and total number of users increase. From the results, *B2P-Stream* achieves a much smaller zero-hit compared to the RR and Uniform scheduler, which is due to the fact that



(a) Low Quality (480p) (b) Medium Quality (1080p) (c) High Quality (2160p)

Fig. 4: The average of measured QoE for 200 users in the network and 4 chains.

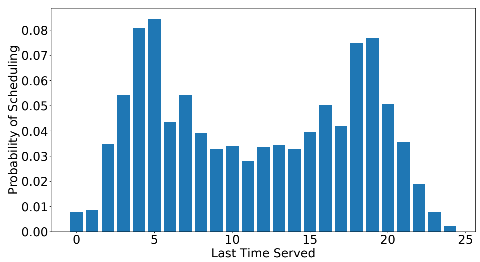


Fig. 5: Empirical distribution of the scheduled users with respect to the last time they have been served.

the *B2P-Stream* utilizes a buffer level prediction. Figure 2 also demonstrate that the *B2P-Stream* generalizes well as we use a more powerful BS (i.e. number of RF chains, K, increases). The *B2P-Stream*, as shown in Figure 3, also provides a better beam alignment overhead compared to RR, which is a result of considering the last time we served a user. On the other hand, the Uniform policy maintains a lower beam alignment overhead, but as shown in Figure 2 it fails to provide an acceptable QoE in terms of number of zero-hits.

QoE Comparison: Figure 4 shows the average QoE for 200 users connected to the network. The BS is equipped with 4 RF chains and users play videos with different resolutions. Each of the columns corresponds to a different group of users with different video resolutions: low resolution (480p), medium resolution (1080p), and high resolution (4K), respectively. In this set of results, we ignore the impact of resolution on the QoE (i.e., the first term in Eq. 3) to provide a fair comparison across different resolutions. Also, the dark grey and slate grey, in Figure 4, correspond to the critical and highly critical region, respectively. The users scheduled by *B2P-Stream* exit the critical regions much earlier than the users scheduled with Uniform or RR policy. Even though only 5% of the users in the simulation are playing a 4K video, both RR and Uniform policies cannot provide a satisfying experience. The situation is the same for users who are playing a 2K video.

Although the *B2P-Stream* provides a better zero-hit statistics compared to two other baselines, there is some cost needs to be paid. This cost is lower QoE for users of lower resolution. From the results, we note that the *B2P-Stream* algorithm maintains a lower QoE for users of lower resolution to compensate for the users of higher resolutions, as they need to be served more often because of the higher bit rate requirements. This does not mean that users with lower video quality would experience a significantly lower QoE because the QoE has a diminishing effect and there is not much of QoE difference as long as they are out of critical regions.

Intuitions Behind *B2P-Stream*: The *B2P-Stream*, as shown in Figure 5, recognizes two groups of users (i.e. two peaks). To balance the trade-off between beam alignment overhead and playback buffer levels in order to optimize the QoE metric, we can intuitively distinguish two groups of users. The first group corresponds to those users who were served recently, thus the beam alignment overhead would be small for them (small τ_t^i), and we can stream more data to this group. The second group are those users that their buffer levels are approaching zero, meaning that it has been a long time since the last time we served them (large x_t^i). Figure 6 also supports this claim by illustrating the time interval between two consecutive schedules of different groups of users in the simulation. The *B2P-Stream* may allocate more resources to the users with higher resolution in comparison with users with lower video quality. This policy results in a more fair QoE for users of different resolutions, as supported by Figure 4 as well.

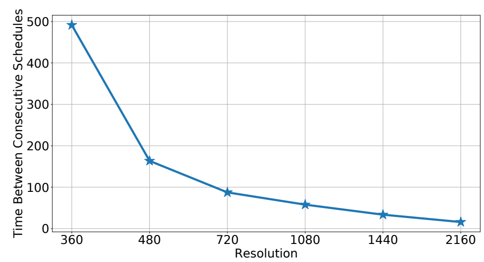


Fig. 6: Time interval between consecutive schedules of users of different video resolution.

