Joint User Association and Wireless Scheduling with Smaller Time-Scale Rate Adaptation

Xiaoyi Wu* Jing Yang* Huacheng Zeng† Bin Li*
*Department of Electrical Engineering, Pennsylvania State University, University Park, PA, USA
†Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, USA

***Abstract*—Rate adaptation is a key mechanism in current IEEE 802.11 networks and next-generation cellular systems. Observing that the operating time scale of rate adaptation is usually much smaller than the user association and scheduling, we study a joint design of wireless user association and scheduling and rate adaptation with different time scales to maximize cumulative system throughput while guaranteeing desired fairness among users. We develop a maximum-weight type user association and scheduling algorithm that combines the virtual queues (tracking the scheduling debt for each user to ensure the desired fairness guarantee) and Upper Confidence Bound (UCB) estimates in its weight measure; each selected user then adopts the UCB algorithm to perform rate adaptation in a smaller time scale. We show that our proposed algorithm yields a cumulative regret growing with the square root of the time horizon up to a logarithmic factor, and achieves zero cumulative fairness violation after a certain number of time frames. We demonstrate the efficiency of the proposed algorithm via simulations using synthetic and realistic data traces.**

I. Introduction

Multiple access points (AP) are typically deployed to ensure sufficient communication capacity for reliable and fast wireless transmissions in crowded areas such as campuses, stadiums, airports, subways, and shopping centers. Current APs are equipped with rate adaptation capability that allows the transmitter to adapt the transmission rate, via different channel coding and modulation schemes, to the time-varying wireless channel, which significantly improves system throughput. Moreover, the rate adaptation will be a key physical-layer mechanism for next-generation millimeter-wave (mmWave) communication systems that typically have a large and unpredictable throughput fluctuation (e.g., [1]). Rate adaptation typically operates every 100ms in IEEE 802.11 systems [2] and on a much smaller time scale (e.g., less than 10ms or even 1ms) in mmWave-based wireless systems. Such an operation time scale is typically smaller than the transmission session of each user. This necessitates a joint user association and scheduling design and rate adaptation: determines when and which AP each user should associate with and is scheduled for wireless transmissions and then which transmission rate each selected user should choose on a small time scale during its scheduling period. The goal is to maximize system throughput while guaranteeing desired fairness among users (i.e., each user should be at least scheduled for a certain fraction of time on average).

This work has been supported in part by NSF under the grants CNS-2152610, CNS-2152658, CNS-2100112, CNS-1956276, and ECCS-2030026.

User association and scheduling design is important for efficiently managing interference in wireless networks and has received extensive research efforts (e.g., [3], [4], [5], [6], [7], [8]). However, the combination of joint user association and scheduling design and rate adaptation with different operation time scales is far less explored. In this paper, we consider the case that user association and scheduling decisions are made every T time slots. The user dissociation and association involve disconnection and reconnection and thus typically require non-trivial costs such as communication interruption and network delay increment. Intuitively, a small T provides more flexibility for APs to make association and scheduling decisions and thus could potentially improve network performance. However, it causes a more frequent dissociation and association and thus introduces a high handover cost. While the frame-based scheduling design (e.g., [9], [10]) for deadline-constrained traffic shares some similarities with our problem context, the user schedule is not fixed within the entire frame and thus is fundamentally different from our problem.

On the other hand, rate adaptation is a key mechanism for wireless communication systems to approach wireless channel capacity, and has been widely studied in the literature. Earlier works (e.g., [11], [12], [13]) developed heuristic rate selection algorithms to strike a balance between exploration and exploitation. In [14], the authors formulated rate adaptation as a Multi-Armed Bandit (MAB) problem, where each arm corresponds to a rate and is associated with an unknown link successful transmission probability. As such, all classical MAB algorithms, such as Upper Confidence Bound (UCB [15]), Kullback-Leibler UCB (KL-UCB [16]), and Thompson Sampling [17], can be directly utilized to develop rate adaptation algorithms with provable performance guarantees. The authors in [14] further developed a KL-UCB-based rate adaptation algorithm by exploiting the unimodal structure of the system throughput with respect to the transmission rate. Subsequent works [18], [19] further developed more efficient rate adaptation algorithms based on Thompson Sampling. However, all these rate adaptation algorithm designs focused on a single wireless link and have not yet been integrated into the wireless user scheduling design that operates on a much larger time scale.

In this paper, we develop a joint wireless user association and scheduling and rate adaptation algorithm with different operating time scales. The goal is to maximize the cumulative throughput over a finite time horizon while guaranteeing the desired fairness among users, i.e., each user is scheduled at least a certain fraction of time on average. Such a joint algorithm design shares a similarity with the combinatorial bandits with fairness constraints (e.g., [20], [21]), where each arm corresponds to a user and fairness is ensured among users. In particular, [20] introduced the virtual queues to address fairness constraints and incorporated it into the algorithm design that yields a cumulative regret growing with the square root of the time horizon up to a logarithmic factor while guaranteeing long-term fairness among users. [21] further developed an algorithm with regret that grows with the square root of the time horizon and agrees with the state-of-the-art instance-independent bound, and zero cumulative fairness violation after a certain time. Our problem setup differs from those works in the following aspects. First, we consider joint user association, scheduling, and rate adaptation with fairness constraints as opposed to user scheduling in [20], [21]. Second, our user association and scheduling are mainly based on the maximum weight-based algorithm, instead of MAB-based online learning. Rather, we utilize the MAB-based online learning approach to adapt the rate for each scheduled user. In other words, a MAB is embedded for each user in the joint user association and scheduling and rate adaptation framework. Third, user association and scheduling and rate adaptation operate at different time scales, which brings unique challenges for the algorithm design and theoretical analysis. To the best of our knowledge, this is the first work for handling wireless user association and scheduling and rate adaptation in different time scales. Our main contributions to this work are summarized as follows:

- We develop an online-learning-based joint user association and scheduling and rate adaptation that operate on different time scales (cf. Section III). In particular, a MaxWeight-type algorithm with the weight combining the virtual queues and UCB estimates is utilized to determine the user association and scheduling, while the UCB algorithm is employed to determine the transmission rate of each selected user.
- We show that our proposed algorithm achieves $O(\sqrt{K \log K})$ regret over K time frames while zero cumulative fairness violation can be achieved after a certain number of time frames independent of K.
- We demonstrate the superior performance of our proposed algorithm via simulations using synthetic data and the data collected in realistic wireless networks.

