*Computers, Materials & Continua* CMC, vol.65, no.3, pp.1987-2007, 2020

# Millimeter-Wave Concurrent Beamforming: A Multi-Player Multi-Armed Bandit Approach

**Ehab Mahmoud Mohamed[1, 2, *], Sherief Hashima[3, 4], Kohei Hatano[3, 5], Hani Kasban[4] and Mohamed Rihan[6]**

**Abstract:** The communication in the Millimeter-wave (mmWave) band, i.e., 30~300 GHz, is characterized by short-range transmissions and the use of antenna beamforming (BF). Thus, multiple mmWave access points (APs) should be installed to fully cover a target environment with gigabits per second (Gbps) connectivity. However, inter-beam interference prevents maximizing the sum rates of the established concurrent links. In this paper, a reinforcement learning (RL) approach is proposed for enabling mmWave concurrent transmissions by finding out beam directions that maximize the long-term average sum rates of the concurrent links. Specifically, the problem is formulated as a multiplayer multiarmed bandit (MAB), where mmWave APs act as the players aiming to maximize their achievable rewards, i.e., data rates, and the arms to play are the available beam directions. In this setup, a selfish concurrent multiplayer MAB strategy is advocated. Four different MAB algorithms, namely, $\epsilon$-greedy, upper confidence bound (UCB), Thompson sampling (TS), and exponential weight algorithm for exploration and exploitation (EXP3) are examined by employing them in each AP to selfishly enhance its beam selection based only on its previous observations. After a few rounds of interactions, mmWave APs learn how to select concurrent beams that enhance the overall system performance. The proposed MAB based mmWave concurrent BF shows comparable performance to the optimal solution.

**Keywords:** Millimeter wave (mmWave), concurrent transmissions, reinforcement learning, multiarmed bandit (MAB).

[1] Electrical Engineering Department, College of Engineering, Prince Sattam Bin Abdulaziz University, Wadi Addwasir, 11991, Saudi Arabia.

[2] Electrical Engineering Department, Faculty of Engineering, Aswan University, Aswan, 81542, Egypt.

[3] Computational Learning Theory Team, RIKEN-Advanced Intelligent Project, Fukuoka, 819-0395, Japan.

[4] Engineering Department, Nuclear Research Center, Egyptian Atomic Energy Authority, Cairo, 13759, Egypt.

[5] Faculty of Arts and Science, Kyushu University, Fukuok, 819-0395, Japan.

[6] Electronics and Electrical Communication Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf, 32952, Egypt.

[*] Corresponding Author: Ehab Mahmoud Mohamed. Email: ehab_mahmoud@aswu.edu.eg.

Received: 30 May 2020; Accepted: 10 July 2020.

## 1 Introduction

The communication in the millimeter-wave band, i.e., 30~300 GHz, is considered as a crucial enabler of fifth-generation (5G) and beyond 5G (B5G) wireless networks due to its swath of available unlicensed spectrum [Wang, Kong, Kong et al. (2018)]. This large chunk of available bandwidth enables multi-gigabits per second (Gbps) connectivity, which can support intensive bandwidth 5G/B5G applications such as virtual reality (VR), 3D video streaming, etc., [Huo, Dong, Xu et al. (2019)]. The ratified IEEE 802.11ad (ay) standards define a set of 60 GHz wireless network protocols, also known as WiGig for WLAN applications [Ghasempour, Silva, Cordeiro et al. (2017a)]. In this paper, we call a WiGig device for the mmWave based device, which has 60 GHz capability. However, mmWave band suffers from several technical challenges due to its high operating frequencies. About 28 (21.6) dB attenuation losses are predicted using 60 GHz over using the legacy 2.4 (5) GHz band, respectively, and oxygen absorption peaks at the 60 GHz band reaching about 15 dB/Km [Rappaport, Sun, Mayzus et al. (2013)]. Moreover, mmWave transmission is highly susceptible to shadowing and path blockage even a human body can obstruct the mmWave path [Rappaport, Xing, MacCartney et al. (2017)]. To overcome such tough channel conditions, WiGig standards advocate the use of antenna beamforming (BF) by means of antenna weight vectors (AWVs) using structured codebooks to increase the channel gains, especially for the non-line of sight (NLoS) paths. Also, they defined a medium access control (MAC) based exhaustive search analog BF as a suitable BF training mechanism for WiGig transmissions [Ghasempour, Silva, Cordeiro et al. (2017b)]. Recently, a lot of advanced BF strategies are proposed in the literature to reduce the incredible complexity of the exhaustive search BF training while obtaining a near performance [Ahmed, Khammari, Shahid et al. (2018)]. The high propagation losses accompanied by path blockage confines the WiGig AP coverage in a short-range. Thus, multiple numbers of WiGig access points (APs) should be installed to cover a target area with Gbps connectivity fully. However, mutual inter-beam interference among the constructed WiGig concurrent links affects the spatial reuse capability of the WiGig APs, and it degrades the total system rate of the WiGig WLAN [Mohamed, Sakaguchi and Sampei (2017a)]. Thus, an efficient concurrent BF strategy should be considered for enabling WiGig simultaneous transmissions, where mutual inter-beam interference should be relaxed when selecting the concurrent beams. Though exhaustively searching all combinations of concurrent beams and choosing the best configuration gives the optimal performance, it results in incredible complexity, and tremendous BF overhead gets it an infeasible solution in real scenarios.