V. Conclusion

In this paper, we considered the problem of multi-user mmWave scheduling (K users out of N) who are streaming videos with different resolutions. The overall objective is to optimize the QoE across all the users. Leveraging the contextual MAB models, we developed a QoE-centric scheduling policy that considers the physical layer characteristics of the mmWave networks. The proposed *B2P-Stream* algorithm is able to optimally balance the trade-off between the beam alignment overhead and the users' playback buffer level. In particular, *B2P-Stream* uses an estimated buffer level as an input for a *trend function* that biases the scheduling policy towards those users with exhausted buffer levels. Overall, mmWave networks are considered as one of the key enablers for data-intensive applications such as high quality video streaming. As such, developing efficient and reliable multi-user management algorithms that guarantee high QoE for all the users, is of utmost importance to enable ubiquitous mmWave technologies.

Acknowledgments

This work was partially supported by the NSF grants CNS-1955561 and CNS-1948511.

References

[1] G. M. D. T. Forecast, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2017–2022," https://s3.amazonaws.com/media.mediapost.com/uploads/CiscoForecast.pdf, [Online].

[2] C. Dehos, A. D. Domenico, and L. Dussopt, "Millimeter-wave access and backhauling: the solution to the exponential data traffic increase in 5G mobile communications systems?" *IEEE Communications Magazine*, 2014.

[3] E. M. Report, "Mobile Data Traffic Outlook," https://www.ericsson.com/en/reports-and-papers/mobility-report/dataforecasts/mobile-traffic-forecast, [Online].

[4] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, 2013.

[5] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proceedings of the IEEE*, 2014.

[6] S. Han, I. Chih-Lin, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Communications Magazine*, 2015.

[7] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive MIMO: A survey," *IEEE Communications magazine*, 2017.

[8] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan *et al.*, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, 2016.

[9] P. Zhou, X. Fang, and Y. Long, "Throughput and robustness guaranteed beam tracking for mmWave wireless networks," *IEEE/CIC International Conference on Communications in China (ICCC)*, 2017.

[10] G. R. Muns, K. V. Mishra, C. B. Guerra, Y. C. Eldar, and K. R. Chowdhury, "Beam alignment and tracking for autonomous vehicular communication using IEEE 802.11ad-based radar," *IEEE Conference on Computer Communications Workshops (INFOCOM)*, 2019.

[11] J. Lee and E. Ekici, "User Scheduling and Beam Alignment in mmWave Networks with a Large Number of Mobile Users," *IEEE Transactions on Wireless Communications*, 2021.

[12] ——, "Beam Alignment and User Scheduling in mmWave Networks under Mobility," *International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*, 2019.

[13] P.-C. Hsieh, I. Hou, X. Liu *et al.*, "Delay-optimal scheduling for queueing systems with switching overhead," *ArXiv:1701.03831*, 2017.

[14] J. Jiang, Y. Li, L. Chen, J. Du, and C. Li, "Multitask deep learning-based multiuser hybrid beamforming for mm-wave orthogonal frequency division multiple access systems," *Science China Information Sciences*, 2020.

[15] J. Wang, L. He, and J. Song, "Stochastic Optimization Based Dynamic User Scheduling and Hybrid Precoding for Broadband MmWave MIMO," in *IEEE International Conference on Communications (ICC)*, 2019.

[16] V. Adarsh, M. Nekrasov, U. Paul, A. Ermakov, A. Gupta, M. Vigil-Hayes, E. Zegura, and E. Belding, "Too Late for Playback: Estimation of Video Stream Quality in Rural and Urban Contexts," *Passive and Active Measurement*, 2021.

[17] Y. Hou, W. Zhou, L. Song, and M. Gao, "A QoE Estimation Model for Video Streaming over 5G Millimeter Wave Network," *Advances on Broad-Band Wireless Computing, Communication and Applications*, 2017.