Note on Notation: We use bold and script font of a variable to denote a vector and a set, respectively. Let $\|\mathbf{x}\|_1$ and $\|\mathbf{x}\|$ denote the l_1 and l_2 norm of the vector $\mathbf{x}$, respectively. Let $f(x) = O(x)$ if $f(x) \leq Cx, \forall x \geq 0$ for some positive real number C.

II. SYSTEM MODEL

We consider a wireless system with N users and L access points (APs). We divide time into frames, each having T time slots. The access point makes the decision on serving users at the beginning of each time frame and serves the selected users over the entire frame. Here, a smaller frame size T allows the user to have more flexibility to switch the AP with better channel quality, but is at the cost of increasing the handover overhead. As such, the frame size T balances the handover cost and network performance. Due to the wireless interference constraints, only a subset of users can transmit simultaneously in each time frame. We define $S_{l,n}(kT) = 1$ if user n is associated with AP l for wireless transmission in time frame k, and $S_{l,n}(kT) = 0$ otherwise. We call $\mathbf{S}(kT) \triangleq (S_{l,n}(kT), \forall l, \forall n)$ the *feasible schedule* denoting the set of users that can be served by each AP simultaneously in frame k. Let $\mathcal{S}$ be the collection of all feasible schedules.

Within each time slot of a frame, each selected user transmits at the rate chosen from the set $\{r_1, r_2, \cdots, r_M\}$, where $0 < r_1 < r_2 < \cdots < r_M$ and M is the number of available transmission rates. Let $X_{l,n,m}(t) = 1$ denote that the wireless transmission of user n associated with AP l at rate r_m is successful in time slot t, and $X_{l,n,m}(t) = 0$ otherwise. Here, $r_m X_{l,n,m}(t)$ represents the throughput when user n associated with AP l transmits at rate r_m in time slot t. We assume that $X_{l,n,m}(t)$ is independently and identically distributed (i.i.d.) with an *unknown* mean $\mu_{l,n,m} \in [0,1]$. Let $I_{l,n,m}(t) = 1$ denote that user n associated with AP l transmits at rate r_m in time slot t, and $I_{l,n,m}(t) = 0$ otherwise. Hence, the received throughput of all users in frame k is $R(kT) \triangleq \sum_{t=kT}^{(k+1)T-1} \sum_{l,n,m} S_{l,n}(kT) r_m X_{l,n,m}(t) I_{l,n,m}(t)$.

Our goal is to maximize the cumulative expected throughput $\sum_{k=0}^{K-1} \mathbb{E}[R(kT)]$ while guaranteeing fairness among users, i.e., each user is scheduled at least $\lambda_n \in (0,1)$ fraction of time on average. If the statistics of throughput (i.e., $\mu_{l,n,m}, \forall l = 1, 2, \cdots, L, \forall n = 1, 2, \cdots, N, \forall m = 1, 2, \cdots, M$) are known in advance, then our objective can be achieved by choosing a randomized stationary schedule[1] $\{q^*(\mathbf{S}), \forall \mathbf{S} \in \mathcal{S}\}$, where $q^*(\mathbf{S})$ is the probability of selecting a feasible schedule $\mathbf{S}$ and solves the following optimization problem:

$$\max_{q(\mathbf{S})} \quad \sum_{\mathbf{S} \in \mathcal{S}} q(\mathbf{S}) \sum_{l=1}^{L} \sum_{n=1}^{N} S_{l,n} T \max_m r_m \mu_{l,n,m} \tag{1}$$

$$\text{s.t.} \quad \sum_{\mathbf{S} \in \mathcal{S}} q(\mathbf{S}) \sum_{l=1}^{L} S_{l,n} \geq \lambda_n + \delta, \forall n = 1, 2, \ldots, N, \tag{2}$$

where $\delta > 0$ is a "tightness" constant. Here, $\max_m r_m \mu_{l,n,m}$ is the maximum achievable throughput for user n communicating with AP l in each time slot, and thus $\sum_{l,n} S_{l,n} T \max_m r_m \mu_{l,n,m}$ is its maximum throughput in one time frame if the schedule $\mathbf{S} = (S_{l,n})$ is selected. In the rest of the paper, we let $\mathbf{S}^*$ denote the feasible schedule selected by the optimal randomized stationary schedule $q^*(\mathbf{S})$. Within each frame, each selected user n chooses $\mathbf{I}^* \triangleq (I^*_{l,n,m})_{l,n,m} \in \arg\max_{\mathbf{I}} \sum_{m=1}^{M} r_m \mu_{l,n,m} I_{l,n,m}$, i.e., selecting the rate with the maximum throughput, i.e., $\max_m r_m \mu_{l,n,m}$, in each time slot.

However, the throughput statistics are unknown *a priori* in practice. As such, each user needs to learn these statistics (also known as (a.k.a.) exploration) and selects the empirically best transmission rate so far (a.k.a. exploitation). This inevitably

[1]The existence of such a randomized stationary policy can be shown by using the similar argument in [22] and its proof is omitted for brevity.

leads to the throughput loss compared to the case when the throughput statistics are known beforehand. Our goal is to design a joint user scheduling and rate adaptation algorithm so that it not only meets the desired fairness requirement but also minimizes the cumulative regret over consecutive K time frames, which is the gap between the expected accumulated throughput and the optimal throughput with known throughput statistics, i.e.,

$$\begin{aligned}\mathrm{Reg}(KT) \triangleq \sum_{k,l,n} \mathbb{E}\left[S^*_{l,n} \sum_{t=kT}^{(k+1)T-1} \sum_{m=1}^{M} r_m \mu_{l,n,m} I^*_{l,n,m}\right] \\ - \sum_{k,l,n} \mathbb{E}\left[S_{l,n}(kT) \sum_{t=kT}^{(k+1)T-1} \sum_{m=1}^{M} r_m \mu_{l,n,m} I_{l,n,m}(t)\right].\end{aligned}$$

III. Algorithm Design and Performance Analysis

In this section, we develop an online-learning-based joint user scheduling and rate adaptation algorithm by integrating the key idea of the well-known UCB algorithm and virtual queue techniques while respecting the different time scales of user scheduling and rate adaptation. In particular, in the time scale for user schedule, the virtual queues are introduced to guarantee the desired fairness constraint (see [23] for an overview). In contrast, in the time scale for rate adaptation, the UCB approach is utilized to deal with the fundamental exploitation-exploration tradeoff in online learning for each user to identify the best transmission rate while achieving a minimum cumulative regret.