In this paper, a machine learning (ML) tool will be used to address the problem of concurrent BF in mmWave networks efficiently. ML is a talented approach that can resolve many of the wireless communication challenges using different learning approaches, i.e., supervised learning, unsupervised learning, and reinforcement learning (RL) [Wang, Jiang, Zhang et al. (2020)]. In supervised learning, the task of the ML algorithms is to model the relation between labeled inputs and their corresponding labeled outputs. The most popular supervised learning problems are regression and classification techniques. This approach can be applied for channel estimation, modulation/demodulation, and spectrum sensing in cognitive radios (CRs), etc., In unsupervised learning, only the inputs are available for the machine, and the task is to find out the hidden patterns in the input data. This type of ML can be applied for users' behavior learning and classification, resource allocation and association, optimal cell deployment, etc. In RL, the RL algorithm does not have any prior information about the environment, and it tries to maximize its long-term reward based on the interactions/observations with/from the environment. The RL algorithm tries to compromise between exploiting the best action taken so far or exploring new actions, formally known as exploitation-exploration trade-off. Q-learning and Multiarmed bandit (MAB) are famous RL algorithms. RL can be used in users' behavior prediction, channel/relay/base station (BS) selections, handover decisions, etc., In this paper, mmWave concurrent BF is formulated as a multiplayer MAB problem. In this formulation, the distributed WiGig APs act as the multiple players aiming to maximize their long-term average data rates, i.e., the rewards. This is done by selecting the appropriate beam directions maximizing the received power, and undergo low mutual interference, where the available beam directions will act as the arms of the bandit. In this WiGig WLAN, all WiGig APs are operating autonomously without installing a central management entity or exchanging information among APs. This setting profoundly relaxes the complexity of the network central management operation, including the need for an AP controller (APC) entity, and it eliminates a large number of management frames that need to be exchanged among the APs. Moreover, no global synchronization is needed as the concurrent links are established autonomously within the WiGig WLAN. Based on this fully decentralized setting, the WiGig APs will play the game selfishly, i.e., each AP will select its beam direction independently from the other APs selections. At each round, every AP will learn to choose beam direction, maximizing its achievable reward only based on its previous interactions, i.e., beam selections and reward observations. Despite the selfish behavior of the WiGig APs, the WiGig APs successively learn actions that enhance the overall average sum rate of the WLAN. In this paper, four major MAB algorithms, namely $\epsilon$-greedy, upper confidence bound (UCB), Thompson sampling (TS), and exponential weight algorithm for exploration and exploitation (EXP3) [Wilhemi, Cano, Neu et al. (2019a)] are modified to be used independently by each WiGig AP to find out its best beam direction for concurrent WiGig transmissions. The main contributions of this paper can be summarized as follows:

- The mmWave concurrent BF is formulated as an optimization problem that maximizes the total sum rate of the WiGig links. Fully decentralized WiGig WLAN is considered where no information is exchanged among the WiGig APs, or a central management entity is installed inside it.
- A selfish multiplayer MAB model is introduced to efficiently address the problem where each WiGig AP will play the game independently irrespective of the beam selections/observations of the other APs. Towards that, four MAB algorithms, i.e., $\epsilon$-greedy, UCB, TS, and EXP3 are modified and exploited by WiGig APs to interact with the environment and select their appropriate beam directions concurrently. The implemented MAB algorithms will learn from previous observations to proactively enhance the overall system performance.
- Numerical analysis is conducted to compare the performances of the adopted MAB algorithms. Also, we will prove the effectiveness of the proposed multiplayer MAB

approach compared to the optimal performance comes from exhaustively searching all available concurrent beam combinations in different scenarios. Simulation results show that the proposed scheme has comparable performances to the optimal one in terms of average total sum rate and spatial reuse factor.

The rest of this paper is prepared as follows; Section 2 gives the related works. Section 3 previews the system model, including the used mmWave link model. Section 4 introduces the optimization problem formulation. In Section 5, the proposed multiplayer MAB solution is presented, including the four modified MAB algorithms. Numerical analysis is given in Section 6, followed by the concluded remarks in Section 7.

## 2 Literature review

The authors in Mohamed et al. [Mohamed, Sakaguchi and Sampei (2017b)] boosted the performance of concurrent transmissions in random access mmWave WLANs. They did so by proposing a greedy scheme that operates with Wi-Fi assisted WiGig WLANs to receive a sub-optimal solution. They utilized statistical learning using Wi-Fi received signal strengths (RSSs) to control the operation inside the WiGig WLAN and assist the establishment of WiGig concurrent links in random access scenarios. Also, a central control algorithm that selects the suboptimal AP and its suboptimal beam direction that is expected to have low interference with the existing links has been proposed. Our scheme is different as the proposed online learning approach will be utilized without any out-band/external assistance. Also, the proposed WLAN is autonomously operated, which is entirely different from the fully centralized WLAN architecture given in Mohamed et al. [Mohamed, Sakaguchi and Sampei (2017c)], where APC was an essential entity to manage its operation adequately. Another efficient concurrent BF technique is highlighted in Qiao et al. [Qiao, Shen, Mark et al. (2015a)] to increase the capacity of indoor mmWave channels in time division multiple access (TDMA) scenarios. The network sum rate is maximized through optimizing the concurrent beamformers to get rid of the mutual interference between APs. They proposed an iterative searching algorithm that suppresses the BF complexity and setup duration. Moreover, they introduced a codebook-based BF protocol that operates at the MAC layer to define the beam sets. Although this work is highly related to the work presented in this paper, they assumed the existence of a piconet controller (PNC) that organizes the concurrent BF process among the distributed mmWave links. This is different from our setting, where a fully autonomous and decentralized WiGig WLAN is assumed. Accordingly, their scheme requires highly complicated management frames, the condition that is highly relaxed through the proposed MAB based approach. The mmWave blockage problem was investigated by proposing a solution using a coordinated multi-point reception (CoMP) scheme in Kumar et al. [Kumar, Saloranta, Kaleva et al. (2018)]. More precisely, this is done by using a trusted hybrid BF approach that deals with the unavailability of predominant linkages. A new coordinated BF technique that is based on stochastic optimization methods was also proposed in Gatzianas et al. [Gatzianas, Kalfas, Vagionas et al. (2019)], which is appropriate for highly dense urban mmWave networks. These techniques and other BF coordination approaches did not consider the problem of multi-point to multi-point concurrent BF, including mutual interference mitigations.