[18] J. Nightingale, P. Salva-Garcia, J. M. A. Calero, and Q. Wang, "5G-QoE: QoE Modelling for Ultra-HD Video Streaming in 5G Networks," *IEEE Transactions on Broadcasting*, 2018.

[19] D. Pérez-Adán, Ó. Fresnedo, J. P. González-Coma, and L. Castedo, "Wideband User Grouping for Uplink Multiuser mmWave MIMO Systems With Hybrid Combining," *IEEE Access*, 2021.

[20] ——, "Wideband User Grouping for Uplink Multiuser mmWave MIMO Systems With Hybrid Combining," *IEEE Access*, 2021.

[21] C. Xu, S. Liu, C. Zhang, Y. Huang, and L. Yang, "Joint user scheduling and beam selection in mmWave networks based on multi-agent reinforcement learning," in *IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2020.

[22] H. Wu, D. Liu, W. Wu, C. Na, and M. Liu, "A low complexity two-stage user scheduling scheme for MmWave massive MIMO hybrid beamforming systems," in *3rd IEEE International Conference on Computer and Communications (ICCC)*, 2017.

[23] F. Firyaguna, A. Bonfante, J. Kibiłda, and N. Marchetti, "Performance Evaluation of Scheduling in 5G-mmWave Networks under Human Blockage," *ArXiv:2007.13112*, 2020.

[24] R. Singh and P. Kumar, "Optimal Decentralized Dynamic Policies for Video Streaming over Wireless Channels," *ArXiv:1902.07418*, 2019.

[25] M. F. Tuysuz and M. E. Aydin, "QoE-Based Mobility-Aware Collaborative Video Streaming on the Edge of 5G," *IEEE Transactions on Industrial Informatics*, 2020.

[26] C. Li, L. Toni, J. Zou, H. Xiong, and P. Frossard, "QoE-Driven Mobile Edge Caching Placement for Adaptive Video Streaming," *IEEE Transactions on Multimedia*, 2018.

[27] K. Sun, H. Zhang, Y. Gao, and D. Wu, "Delay-aware fountain codes for video streaming with optimal sampling strategy," *Journal of Communications and Networks*, 2019.

[28] N. Carlsson, D. Eager, V. Krishnamoorthi, and T. Polishchuk, "Optimized Adaptive Streaming of Multi-video Stream Bundles," *IEEE Transactions on Multimedia*, 2017.

[29] R. Mahmud, S. N. Srirama, K. Ramamohanarao, and R. Buyya, "Quality of Experience (QoE)-aware placement of applications in Fog computing environments," *Journal of Parallel and Distributed Computing*, 2019.

[30] T. 38.213, *NR:Physical layer procedures for control*. 3GPP TSG RAN: V15.0.0, 2018.

[31] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3gpp nr at mmwave frequencies," *IEEE Communications Surveys Tutorials*, 2019.

[32] T. Lattimore and C. Szepesvári, *Bandit Algorithms*. Cambridge University Press, 2020.

[33] L. Li, W. Chu, J. Langford, and R. Schapire, "A contextual-bandit approach to personalized news article recommendation," *Proceedings of the 19th international conference on World wide web*, 2010.

[34] D. Bouneffouf and R. Féraud, "Multi-armed bandit problem with known trend," *Neurocomputing*, 2016.

[35] S. Ontañón, "Combinatorial Multi-armed Bandits for Real-Time Strategy Games," *ArXiv:1710.04805*, 2017.

[36] W. Chen, Y. Wang, and Y. Yuan, "Combinatorial Multi-Armed Bandit: General Framework and Applications," *Proceedings of the 30th International Conference on Machine Learning*, 2013.

[37] Youtube Help, "Recommended Upload Encoding Settings for Youtube," https://support.google.com/youtube/answer/1722171?hl=en, [Online].

[38] R. N. Adams, "4K Gaming Used By Tiny Percent Of Players According to Steam," https://techraptor.net/gaming/news/4k-gaming-still-in-minority-according-to-steam, 2019, [Online].