To ensure fairness among users, we maintain a virtual queue for each user that tracks its scheduling "debt" over time frames. In particular, let $Q_n(kT)$ be the virtual queue-length of user n at the beginning of time frame k, and its evolution over time frames is described as follows:

$$Q_n((k+1)T) = \left(Q_n(kT) + \lambda_n - \sum_{l=1}^{L} S_{l,n}(kT) + \epsilon_k\right)^+, \tag{3}$$

for $k = 0, 1, 2, \ldots$, where $(x)^+ \triangleq \max\{x, 0\}$ and $\epsilon_k > 0$ is some control parameter that will be specified later. We set $Q_n(0) = 0$ as the system starts at $k = 0$.

Let $H_{l,n,m}(t)$ be the number of time slots that user n is associated with AP l and transmits at rate r_m until time slot t, i.e., $H_{l,n,m}(t) \triangleq \sum_{\tau=0}^{t-1} S_{l,n}(\lfloor \tau/T \rfloor T) I_{l,n,m}(\tau)$, where $\lfloor x \rfloor$ denotes the maximum integer that is not greater than x. We set $H_{l,n,m}(0) = 0$ due to the fact that the system starts at $t = 0$. We use $\overline{\mu}_{l,n,m}(t)$ to denote the fraction of successful transmissions when user n is associated with AP l and transmits at rate r_m until time slot t, i.e., $\overline{\mu}_{l,n,m}(t) \triangleq \left(\sum_{\tau=0}^{t-1} S_{l,n}(\lfloor \tau/T \rfloor T) X_{l,n,m}(t) I_{l,n,m}(t)\right) / H_{l,n,m}(t)$. If $H_{l,n,m}(t) = 0$, we set $\overline{\mu}_{l,n,m}(t) = 1$. Let $w_{l,n,m}(t)$ denote the UCB estimate of user n associated with AP l using rate r_m in time slot t, which can be defined below:

$$w_{l,n,m}(t) \triangleq \min\left\{\overline{\mu}_{l,n,m}(t) + \sqrt{\frac{3\log t}{2H_{l,n,m}(t)}}, 1\right\}, \tag{4}$$

where $\sqrt{3\log t/(2H_{l,n,m}(t))}$ is the exploration bonus term that measures the uncertainty of the sample mean estimate $\overline{\mu}_{l,n,m}(t)$. Note that a smaller $H_{l,n,m}(t)$ implies less exploration on user n using rate r_m and thus more inaccuracy in the estimate $\overline{\mu}_{l,n,m}(t)$, in which case user n is encouraged to transmit at rate r_m for further exploration. In (4), we use the truncated version of the UCB estimate, since the successful transmission probability is at most 1. When $H_{l,n,m}(t) = 0$, we set $w_{l,n,m}(t) = 1$, i.e., if user n has not transmitted at rate r_m until time slot t, it should have the highest priority to be served.

On one hand, we would like to schedule users with large virtual queue-lengths in each time frame to meet the desired fairness constraint. On the other hand, in order to achieve a low cumulative regret, we prefer to schedule users and select their rates with large UCB weights in each time slot. This motivates the following online-learning-based joint user scheduling and rate adaptation algorithm, as shown in Algorithm 1.

Algorithm 1 Online-Learning-based Joint User Association and Scheduling and Rate Adaptation (OL-JUASRA) Algorithm

---

At the beginning of frame k, select a feasible schedule $\widehat{\mathbf{S}}(kT) \triangleq (\widehat{S}_{l,n}(kT), \forall l, \forall n)$ satisfying

$$\begin{aligned}\widehat{\mathbf{S}}(kT) \in \arg\max_{\mathbf{S}\in\mathcal{S}} \sum_{l,n} S_{l,n}\Big(Q_n(kT) \\ + \eta_k T \max_m r_m w_{l,n,m}(kT)\Big),\end{aligned}$$

where $\eta_k = \delta\sqrt{k}/T$ (with $\eta_0 = \delta/(2T)$). Then, update the virtual queue-lengths according to (3) with $\epsilon_k = (4r_M L N^{1.5} + 1)/(2\sqrt{k+1})$.

Within each time slot t in frame k, i.e., $t = kT, kT + 1, \ldots, (k+1)T - 1$, each selected user n associated with AP l (i.e., $\widehat{S}_{l,n}(kT) = 1$) chooses the rate index $\widehat{m}_n(t)$ (i.e., $\widehat{I}_{l,n,\widehat{m}_n(t)}(t) = 1$ and $\widehat{I}_{l,n,m}(t) = 0, \forall m \neq \widehat{m}_n(t)$) such that

$$\widehat{m}_n(t) \in \arg\max_m r_m w_{l,n,m}(t).$$

In the proposed OL-JUASRA algorithm, the increasing sequence $\{\eta_k\}_{k\geq 0}$ balances the virtual queue-lengths and the UCB estimates for throughput statistics over time frames. Initially, the OL-JUASRA algorithm puts a larger weight on the virtual queue-lengths to quickly guarantee desired fairness while learning the best transmission rate for each user, and then emphasizes more on the UCB weight to ensure a smaller cumulative regret. The parameter η_k requires the exact knowledge of the slackness constant, which is usually unavailable in practice. We will demonstrate that the OL-JUASRA algorithm with inaccurate slackness constants still performs well via simulations in Section IV. Different from prior works on combinatorial bandits with fairness constraints (e.g., [21]), the user scheduling and rate selection have different time scales. This requires carefully manipulating the virtual queue-lengths and UCB weights and decoupling them in an appropriate way in the performance analysis.

Next, we characterize the cumulative fairness violation of the proposed OL-JUASRA algorithm.

Proposition 1 (Cumulative Fairness Violation): Under the OL-JUASRA algorithm, if $\epsilon_k \leq \delta/2$, the cumulative fairness violation over K time frames can be upper bounded below:

$$\sum_{n=1}^{N}\left(\mathbb{E}\left[\sum_{k=0}^{K-1}\sum_{t=kT}^{(k+1)T-1}(\lambda_n - S_n(t))\right]\right)^+ \leq NT\left(g(N,\delta,r_M)-\sqrt{K}\right)^+,$$

where $g(N,\delta,r_M) = \frac{74N^{2.5}}{\delta}\log\left(\frac{18N}{\delta}\right) + (4r_M L + 3)N^{1.5} + \frac{N^{1.5}(6+\delta^2)}{\delta} + 1 + \frac{N^{1.5}(4r_M LN^{1.5}+1)^2}{\delta}\left(\frac{1}{\delta}+1\right)$ is a constant depending on system parameters such as N, δ and r_M.