MmWave BF poses distinct difficulties because of the large available bandwidth and uncommon channel characteristics in addition to hardware restrictions. A brief survey of mmWave BF in indoor and outdoor scenarios is discussed in Kutty et al. [Kutty and Sen (2016)]. Due to its powerful capabilities, ML algorithms have been applied to overcome the propagation difficulties of mmWave. A brief overview of potential solutions to 5G/B5G problems from ML point of view are addressed in Chen et al. [Chen, Xiong, Xu et al. (2019); Li, Li, Zhang et al. (2019); Gui and Zeng (2020); Liu, Peng, Wang et al. (2019); Song, Yang, Xie et al. (2017); Zhang, Li, Wang et al. (2018); Cayamcela, Lee and Lim (2019)]. The problem of neighborhood discovery of mmWave based D2D communications was formulated as stochastic MAB in Hashima et al. [Hashima, Hatano, Takimoto et al. (2020)]. A group of MAB based algorithms was modified to reflect the remaining energies of the devices in the selection process, which improves the mmWave D2D link performance. A MAB training beam selection and Bayesian learning channel tracking approach for time-varying mmWave channels were proposed in Booth et al. [Booth, Suresh, Michelusi et al. (2019)]. The proposed techniques rapidly aligned the beams, which strongly supports such dynamic channels environments. The overhead produced from mmWave beam alignment is reduced using a proposed hierarchical beam alignment (HBA) algorithm [Wu, Cheng, Zhang et al. (2019)]. In such an algorithm, the BA problem was formulated as a stochastic MAB that attains to maximize the long-term RSS during a specific time. HBA selects the optimal beam with considerably low time compared to standard BA techniques in addition to latency reduction. Another BA optimization solution based on MAB for mmWave systems was proposed in Hashemi et al. [Hashemi, Sabharwal, Koksal et al. (2018)]. Besides formulating the problem as online stochastic MAB, the authors proposed an equivalent structured MAB model that optimally solves the problem. For BF of high-speed trains (HSTs), a mmWave MAB inspired beam searching algorithm, that reduces the overhead and searching trials as much as possible, was proposed in Wang et al. [Wang, Cheng, Wu et al. (2018)]. A novel efficient ML coordinated BF algorithm suitable for highly mobile mmWave scenarios was proposed in Alkhateeb et al. [Alkhateeb, Alex, Varkey et al. (2018)], where a deep learning (DL) model is trained to predict the BF vectors at the BSs. A DL based beam selection that is adaptable with 5G standards was proposed in Sim et al. [Sim, Lim, Park et al. (2020)], in which a deep neural network (DNN) is utilized to calculate a power delay profile of mmWave channel. In Aykin et al. [Aykin, Akgun, Feng, et al. (2020)], a MAB framework assigned for beam tracking in mmWave systems was suggested. The authors proposed an adaptive TS that selects suitable beams and its transmission rates by making use of former beam quality information. To the best of our knowledge, despite the existing ML applications in mmWave transmissions stated above, modeling mmWave concurrent BF using autonomously operating WiGig APs as a multiplayer MAB problem is first introduced in this paper.

## 3 System model

Herein, we will explain the network architecture of the WiGig WLAN under consideration in addition to the mmWave link model, including antenna BF gain.

### *3.1 WiGig WLAN network architecture*

Assume a WiGig WLAN network architecture that contains an $M$ WiGig APs-user equipments (UEs) links. Each of the *M* WiGig UEs is associated with one of the *M*

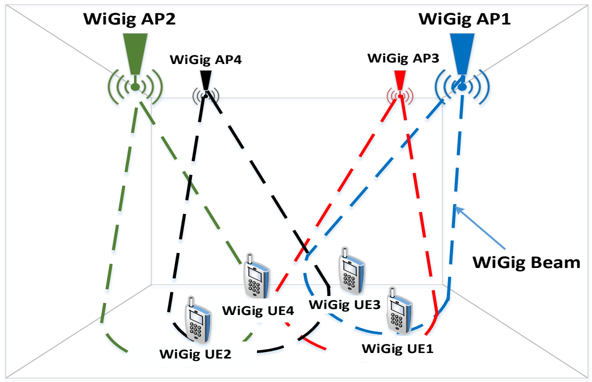


**Figure 1:** WiGig WLAN architecture

WiGig APs, as shown in Fig. 1, where 4 WiGig UEs are associated with 4 WiGig APs. The WiGig APs are operating autonomously and independently from each other. Also, the UEs are associated with the APs based on maximum received power. Without loss of generality, downlink transmission is assumed, and WiGig APs are able to send directional transmissions using antenna BF while UEs are using quasi-Omni antenna patterns. Conventionally, due to the lack of coordination among the WiGig APs, antenna beams are selected based on the maximum received power criteria, leading to high mutual- interference among the established WiGig links. This causes a degradation of the total sum rate of the concurrent WiGig connections. To resolve this inter-beam interference, WiGig APs can use too sharp beams directed towards their associated UEs. However, this requires a large antenna array containing antenna elements that needs to search over huge number of beam patterns. Moreover, direct line-of-sight (LoS) beam may influence blockage, which necessitates the use of non-LoS (NLoS) beams that might experience high mutual interference with the other concurrent links. Thus, improving the concurrent BF process via selecting low mutual-interfering concurrent beams seems to be a more practical solution valid for any type of antenna arrays, even using wide beams.

### *3.2 WiGig link model*

In this paper, we will utilize the mmWave 3D channel model mentioned in the IEEE 802.11ad standard and used by the authors in Mohamed et al. [Mohamed, Sakaguchi and Sampei (2017d)]. In this model, the channel response, including the BF gain, is given as:

$$g(\tau) = \int_0^{2\pi} \int_0^{\pi} G(\theta - \theta_b, \varphi - \varphi_b) h(\theta, \varphi, \tau) \sin(\theta)\, d\theta d\varphi \quad (1)$$

where $G(\theta - \theta_b, \varphi - \varphi_b)$ is the 3D BF gain. $\theta$ and $\varphi$ are the elevation and azimuth directions, and $\theta_b$ and $\varphi_b$ are the boresight angles of the directed beam. $h(\theta, \varphi, \tau)$ is the multi-path channel represented by:

$$h(\theta, \varphi, \tau) = \sum_{l=1}^{L} \rho_l \delta(\tau - \tau_l)(\theta - \theta_l)(\varphi - \varphi_l), \quad (2)$$

where $L$ is the total number of paths, $l = 1$ indicates the LoS path and $2 \leq l \leq L$ represents the NLoS paths. $\rho_l, , \tau_l, \theta_l$ and $\varphi_l$ represent the gain, delay, elevation, and azimuth angles of path $l$. Based on the channel model given in Eq. (1), the received power $P_r$ including the LoS blockage, which is modeled as a Bernoulli random variable (RV) [Mohamed, Elhalawany, Khallaf et al. (2020)], can be expressed as:

$$P_r = P_t\left(\mathcal{X}_{LOS}|g_1(\tau)|^2 + (1-\mathcal{X}_{LOS})\overline{|g_{2\leq l\leq L}(\tau)|^2}\right) \tag{3}$$

where $P_t$ is the transmit power, and $\mathcal{X}_{LOS}$ is Bernoulli RV that either equals 1 for LoS with probability $p$ or zero for NLoS with probability 1-$p$. $|g_1(\tau)|^2$ is the channel gain of the LoS path while $\overline{|g_{2\leq l\leq L}(\tau)|^2}$ is the average channel gain over the NLoS paths. For the mmWave BF gain, we will utilize the steering antenna model defined in IEEE 802.11ad, which is based on circularly symmetric Gaussian function to identify the main loop, represented as:

$$G(\theta-\theta_b,\varphi-\varphi_b)[dB] = G_0[dB] - \min\left[-\left(G_H(\varphi-\varphi_b)+G_V(\theta-\theta_b)\right), A\right], \tag{4}$$

$$A[dB] = 12 + G_0[dB], \tag{5}$$

$$G_0[dB] = 20\log_{10}\left(\frac{1.6162}{\sin\left(\frac{\theta_{-3dB}}{2}\right)}\right), \tag{6}$$

where $G_0[dB]$ is the maximum BF gain in decibel, and $G_H(\varphi-\varphi_b)$, $G_V(\theta-\theta_b)$ are the beam gains in horizontal and vertical directions, which are defined :

$$G_H(\varphi-\varphi_b) = -\min\left[12\left(\frac{\varphi-\varphi_b}{\varphi_{-3dB}}\right)^2, A\right], \qquad 0\leq \ \varphi \leq 2\pi \tag{7}$$

$$G_V(\theta-\theta_b) = -\min\left[12\left(\frac{\theta-\theta_b}{\theta_{-3dB}}\right)^2, A\right], \qquad 0\leq \ \theta \leq \pi \tag{8}$$

where $\varphi_{-3dB}$ and $\theta_{-3dB}$ are the half-power beamwidths of the azimuth and elevation angles.

### *4 Concurrent BF problem formulation*

Typically, BF is the process of adjusting the beam direction of the steerable antenna array to optimize a predefined cost function, e.g., signal-to-interference plus noise ratio (SINR) and the achievable data rate in consequence. In the case of concurrent BF, the selected beam direction of a WiGig link affects the achievable SINR of this link in addition to the attainable SINRs of the other links due to possible inter-beam interference. In this case, the cost function is to globally optimize the sum rate of all concurrent links taking mutual interference into account. Thus, the set of beams pattern of all $M$ links, i.e., $\{b_1, b_2, \dots, b_M\}$, should be selected to maximize the total sum rate of the concurrent connections, which can be expressed as:

$$\max_{S} \sum_{m=1}^{M} BW\log_2\left(1+\frac{P_{rm}(b_m)}{\sum_{k=1,k\neq m}^{M} P_{rm}(b_k)+N_0}\right) \tag{9}$$

$$\text{s.t} \qquad S\in\phi_S,\quad b_m, b_k\in\phi_B,\quad M\in\mathcal{Z}^+$$

where $BW$ is the assigned bandwidth and $N_0$ reflects the noise power. $P_{rm}(b_m)$ indicates the power received at the UE of link *m* from its corresponding AP using beam identification (ID) $b_m$, and $P_{rm}(b_k)$ indicates the power received at the UE of link *m* from the AP of link *k* using its beam ID $b_k$. $\phi_B$ refers to the available beam space of the WiGig AP. $P_{rm}(b_m)$ and $P_{rm}(b_k)$ can be calculated using Eq. (3) utilizing their related channel responses and the boresight angles ($\theta_b$,$\varphi_b$) of $b_m$ and $b_k$ for calculating the BF gains. $S=\{b_1,b_2,\ldots,b_M\}$ is the set of concurrent beam IDs, $\phi_S$ is the space of all available groups of concurrent beam IDs, and $\mathcal{Z}^+$ indicates the set of all positive integers. It is proved in Qiao et al. [Qiao, Shen, Mark et al. (2015b)] that the optimization problem in Eq. (9) mimics the Knaspasck situation [Zeng and Cremaschi (2018)], which is an NP-complete. Even concurrent BF is more difficult than Knaspasck problem because the rates of the links corresponding to a certain beams set $S$ are unknown unless this beams configuration is implemented. Exhaustively searching all available beam sets results in obtaining the optimal solution, but it requires exponential complexity of $|\phi_S| = |\phi_B|^M$. This exponential increase with respect to the number of concurrent links makes the exhaustive search an infeasible solution. Instead, the authors in Qiao et al. [Qiao, Shen, Mark et al. (2015c)] addressed this problem using an iterative search algorithm. At each round, BF training is done by one of the links while the other links using their previously selected beams. Then, the iterative search is conducted one by one till the convergence of the set of selected concurrent beams is achieved. Although this proposal relaxes the complexity of the optimization problem, still it needs a considerable amount of BF training till the set of concurrent beams converges. Also, this scheme mandates the use of PNC to fully control the WiGig network using a large number of management frames. Instead, in this paper, a RL based solution will be introduced using multiplayer MAB.

## 5 Proposed multi-player MAB approach

Herein, we will devise four MAB algorithms namely, $\epsilon$-greedy, UCB, TS and EXP3 to be utilized by each WiGig AP to selfishly select its concurrent beam.

### *5.1 Proposed multi-player MAB based mmWave concurrent BF*

Because all WiGig APs/UEs are operating autonomously without either using APC/PNC or permitting information exchange, selfish multiplayer MAB will be employed to achieve the sub-optimal set of concurrent beams. Specifically, a MAB algorithm will be implemented in each WiGig AP to interact with the environment independently, and timely enhance its concurrent beam selection based on its successive observations. In the proposed MAB modeling, a WiGig AP *m* is acting as the player trying to maximize its own long- term profit at each time *t* via playing over its available beam space, i.e., the arms of the bandit. This is done through utilizing its own observations irrespective of the other APs selections/observations. In this scenario, the profit of a WiGig link *m* is its achievable spectral efficiency in bps/Hz at time *t* using beam ID $b_{m,t}$, which can be expressed as:

$$\mathcal{R}_{b_{m,t}} = \log_2\left(1 + \frac{P_{rm}(b_{m,t})}{\sum_{k=1,k\neq m}^{M} P_{rm}(b_{k,t}) + N_0}\right), \tag{10}$$

where $b_{m,t}$ and $b_{k,t}$ are the selected beam IDs of links *m* and *k* at time *t*. In the following, we will adopt four famous MAB algorithms, $\epsilon$-greedy, UCB, TS, and EXP3, for performing concurrent BF selfishly.