Proof: We first select the Lyapunov function

$$V(kT) \triangleq \|\mathbf{Q}(kT)\|,$$

and prove that the Lyapunov function has an expected negative drift when $V(kT)$ is sufficiently large and its drift is absolutely bounded. Then according to [21, Lemma 11], $\mathbb{E}\left[\|\mathbf{Q}(kT)\|_1\right]$ can be upper bounded. Finally, we can derive the upper bound of the cumulative fairness violation by combining the dynamics of virtual queue-lengths and the analysis of $\mathbb{E}\left[\|\mathbf{Q}(kT)\|_1\right]$. The sketch of the proof is available in our technique report [24] due to the space limit. ■

Remarks 1: We can see from Proposition 1 that the OL-JUASRA algorithm achieves zero cumulative fairness violation when $K \geq g^2(N,\delta,r_M)$. Moreover, the number of frames required for achieving zero cumulative fairness violation is independent of frame size T and thus the required number of time slots for achieving zero cumulative fairness violation linearly increases with the frame size T. Furthermore, the amount of cumulative fairness violation linearly increases with the frame size T. All these observations will be demonstrated via simulations in Section IV.

We derive an upper bound on the cumulative regret under the OL-JUASRA algorithm.

Proposition 2 (Cumulative Regret): Under the OL-JUASRA algorithm with $\epsilon_k \leq \delta$, the cumulative regret $\mathrm{Reg}(KT)$ over K time frames can be upper bounded as follows:

$$\begin{aligned}\mathrm{Reg}(KT) \leq\ & \frac{Nr_M T(4r_M LN^{1.5}+1)^2}{4\delta^2} + 2\sqrt{K}NT(\delta + \frac{3}{2\delta})\\ &+ \frac{NT(2\delta+1)^2(4r_M LN^{1.5}+1)^3}{16\delta^4}\\ &+ LMNr_M\left(T+3+\frac{5\pi^2}{6}+(T+1)\log(KT)\right)\\ &+ (T+4)r_M\sqrt{6LMNS_{\max}KT\log(KT)}\\ &+ r_M\sqrt{\frac{3LMNS_{\max}KT}{2\log T}}\\ =\ & O\bigg(NT\sqrt{K} + LMNT\log(KT)\\ &\qquad + T\sqrt{LMNKT\log(KT)}\bigg).\end{aligned}$$

Proof: We perform the drift-plus-penalty analysis. Unlike prior work on the regret analysis (e.g., [21]), the different time scales of user scheduling and rate selection impose unique challenges on the corresponding regret bound analysis. In particular, we need to upper bound the regret by carefully decoupling the user decision and rate adaption in different time scales. Please see Appendix A for the detailed proof. ■

Remarks 2: For the impact of the number of frames K on the regret performance, our derived regret upper bound has the same order $O(\sqrt{K\log K})$ as the instance-independent upper bound for the classical UCB algorithm. While the derived regret upper bound increases with the frame size T, the simulations demonstrate that the frame size has a marginal impact on the regret performance. The reason is that each user has sufficient time to identify its best transmission rate under different frame sizes.

IV. Simulations

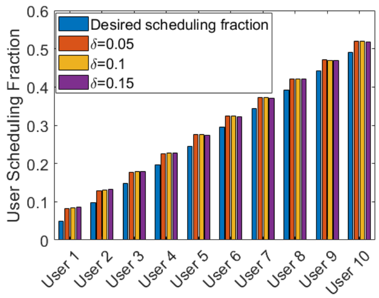
In this section, we evaluate the performance of our proposed OL-JUASRA algorithm via simulations based on synthetic and real-world data.

A. Synthetic Traces

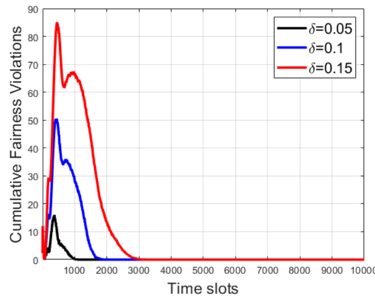
We consider $N = 10$ users and $L = 3$ APs in an IEEE 802.11g wireless system, where each user can associate at most one AP and each AP can schedule at most one user in each time frame. We let $\boldsymbol{\lambda} = \frac{2.7}{55} \times [1,2,3,4,5,6,7,8,9,10]$, where we recall that the n^{th} element denotes the desired scheduling fraction for user n. Each user has eight available rates as follows (in Mbps): $r_1 = 6$, $r_2 = 9$, $r_3 = 12$, $r_4 = 18$, $r_5 = 24$, $r_6 = 36$, $r_7 = 48$, and $r_8 = 54$, and its corresponding successful transmission probability vector $\boldsymbol{\mu_1} = [0.95, 0.9, 0.8, 0.65, 0.45, 0.25, 0.15, 0.10]$ at AP 1, $\boldsymbol{\mu_2} = [0.92, 0.87, 0.77, 0.62, 0.42, 0.22, 0.12, 0.07]$ at AP 2 and $\boldsymbol{\mu_3} = [0.90, 0.85, 0.75, 0.6, 0.4, 0.2, 0.1, 0.05]$ at AP 3.

Fig. 2 shows the impact of the frame size on the performance of the OL-JUASRA algorithm with an exact slackness constant (i.e., $\delta = 0.1$). We observe from Fig. 2a that the scheduling fraction of each user is larger than its desired value under different frame sizes, demonstrating that our algorithm achieves long-term fairness. Moreover, our algorithm can guarantee zero cumulative fairness violation, as shown in Fig. 2b, demonstrating that our algorithm can also achieve short-term fairness. The larger the frame size, the longer time required to achieve zero cumulative fairness violation and the larger amount of cumulative fairness violation. Furthermore, we can see from Fig. 2c that our algorithm achieves almost the same regret under different frame sizes, indicating that the impact of the frame size on the regret is marginal.

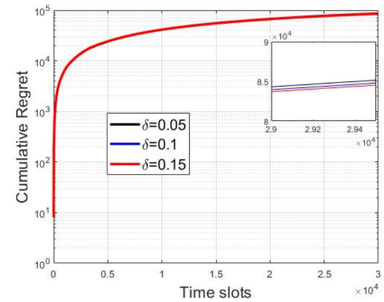
The parameter η_k in the OL-JUASRA algorithm requires the knowledge of the slackness constant δ, which is usually not available in practice. As such, we study the impact of the imperfect knowledge of the slackness constant on the system performance. In particular, we compare the performance of our algorithm with exact knowledge of the slackness (i.e., $\delta = 0.1$) and its 50% inaccurate estimates ($\delta = 0.05$ and $\delta = 0.10$) when



(a) User Scheduling Fraction

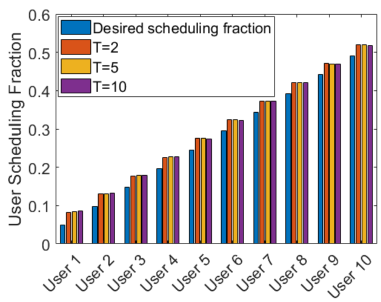


(b) Cumulative Fairness Violation

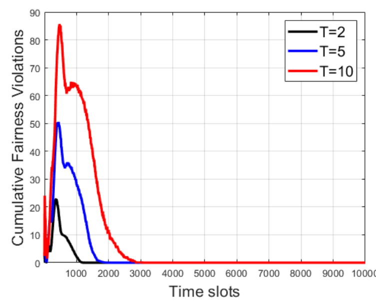


(c) Cumulative Regret

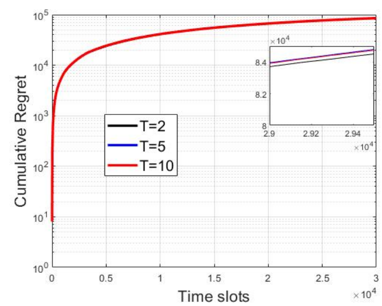
Fig. 1: Impact of slackness constant δ.



(a) User Scheduling Fraction



(b) Cumulative Fairness Violation



(c) Cumulative Regret

Fig. 2: Impact of frame size T.

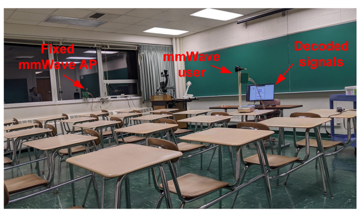


Fig. 3: Experimental setup in a classroom.

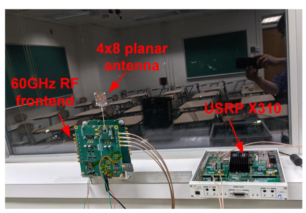


Fig. 4: A picture of mmWave AP.

the frame size T is set to 5, as shown in Fig. 1. From Fig. 1, we can see that the OL-JUASRA algorithm with inaccurate slackness constants can still guarantee long-term fairness (cf. Fig. 1a), zero cumulative fairness violation (cf. Fig. 1b), and logarithmic regret (cf. Fig. 1c). Moreover, the overestimate of the slackness constant (i.e., $\delta = 0.15$) leads to a larger cumulative fairness violation and a longer time to achieve zero cumulative fairness violation (cf. Fig. 1b), while it results in a slight improvement in cumulative regret (cf. Fig. 1c).

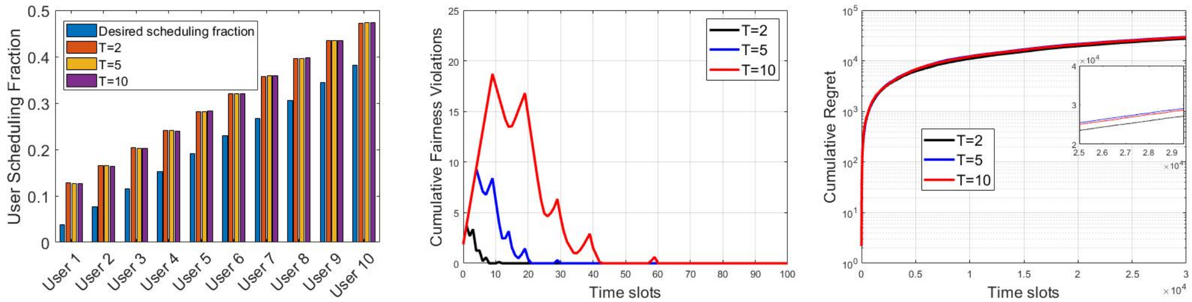
B. Real-World Data Traces

Experimental Setup: We consider a 60 GHz mmWave short-range communication network (e.g., 802.11ad [25]) in a classroom as shown in Fig. 3, where one AP was placed on the wall while 10 user devices are uniformly distributed over the whole classroom. Ideally, commodity off-the-shelf (COTS) 802.11ad devices should be used to evaluate our algorithm. However, COTS 802.11ad routers do not provide the control of rate adaptation. To overcome this challenge, we built a 60 GHz mmWave testbed to measure the end-to-end channel quality of communication traces and conduct simulations based on collected real-world data. The mmWave testbed consists of one transmitter and one receiver, each of which was built using a computer for baseband signal processing, a USRP X310 for signal shaping, and ADI's EVAL-HMC6300 Boards for signal frequency conversion. Simplified 802.11ad PHY-layer signal processing modules were implemented on the transmitter and receiver to conduct real-time data packet transmission from the AP to each user device.

TABLE I: EVM table specified in IEEE 802.11ad standard [25] (B: BPSK; Q: QPSK; 16Q: 16-QAM).

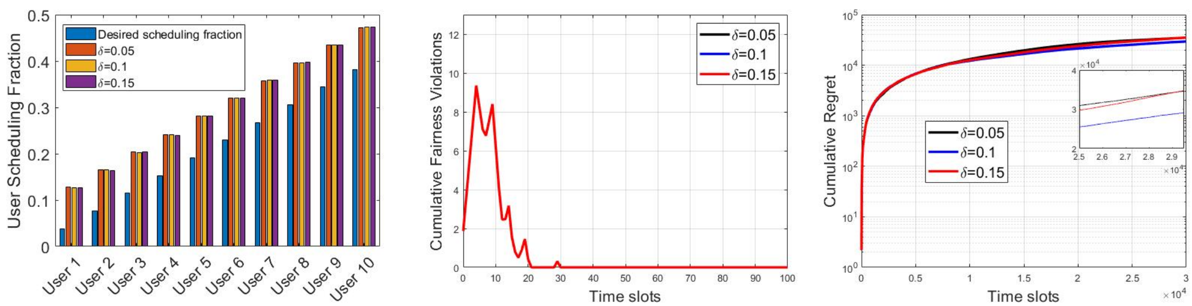
index (m)	1	2	3	4	5	6	7	8	9
postSNR (dB)	-7	-9	-10	-11	-12	-14	-15	-16	-17
Modulation	B	B	Q	Q	Q	Q	16Q	16Q	16Q
Coding rate	1/2	5/8	1/2	5/8	3/4	13/16	1/2	5/8	3/4
γ (postSNR)	0.5	0.63	1	1.25	1.5	1.63	2	2.5	3
Rate (Gbps)	0.73	0.91	1.46	1.825	2.19	2.37	2.92	3.65	4.38