**Algorithm 1:** $\epsilon$-greedy based mmWave concurrent BF

**Inputs:** $\phi_B$, $\epsilon$

**Initialize:** $t = 0,\ \bar{\mathcal{R}}_{b_{m,t}} = 0, x_{b_{m,t}} = 0, 1 \leq b_m \leq |\phi_B|$

**For** $t = 1{:}T$

1. Draw a beam ID, and obtain the reward:
   - $b^*_{m,t} = \begin{cases} \arg \max\limits_{1 \leq b_m \leq |\phi_B|} \left(\bar{\mathcal{R}}_{b_{m,t-1}}\right) & \text{with probability } 1-\epsilon \\ \mathcal{U}(1, |\phi_B|) & \text{with probability } \epsilon \end{cases}$
   - Obtain $\mathcal{R}_{b^*_{m,t}}$
2. $x_{b^*_{m,t}} = x_{b^*_{m,t-1}} + 1$
3. $\bar{\mathcal{R}}_{b^*_{m,t}} = \frac{1}{x_{b^*_{m,t}}} \sum_{j=1}^{x_{b^*_{m,t}}} \mathcal{R}_{b^*_{m,j}}$

**END For**

*5.1.1 Proposed $\epsilon$-greedy based mmWave concurrent BF*

$\epsilon$-greedy is the simplest MAB algorithm in dealing with exploitation-exploration trade-off. At each time *t* in the time horizon *T*, the algorithm selects with a probability $1-\epsilon$ the best arm having the highest average reward up to (but excluding) time *t*, and it explores random arm from the available arm space with a probability $\epsilon$, where $\epsilon$ is a design parameter. Algorithm 1 gives the proposed $\epsilon$-greedy based concurrent BF MAB algorithm implemented in each WiGig AP, where the inputs to the algorithm are the beam space $\phi_B$ and the value of $\epsilon$. At each time *t*, the beam ID for AP *m* is selected based on the following criteria:

$$b^*_{m,t} = \begin{cases} \arg \max\limits_{1 \leq b_m \leq |\phi_B|} \left(\bar{\mathcal{R}}_{b_{m,t-1}}\right) & \text{with probability } 1-\epsilon \\ \mathcal{U}(1, |\phi_B|) & \text{with probability } \epsilon \end{cases} \tag{11}$$

where with a probability $1-\epsilon$, beam ID $b^*_{m,t}$ is set to that giving the maximum average spectrum efficiency up to (but excluding) time *t* or it is dropped randomly from uniform distribution $\mathcal{U}(1, |\phi_B|)$, otherwise. After selecting $b^*_{m,t}$ for constructing the concurrent link, its corresponding spectral efficiency $\mathcal{R}_{b^*_{m,t}}$ is observed. Then, its number of selections $x_{b^*_{m,t}}$ as well as its average achievable spectrum efficiency $\bar{\mathcal{R}}_{b^*_{m,t}}$ are updated for the next round of selection, as given in Steps 2 and 3 in Algorithm 1.

*5.1.2 Proposed UCB based mmWave concurrent BF*

UCB can efficiently address the exploitation-exploration trade-off by increasing the confidence of the selected arm [Wilhemi, Cano, Neu et al. (2019b)]. It compromises between the arms having the best average rewards and that less being explored when taking an arm choice decision. Algorithm 2 gives the proposed UCB based mmWave concurrent BF algorithm, where the input to the algorithm is the available beam space

**Algorithm 2:** UCB based mmWave concurrent BF

**Inputs:** $\phi_B$

**Initialize:** each $b_m$, $1 \leq b_m \leq |\phi_B|$, will be selected once, and its corresponding $\mathcal{R}_{b_{m,t}}$ is evaluated, $1 \leq t \leq |\phi_B|$.

**For** $t = |\phi_B| + 1{:}T$

1. Draw a beam ID and obtain the reward:
   - $b^*_{m,t} = \arg \max_{1 \leq b_m \leq |\phi_B|} \left( \bar{\mathcal{R}}_{b_{m,t-1}} + \sqrt{\frac{2\ln(t)}{x_{b_{m,t-1}}}} \right)$
   - Obtain $\mathcal{R}_{b^*_{m,t}}$
2. $x_{b^*_{m,t}} = x_{b^*_{m,t-1}} + 1$
3. $\bar{\mathcal{R}}_{b^*_{m,t}} = \frac{1}{x_{b^*_{m,t}}} \sum_{j=1}^{x_{b^*_{m,t}}} \mathcal{R}_{b^*_{m,j}}$

**END For**

$\phi_B$. For initialization, every beam ID is selected once, and its corresponding $\mathcal{R}_{b_{m,t}}$ is observed where $1 \leq t \leq |\phi_B|$. After initialization, the beam ID maximizing the following equation is selected by WiGig AP *m* at each time *t*:

$$b^*_{m,t} = \arg \max_{1 \leq b_m \leq |\phi_B|} \left( \bar{\mathcal{R}}_{b_{m,t-1}} + \sqrt{\frac{2\ln(t)}{x_{b_{m,t-1}}}} \right), \ |\phi_B| + 1 \leq t \leq T, \quad (12)$$

where $\bar{\mathcal{R}}_{b_{m,t-1}}$ is the average spectrum efficiency resulted from using beam ID $b_m$ up to (but excluding) time *t*, which represents the exploitation term in the UCB equation. Yet, the term $\sqrt{\frac{2\ln(t)}{x_{b_{m,t-1}}}}$ indicates the exploration term in the UCB hypothesis, where $x_{b_{m,t-1}}$ indicates the number of time beam ID $b_m$ was selected. The idea behind UCB is to compromise between selecting the beam ID having maximum average spectrum efficiency or exploring new less investigated ones. After selecting beam ID $b^*_{m,t}$ to be played at time *t*, its corresponding number of selections $x_{b^*_{m,t}}$ and average spectral efficiency $\bar{\mathcal{R}}_{b^*_{m,t}}$ are updated accordingly for the next round of selection, see Steps 2 and 3 in Algorithm 2.

### *5.1.3 Proposed TS based mmWave concurrent BF*

TS is a Bayesian algorithm, where posterior distributions are constructed for the gained rewards based on a predefined probabilistic model. TS achieves good empirical performance with guarantees even better than those warranted by UCB, especially when the said model highly matches the actual distribution of the rewards. At the beginning of the TS algorithm, prior distributions are constructed for rewards based on parameter initialization of the said model. Then, the TS policy keeps track of the posterior distributions of the rewards based on the collected data during the learning process, which is achieved by updating the parameters of the probabilistic models. Algorithm 3 provides the proposed TS based mmWave concurrent BF, where gaussian distribution is

**Algorithm 3:** TS based mmWave concurrent BF

**Inputs:** $\phi_B$

**Initialize:** $t = 0,\ \bar{\mathcal{R}}_{b_{m,t}} = 0, x_{b_{m,t}} = 0, \sigma^2_{b_{m,t}} = 1.$

**For** $t = 1{:}T$

Sample $\Delta_{b_{m,t-1}}$, $1 \le b_m \le |\phi_B|$, from normal distributions $\mathcal{N}\left(\bar{\mathcal{R}}_{b_{m,t-1}}, \sigma^2_{b_{m,t-1}}\right)$

4 Draw a beam ID and obtain the reward:

- $b^*_{m,t} = \arg \max_{1\le b_m \le |\phi_B|} \left(\Delta_{b_{m,t-1}}\right)$
- Obtain $\mathcal{R}_{b^*_{m,t}}$

5 $x_{b^*_{m,t}} = x_{b^*_{m,t-1}} + 1$

6 $\bar{\mathcal{R}}_{b^*_{m,t}} = \frac{1}{x_{b^*_{m,t}}} \sum_{j=1}^{x_{b^*_{m,t}}} \mathcal{R}_{b^*_{m,j}}$

7 $\sigma^2_{b^*_{m,t}} = \frac{1}{x_{b^*_{m,t}}+1}$

**END For**

assumed for the spectrum efficiency, i.e., reward, obtained by each beam ID, i.e., $\mathcal{N}\left(\bar{\mathcal{R}}_{b_{m,t}}, \sigma^2_{b_{m,t}}\right)$, where $\bar{\mathcal{R}}_{b_{m,t}}$ and $\sigma^2_{b_{m,t}}$ are the mean and variance of the gaussian distribution. Based on the assumptions given by the authors in Wilhemi et al. [Wilhemi, Cano, Neu et al. (2019c)], $\bar{\mathcal{R}}_{b_{m,t}}$ is set to $\frac{1}{x_{b_{m,t}}} \sum_{j=1}^{x_{b_{m,t}}} \mathcal{R}_{b_{m,t}}$ and $\sigma^2_{b_{m,t}}$ is equal to $\frac{1}{x_{b_{m,t}}+1}$. Gaussian distribution is a reasonable assumption for modeling the distribution of spectral efficiency because usually, the received power has a normal distribution due to the effect of the additive wight Gaussian noise (AWGN) and the added random interference. A prior distribution is initialized for the rewards by initializing the values of $\bar{\mathcal{R}}_{b_{m,t}}$ and $\sigma^2_{b_{m,t}}$ as given in Algorithm 3. Then, samples $\Delta_{b_{m,t-1}}$ are taken from these distributions, and the beam ID having the maximum sample value is chosen to be played, as follows:

$$b^*_{m,t} = \arg \max_{1\le b_m \le |\phi_B|} \left(\Delta_{b_{m,t-1}}\right) \tag{13}$$

The spectral efficiency corresponding to $b^*_{m,t}$ is then observed and the values of $x_{b^*_{m,t}}$, $\bar{\mathcal{R}}_{b^*_{m,t}}$ and $\sigma^2_{b^*_{m,t}}$ are updated as given in Steps 5-7 in Algorithm 3. Based on the updated values, the posterior distributions of the rewards are enhanced during the learning process contributing in better beam ID selections over the time horizon.

*5.1.4 Proposed EXP3 based mmWave concurrent BF*

EXP3 is a weighted MAB algorithm, where higher weights are given to the best actions as the learning process proceeds. Weights' probabilities are calculated, and actions are randomly taken based on these probabilities. Algorithm 4 demonstrates the proposed

**Algorithm 4:** EXP3 based mmWave concurrent BF

**Inputs:** $\phi_B, \chi$

**Initialization:** $t = 0, \delta(t) = \delta_0, w_{b_{m,t}} = 1$ for $\forall b_m$

**For** $t = 1{:}T$

1. $\Omega_{b_{m,t}} \leftarrow (1-\chi)\frac{w_{b_{m,t}}}{\sum_{b_m=1}^{|\phi_B|} w_{b_{m,t}}} + \frac{\chi}{|\phi_B|}$
2. Draw a gateway UAV and obtain the reward:
   - $b_{m,t}^* \sim \Omega_{b_{m,t}} = \left(\Omega_{1,t}, \Omega_{2,t}, \ldots, \Omega_{|\phi_B|,t}\right)$
   - Obtain $\mathcal{R}_{b_{m,t}^*}$
3. $\bar{\mathcal{R}}_{b_{m,t}^*} = \frac{\mathcal{R}_{b_{m,t}^*}}{\Omega_{b_{m,t}^*}}$
4. $\delta(t) = \frac{\delta_0}{\sqrt{t}}$
5. $w_{b_{m,t+1}^*} = \left(w_{b_{m,t}^*}\right)^{\frac{\delta(t)}{\delta(t-1)}} \exp\left(\delta(t)\bar{\mathcal{R}}_{b_{m,t}^*}\right)$
6. $w_{b_{m,t+1}} = \left(w_{b_{m,t}}\right)^{\frac{\delta(t)}{\delta(t-1)}}, \ \forall b_m \neq b_m^*$

**END For**

EXP3 mmWave concurrent BF algorithm, where the inputs to the algorithm are the beam space $\phi_B$ and exploration parameter $\chi \in (0,1]$. For initialization, the weights of all available beam IDs are set to 1, and the learning rate of the algorithm is also initialized, i.e., $\delta(t) = \delta_0$. At every time t, a probability is given to each beam ID based on its allocated weight and the assigned exploration value $\chi$ as follows:

$$\Omega_{b_{m,t}} \leftarrow (1-\chi)\frac{w_{b_{m,t}}}{\sum_{b_m=1}^{|\phi_B|} w_{b_{m,t}}} + \frac{\chi}{|\phi_B|}, \tag{14}$$

Then, a beam ID is picked randomly based on these probabilities:

$$b_{m,t}^* \sim \Omega_{b_{m,t}} = \left(\Omega_{1,t}, \Omega_{2,t}, \ldots, \Omega_{|\phi_B|,t}\right), \tag{15}$$

After obtaining the spectrum efficiency $\mathcal{R}_{b_{m,t}^*}$ corresponding to the selected beam ID $b_{m,t}^*$, its weighted estimated value $\bar{\mathcal{R}}_{b_{m,t}^*}$ is evaluated as $\bar{\mathcal{R}}_{b_{m,t}^*} = \frac{\mathcal{R}_{b_{m,t}^*}}{\Omega_{b_{m,t}^*}}$ where dividing the actual reward by its probability value when calculating $\bar{\mathcal{R}}_{b_{m,t}^*}$ compensates the beam IDs that unlikely to be chosen. Following the same methodology given in Wilhemi et al. [Wilhemi, Cano, Neu et al. (2019d)], the weights of both the selected beam ID and the other beam IDs are updated as given in Steps 5 and 6 in Algorithm 4. Also, a time-dependent learning rate of $\delta(t) = \frac{\delta_0}{\sqrt{t}}$ is utilized to enhance the learning process, where large values of $\delta$ results in more confident update while small values lead to conservative behavior.