Data Collection and Interpretation: We collect the post signal-to-noise ratio (post-SNR) of decoded signal constellations at each user device (receiver) every 1 ms for a total time duration of 20 seconds. Eventually, the collected data is a $30 \times 30,000$ matrix, with each element representing the instantaneous, end-to-end channel quality (i.e., postSNR) of the communication link from AP to a user device. We will publish our collected dataset for public access. Based on the measured postSNR, we calculate the achievable data rate of 802.11ad link as follows: $r(\text{postSNR}) = f \cdot \frac{\tau_{ofdm}}{\tau_{gi}+\tau_{ofdm}} \cdot \frac{N_{data}}{N_{fft}} \cdot \gamma(\text{postSNR})$, where $f = 2.64$GHz is the sampling rate, $\tau_{gi} = 36.36$ns is the normal guard interval, $\tau_{ofdm} = 194.56$ns is the OFDM symbol duration, $N_{data} = 336$ is the number of subcarriers for data, $N_{fft} = 512$ is the FFT size, and $\gamma(\text{postSNR})$ is the adaptive



(a) User Scheduling Fraction (b) Cumulative Fairness Violation (c) Cumulative Regret

Fig. 5: Impact of frame size T in Trace-based simulation.



(a) User Scheduling Fraction (b) Cumulative Fairness Violation (c) Cumulative Regret

Fig. 6: Impact of slackness constant δ in Trace-based simulation.

rate given in Table I.

Simulation Setup: We set each user's desired scheduling fraction the same as that in synthetic simulations, i.e., $\boldsymbol{\lambda} = \frac{2.1}{55} \times [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$. There are nine rates available for selection, which are shown in the last row of Table I. The data packet transmission result was estimated as follows. Suppose the AP selects r_m for a user device in the scheduling phase and the user device measures snr as its postSNR, if $r(snr) \geq r_m$, the data packet transmission is successful; otherwise, it fails.

Fig. 5 shows the performance of the OL-JUASRA algorithm using our collected wireless channel traces. We again observe from Fig. 5a that each user's scheduling fraction is larger than its desired value under different frame sizes, which shows that our algorithm can guarantee long-term fairness. In addition, as shown in Fig. 5b, our algorithm can also achieve zero cumulative fairness violation and thus yields short-term fairness. Similar to the synthetic simulations, a larger frame size results in a long time to achieve zero cumulative fairness and a larger amount of cumulative fairness violation. Moreover, the OL-JUASRA algorithm can achieve sublinear regret and the frame size has a negligible impact on the regret performance. Different from the synthetic simulations, the impact of the slackness constant δ has a marginal impact on cumulative fairness violation, as shown in Fig. 6.

V. Conclusion

In this paper, we studied the joint design of user association and scheduling and rate adaptation with different time scales in wireless networks to maximize cumulative throughput while guaranteeing desired fairness among users. We developed a MaxWeight-type user association and scheduling algorithm that combines both the virtual queues and UCB estimates in its weight measure. Each selected user utilizes the UCB algorithm to determine a transmission rate on a small time scale. We showed that our proposed algorithm yields $O(\sqrt{K \log K})$ cumulative regret over K time frames and achieves zero cumulative fairness violation after a certain number of time frames. We performed simulations to demonstrate the efficiency of the proposed algorithm based on both synthetic and real-world data.

Appendix A
Proof Sketch of Proposition 2

The proof relies on the following key lemma that establishes the relationship between the UCB weights at the beginning of a time frame and their values in the consecutive time slots within that frame.

Lemma 1: The gap between the UCB estimate in time slot t in frame k (i.e., $w_{l,n,m}(t)$ for $t = kT, kT+1, \ldots, (k+1)T-1$) and the estimate in the first time slot of frame k (i.e., $w_{l,n,m}(kT)$) can be bounded as follows:

$$\begin{aligned} & w_{l,n,m}(kT) \\ \geq & w_{l,n,m}(t) - \frac{T}{H_{l,n,m}(t)} - \frac{\sqrt{3}}{2\sqrt{2H_{l,n,m}(t)}\sqrt{\log T}} \end{aligned} \tag{5}$$

and

$$w_{l,n,m}(kT) \leq w_{l,n,m}(t) + \frac{T}{H_{l,n,m}(t)} + T\sqrt{\frac{3\log t}{2H_{l,n,m}(t)}}. \tag{6}$$

The proof of Lemma 1 requires to develop tight bounds between the UCB weights at the beginning of the frame and their values in each time slot within that frame. The sketch of proof is available in our technique report [24] due to the space limit.

Next, we develop the regret upper bound under our proposed OL-JUASRA algorithm. We rewrite the regret as follows.

$$\begin{aligned}
&\mathrm{Reg}(KT)\\
&\triangleq \sum_{k,l,n} \mathbb{E}\left[S^*_{l,n} \left(\sum_{t=kT}^{(k+1)T-1} \sum_{m=1}^{M} r_m \mu_{l,n,m} I^*_{l,n,m} \right) \right]\\
&- \sum_{k,l,n} \mathbb{E}\left[\widehat{S}_{l,n}(kT) \left(\sum_{t=kT}^{(k+1)T-1} \sum_{m=1}^{M} r_m \mu_{l,n,m} \widehat{I}_{l,n,m}(t) \right) \right]\\
&= \sum_{k=0}^{K-1} \Delta R(kT),
\end{aligned} \tag{7}$$

where $\Delta R(kT) \triangleq \sum_{t=kT}^{(k+1)T-1} \sum_{l,n,m} \mathbb{E}\big[r_m \mu_{l,n,m} S^*_{l,n} I^*_{l,n,m} - r_m \mu_{l,n,m} \widehat{S}_{l,n}(kT) \widehat{I}_{l,n,m}(t)\big]$. Select the Lyapunov function $V_1(\mathbf{Q}) \triangleq \frac{1}{2}\sum_{n=1}^{N} Q_n^2$ and consider its expected drift. In the rest of the proof, we omit the frame index kT associated with virtual queue lengths $\mathbf{Q}$ and schedule $\mathbf{S}$ without causing ambiguity.

$$\begin{aligned}
&\mathbb{E}[V_1(\mathbf{Q}((k+1)T) - V_1(\mathbf{Q}(kT))]\\
&\leq \sum_{n=1}^{N} \mathbb{E}\left[(\lambda_n + \epsilon_k) Q_n\right] - \sum_{l=1}^{L}\sum_{n=1}^{N} \mathbb{E}\left[Q_n \widehat{S}_{l,n}\right] + H_k,
\end{aligned} \tag{8}$$

where $H_k \triangleq N\left(3/2 + \epsilon_k^2\right)$.

Adding the term $\eta_k \Delta R(kT)$ on both sides of (8) and utilizing the fact that the optimal stationary randomized policy $\mathbf{S}^*(kT)$ is independent of the system state and stabilizes the system, i.e., $\mathbb{E}[\sum_{l=1}^{L} S^*_{l,n}(kT)] \geq \lambda_n + \delta, \forall n$, we have

$$\begin{aligned}
&\mathbb{E}[V_1(\mathbf{Q}((k+1)T) - V_1(\mathbf{Q}(kT))] + \eta_k \Delta R(kT)\\
&\leq \sum_{n} \mathbb{E}[(\lambda_n + \epsilon_k) Q_n] - \sum_{l,n} \mathbb{E}[Q_n \widehat{S}_{l,n}] + H_k\\
&+ \eta_k \sum_{t=kT}^{(k+1)T-1} \sum_{l,n,m} \mathbb{E}\big[r_m \mu_{l,n,m} S^*_{l,n} I^*_{l,n,m}\\
&\qquad - r_m \mu_{l,n,m} \widehat{S}_{l,n} \widehat{I}_{l,n,m}(t)\big]\\
&\leq H_k + \sum_{l,n} \mathbb{E}\left[\left(Q_n + \eta_k T \max_m r_m \mu_{l,n,m}\right)\left(S^*_{l,n} - \widehat{S}_{l,n}\right)\right]\\
&+ \eta_k \sum_{l,n,m} \sum_{t=kT}^{(k+1)T-1} \mathbb{E}\left[r_m \mu_{l,n,m}\left(I^*_{l,n,m} - \widehat{I}_{l,n,m}(t)\right)\widehat{S}_{l,n}\right],
\end{aligned} \tag{9}$$

where the last step follows from the fact the optimal stationary randomized policy $\mathbf{S}^*(kT)$ is independent of the system state and stabilizes the system, i.e., $\mathbb{E}[S^*_n(kT)] \geq \lambda_n + \delta, \forall n$ and holds for any $k \geq k_0 \triangleq (4r_M L N^{1.5} + 1)^2/4\delta^2$ such that $\epsilon_{k_0} \leq \delta$.

Dividing η_k on both sides of (9) and summing over $k = k_0, k_0+1, \ldots, K-1$, we have

$$\begin{aligned}
&\mathrm{Reg}(KT) = \sum_{k=0}^{k_0-1} \Delta R(kT) + \sum_{k=k_0}^{K-1} \Delta R(kT)\\
&\overset{(a)}{\leq} k_0 N r_M T + \sum_{k=k_0}^{K-1} \frac{H_k}{\eta_k} + \frac{1}{2\eta_{k_0}} N\left(k_0 + \sum_{k=0}^{k_0-1} \epsilon_k\right)^2\\
&+ \sum_{k=k_0}^{K-1} \frac{1}{\eta_k} \sum_{l,n} \mathbb{E}\Big[\left(Q_n + \eta_k T \max_m r_m \mu_{l,n,m}\right)\\
&\qquad\qquad \cdot \left(S^*_{l,n} - \widehat{S}_{l,n}\right)\Big]\\
&+ \sum_{k=k_0}^{K-1} \sum_{l,n,m} \sum_{t=kT}^{(k+1)T-1} \mathbb{E}\Big[r_m \mu_{l,n,m}\left(I^*_{l,n,m} - \widehat{I}_{l,n,m}(t)\right)\widehat{S}_{l,n}\Big]
\end{aligned} \tag{10}$$

where step (a) uses the fact that η_k is an increasing sequence and $Q_n(k_0 T) \leq k_0 \lambda_n + \sum_{k=0}^{k_0-1} \epsilon_k \leq k_0 + \sum_{k=0}^{k_0-1} \epsilon_k$.

Utilizing the scheduling component of the OL-JUASRA algorithm, we have

$$\begin{aligned}
&\sum_{l,n}\left(Q_n + \eta_k T \max_m r_m \mu_{l,n,m}\right)\left(S^*_{l,n} - \widehat{S}_{l,n}\right)\\
&\leq \eta_k T \sum_{l,n,m} r_m \left(w_{l,n,m}(kT) - \mu_{l,n,m}\right)^+\\
&\quad + \eta_k T \sum_{l,n,m} r_m \left(\mu_{l,n,m} - w_{l,n,m}(kT)\right)^+.
\end{aligned} \tag{11}$$

Similarly, we can bound

$$\sum_{t=kT}^{(k+1)T-1} \sum_{m} r_m \mu_{l,n,m}\left(I^*_{l,n,m} - \widehat{I}_{l,n,m}(t)\right).$$

Substituting (11) into (10), we have

$$\begin{aligned}
&\mathrm{Reg}(KT) \leq k_0 N r_M T + \sum_{k=k_0}^{K-1} \frac{H_k}{\eta_k} + \frac{N}{2\eta_{k_0}}\left(k_0 + \sum_{k=0}^{k_0-1} \epsilon_k\right)^2\\
&+ \underbrace{\sum_{k=k_0}^{K-1} \sum_{l,n,m} \sum_{t=kT}^{(k+1)T-1} r_m \mathbb{E}\left[\left(w_{l,n,m}(t) - \mu_{l,n,m}\right)^+\right]}_{\triangleq G_1(KT)}\\
&+ \underbrace{\sum_{k=k_0}^{K-1} \sum_{l,n,m} \sum_{t=kT}^{(k+1)T-1} r_m \mathbb{E}\left[\left(\mu_{l,n,m} - w_{l,n,m}(t)\right)^+\right]}_{\triangleq G_2(KT)}\\
&+ \underbrace{T \sum_{k=k_0}^{K-1} \sum_{l,n,m} r_m \mathbb{E}\left[\left(w_{l,n,m}(kT) - \mu_{l,n,m}\right)^+\right]}_{\triangleq G_3(KT)}\\
&+ \underbrace{T \sum_{k=k_0}^{K-1} \sum_{l,n,m} r_m \mathbb{E}\left[\left(\mu_{l,n,m} - w_{l,n,m}(kT)\right)^+\right]}_{\triangleq G_4(KT)}.
\end{aligned} \tag{12}$$

For the term $\sum_{k=k_0}^{K-1} H_k/\eta_k$, according to the definition of η_k, we can show that

$$\sum_{k=k_0}^{K-1} \frac{H_k}{\eta_k} \leq 2\sqrt{K}NT \left(\delta + \frac{3}{2\delta}\right). \tag{13}$$