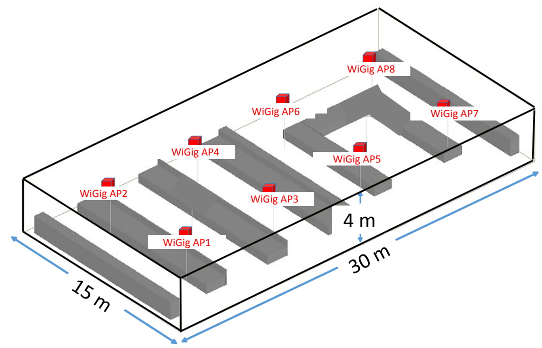


**Figure 2:** Ray tracing WiGig WLAN study area

**Table 1:** Simulation parameters

| Parameter | Value |
|---|---|
| Number of APs/UEs | 8 |
| UE antenna height | 0.75 m |
| $P_t$ | 10 dBm |
| $L$ | 10 |
| $BW$ | 2.16 GHz |
| $N_0$ | $-174 + 10\log_{10}(BW) + 10$ |
| $\varepsilon$ | 0.1 |
| $\delta_0$ | 0.4 |
| $\chi$ | 0.05 |

## 6 Numerical analysis

In this section, extensive numerical simulations are conducted to compare the performances of the proposed multiplayer MAB based WiGig concurrent BF algorithms. Also, the optimal performance results from exhaustively searching all available concurrent beams combinations are provided as a benchmark performance.

### *6.1 Simulation parameters*

In the simulation scenario, an indoor WLAN area of size $30 \times 15 \times 4$ m$^3$ is studied where eight WiGig APs are attached to the ceiling, as shown in Fig. 2. Ray tracing using wireless InSite software is used to generate the 60 GHz channels between WiGig APs and their associated UEs. Each UE is associated with one AP based on maximum received power. Thus, only $M$ WiGig links exist for concurrent transmissions, where $M$ is the total number of used APs. Other simulation parameters are given in Tab. 1 unless otherwise stated.

### 6.2 Performance metrics

The following metrics are used for performance comparisons:

- **Average total sum rate** of the concurrent links, which is equal to

$$\mathbb{R}_t = \frac{BW}{T} \sum_{t=1}^{T} \sum_{m=1}^{M} \mathcal{R}_{b_{m,t}}, \tag{16}$$

- **Spatial reuse factor** $\rho$, which is defined as:

$$\rho = \frac{\text{Total sum rate of the concurrent links}}{\text{Average rate of the isolated links}}, \tag{17}$$

Stronger mutual interference results in smaller values of $\rho$ and vice versa. Ideally, for eight coexisting links, $\rho$ should be equal to 8.

### 6.3 Simulation results

Fig. 3 shows the average total sum rate of the compared schemes against the number of concurrent links which are uniformly distributed inside the WLAN area, where LoS blockage probability of 0 and -3 dB beamwidths of $50^{\circ}$, i.e., 7 beam IDs, are used. As shown in this figure, as the number of concurrent links increases, the average total sum rate increases as well. However, the curves tend to saturate after using 7 concurrent links due to the increase in mutual interference. The proposed multiplayer MAB based concurrent BF algorithms show comparable performances to the optimal one, where TS and UCB provide the best performance while $\epsilon$-greedy gives the worst one. This comes from the Bayesian strategy and the confidential policy of TS and UCB, respectively. Yet, the worst performance of $\epsilon$-greedy comes from its intuitive online learning process based on allocating a fixed probability for both exploitation and exploration. When the number of concurrent links is equal to 8, about 91.7%, 91%, 87%, and 83.33% of the optimal performance are obtained using the proposed TS, UCB, EXP3, and $\epsilon$-greedy, respectively. This is done while the optimal performance needs to search over 5,764,801 concurrent beam combinations, but the proposed MAB schemes need to test only eight beams at a time. Fig. 4 demonstrates the spatial reuse factor of the compared schemes against the number of concurrent links with the same simulation setting used in Fig. 3. It is interesting to note that a near-optimal spatial reuse performance is obtained using the TS algorithm at all tested values of concurrent links. Also, UCB and EXP3 show a good spatial reuse performance compared to the optimal one. However, $\epsilon$-greedy shows weak performance due to the aforementioned reason, especially when increasing the number of concurrent links. Using two concurrent links, about 98%, 95%, 95%, and 93.16% of the optimal spatial reuse factor are obtained using TS, UCB, EXP3, and $\epsilon$-greedy, respectively. These values become 97%, 94.3%, 93.3%, and 90.8% when using 8 concurrent links.

Fig. 5 demonstrates the average total system rate of the compared schemes against the LoS blockage probability using 4 concurrent links by operating APs numbers 3, 4, 5 and 6, and -3dB beamwidth of $50^{\circ}$, i.e., using 7 beam IDs. Generally speaking, the average total sum rate of all compared schemes are decreasing as the LoS blockage probability is increasing. This is because lower power is received from the NLoS paths compared to the LoS ones. Likewise, the proposed MAB based schemes show comparable performances to the optimal one at all tested values of LoS blockage probability. TS and UCB give the best performances among the MAB schemes, while $\epsilon$-greedy shows the worst one. At LoS blockage probability of 0, bout 92.8%, 92.2%, 88%, and 83% of the optimal performance are obtained using the proposed TS, UCB, EXP3, and $\epsilon$-greedy, respectively. These values become 95.5%, 92%, 76%, and 67% at LoS blockage probability of 0.9. This confirms the superior performance of the proposed MAB based schemes, especially TS and UCB, even in a harsh blockage environment.