The analyses of $G_1(KT)$ and $G_2(KT)$ are similar to the line of regret analysis in [20]. In particular, we can show that

$$G_1(KT) \leq LNMr_M \left(1 + \frac{\pi^2}{4}\right) + 2r_M\sqrt{6LNMS_{\max}KT\log(KT)}, \tag{14}$$

$$G_2(KT) \leq \frac{LNMr_M\pi^2}{6}. \tag{15}$$

For $G_3(KT)$ and $G_4(KT)$, we first utilize Lemma 1 and then upper bound $G_3(KT)$ and $G_4(KT)$ by $G_1(KT)$ and $G_2(KT)$, respectively. Thus we can obtain the following results.

$$G_3(KT) \leq LNMr_M \left(T + \frac{\pi^2}{4}\right) + LNMr_M T\log(KT) + (T+2)r_M\sqrt{6LNMS_{\max}KT\log(KT)}, \tag{16}$$

$$G_4(KT) \leq LNMr_M \left(2 + \frac{\pi^2}{6} + \log(KT)\right) + r_M\sqrt{\frac{3LMNS_{\max}KT}{2\log(T)}}. \tag{17}$$

Hence, by substituting (13), (14), (15), (16), and (17) into (12), we have the desired result.

REFERENCES

[1] A. Narayanan, E. Ramadan, R. Mehta, X. Hu, Q. Liu, R. A. Fezeu, U. K. Dayalan, S. Verma, P. Ji, T. Li *et al.*, "Lumos5g: Mapping and predicting commercial mmwave 5g throughput," in *Proceedings of the ACM Internet Measurement Conference*, 2020, pp. 176–193.

[2] R. Queiros, E. Almeida, H. Fontes, J. Ruela, and R. Campos, "Wi-fi rate adaptation using a simple deep reinforcement learning approach," *arXiv preprint arXiv:2202.03997*, 2022.

[3] Y. Bejerano, S.-J. Han, and L. Li, "Fairness and load balancing in wireless lans using association control," in *Proceedings of the 10th annual international conference on Mobile computing and networking*, 2004, pp. 315–329.

[4] H. Gong and J. Kim, "Dynamic load balancing through association control of mobile users in wifi networks," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 2, pp. 342–348, 2008.

[5] W. Li, S. Wang, Y. Cui, X. Cheng, R. Xin, M. A. Al-Rodhaan, and A. Al-Dhelaan, "Ap association for proportional fairness in multirate wlans," *IEEE/ACM Transactions On Networking*, vol. 22, no. 1, pp. 191–202, 2013.

[6] G. Athanasiou, P. C. Weeraddana, C. Fischione, and L. Tassiulas, "Optimizing client association for load balancing and fairness in millimeter-wave wireless networks," *IEEE/ACM Transactions on Networking*, vol. 23, no. 3, pp. 836–850, 2014.

[7] M. H. Dwijaksara, W. S. Jeon, and D. G. Jeong, "A joint user association and load balancing scheme for wireless lans supporting multicast transmission," in *Proceedings of the 31st Annual ACM Symposium on Applied Computing*, 2016, pp. 688–695.

[8] L. Sun, L. Wang, Z. Qin, Z. Ma, and Z. Yuan, "A novel on-line association algorithm in multiple-ap wireless lan," in *International Conference on Wireless Algorithms, Systems, and Applications*. Springer, 2017, pp. 890–902.

[9] I.-H. Hou, "Scheduling heterogeneous real-time traffic over fading wireless channels," *IEEE/ACM Transactions on Networking*, vol. 22, no. 5, pp. 1631–1644, 2013.

[10] J. J. Jaramillo, R. Srikant, and L. Ying, "Scheduling for optimal rate allocation in ad hoc networks with heterogeneous delay constraints," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 5, pp. 979–987, 2011.

[11] J. C. Bicket, "Bit-rate selection in wireless networks," Ph.D. dissertation, Massachusetts Institute of Technology, 2005.

[12] A. Kamerman and L. Monteban, "Wavelan®-ii: a high-performance wireless lan for the unlicensed band," *Bell Labs technical journal*, vol. 2, no. 3, pp. 118–133, 1997.

[13] M. Lacage, M. H. Manshaei, and T. Turletti, "Ieee 802.11 rate adaptation: a practical approach," in *Proceedings of the 7th ACM international symposium on Modeling, analysis and simulation of wireless and mobile systems*, 2004, pp. 126–134.

[14] R. Combes, J. Ok, A. Proutiere, D. Yun, and Y. Yi, "Optimal rate sampling in 802.11 systems: Theory, design, and implementation," *IEEE Transactions on Mobile Computing*, vol. 18, no. 5, pp. 1145–1158, 2018.

[15] P. Auer, N. Cesa-Bianchi, and P. Fischer, "Finite-time analysis of the multiarmed bandit problem," *Machine learning*, vol. 47, no. 2-3, pp. 235–256, 2002.

[16] A. Garivier and O. Cappé, "The KL-UCB algorithm for bounded stochastic bandits and beyond," in *Proceedings of the 24th annual conference on learning theory*, 2011, pp. 359–376.

[17] S. Agrawal and N. Goyal, "Analysis of thompson sampling for the multi-armed bandit problem," in *Conference on learning theory*, 2012, pp. 39–1.

[18] H. Gupta, A. Eryilmaz, and R. Srikant, "Low-complexity, low-regret link rate selection in rapidly-varying wireless channels," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*. IEEE, 2018, pp. 540–548.

[19] ——, "Link rate selection using constrained thompson sampling," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019, pp. 739–747.

[20] F. Li, J. Liu, and B. Ji, "Combinatorial sleeping bandits with fairness constraints," *IEEE Transactions on Network Science and Engineering*, 2019.

[21] X. Liu, B. Li, P. Shi, and L. Ying, "An efficient pessimistic-optimistic algorithm for constrained linear bandits," *arXiv preprint arXiv:2102.05295*, 2021.

[22] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic power allocation and routing for time varying wireless networks," in *IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No. 03CH37428)*, vol. 1. IEEE, 2003, pp. 745–755.

[23] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," *Synthesis Lectures on Communication Networks*, vol. 3, no. 1, pp. 1–211, 2010.

[24] X. Wu, J. Yang, H. Zeng, and B. Li, "Joint user association and wireless scheduling with smaller time-scale rate adaptation," https://sites.psu.edu/binli/files/2023/06/WiOpt2023Report.pdf, 2023.

[25] "IEEE 802.11ad-2012," https://standards.ieee.org/ieee/802.11ad/4527/, Accessed: 08-July-2022.