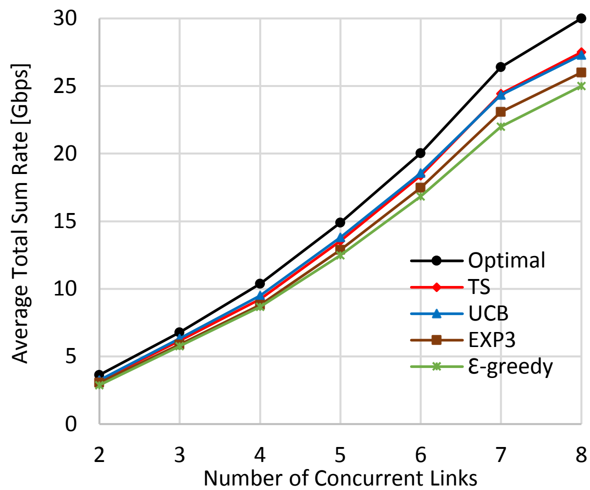


**Figure 3:** Average total system rate against the number of concurrent links using LoS blockage probability of 0 and -3dB beamwidth of 50$^\circ$

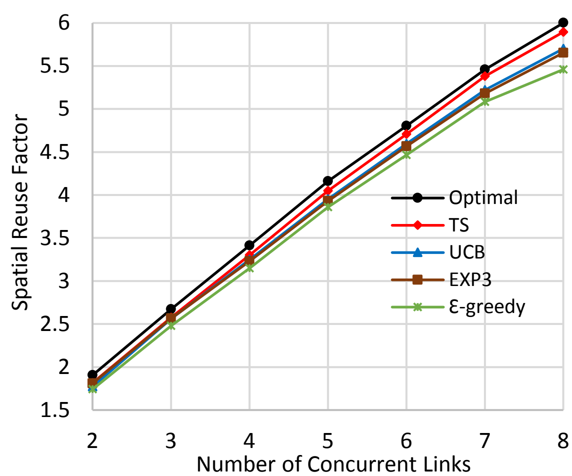


**Figure 4:** Spatial reuse factor against the number of concurrent links using LoS blockage probability of 0 and -3dB beamwidth of 50$^\circ$

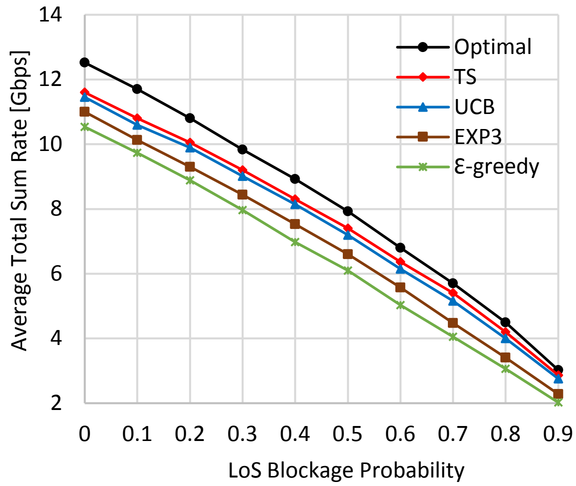


**Figure 5:** Average total system rate against the LoS blockage probability using 4 concurrent links and -3 dB beamwidth of $50^{\circ}$

Fig. 6 shows the average total sum rate performances of the compared schemes against the -3 dB beamwidth, i.e., using a different number of beam IDs ranging from 61 beams when -3 dB beamwidth is set to $20^{\circ}$ to 7 beams when the -3 dB beamwidth is set to $60^{\circ}$. In this simulation, LoS blockage probability of 0 and 4 concurrent links by operating APs number 3, 4, 5, and 6 are used. As shown in Fig. 6, as the -3 dB beamwidth is increasing, i.e., using wider beams, the average total sum rates of all compared schemes are decreasing accordingly. This is due to the low BF gain resulting from using wide beams, which reduces the received power in addition to increasing the mutual interference. The opposite happens when using narrow beams. Again, the proposed MAB based schemes show comparable performances to the optimal one, even in high interfering environment when using wide beams. Also, TS and UCB show the best performances, while $\epsilon$-greedy gives the worst one. At -3 dB beamwidth of $20^{\circ}$, i.e., using 61 beams, about 90.5%, 88.3%,80%, and 76% of the optimal performance are obtained using the proposed TS, UCB, EXP3, and $\epsilon$-greedy, respectively. This is done while the optimal performance requires to exhaustively search over $61^4$ different combinations of concurrent beams while the proposed MAB schemes need to only test 4 concurrent beams at a time. Even in a high interfering environment when the -3 dB beamwidth is equal to $60^{\circ}$, these values become 93%, 90.4%, 85%, and 83%. This means that WiGig APs learn to select concurrent beam IDs that enhance the overall system performance even in high interfering environment by only exploiting their own observations while selfishly interacting with the environment.

Fig. 7 illustrates the average total sum rate performances of the compared schemes against user density in user/m$^2$ using 8 concurrent links, LoS blockage probability of 0, and -3 dB beamwidth of $50^{\circ}$. To simulate the different user densities given in Fig. 7, we shrink the area within which the 8 UEs are distributed around the room center. For example, to obtain a user density of 0.02 user/m$^2$, the 8 users are distributed in the whole area of $30 \times 15$ m$^2$. However, to obtain a user density of 0.8 user/m$^2$, the 8 users are distributed in the area of

$5 \times 2\ m^2$ around the room center, and so on. Low user density means low mutual interference, while high user density indicates high mutual interference. As shown in Fig. 7, as the user density increases, the average total sum rates of all compared schemes are decreased due to the increase in mutual interference. Yet, the proposed MAB schemes show comparable performances to the optimal one even in harsh interfering environment, and TS is still showing the best performance. At low user density of 0.02 user/$m^2$, about 94.8%, 94.1%, 86.6%, and 86.2% of the optimal performance are obtained using TS, UCB, EXP3, and $\epsilon$-greedy, respectively. However, at suffered interfering environment of 0.8 user/$m^2$, these values become 80.14%, 76.12%, 71%, and 68%, respectively. Although the performances of the MAB schemes are decreased in high user density, they still show comparable performances to the optimal one considering the huge difference in the concurrent BF complexity and the strong interfering environment. Compared to the scheme proposed in Qiao et al. [Qiao, Shen, Mark et al. (2015d)], about 95% of the optimal performance is obtained using their proposed iterative search algorithm when using 8 concurrent links. This comes at the expense of about 45% of the normalized setup time consumed by the optimum exhaustive search approach. This high overhead is due to the several management frames needed for organizing the operation among the distributed concurrent links. This is highly relaxed using the proposed MAB approach, as every WiGig AP is operating autonomously without the need for any global management among the WiGig concurrent links. Simultaneously, comparable performances to the optimal one are obtained.

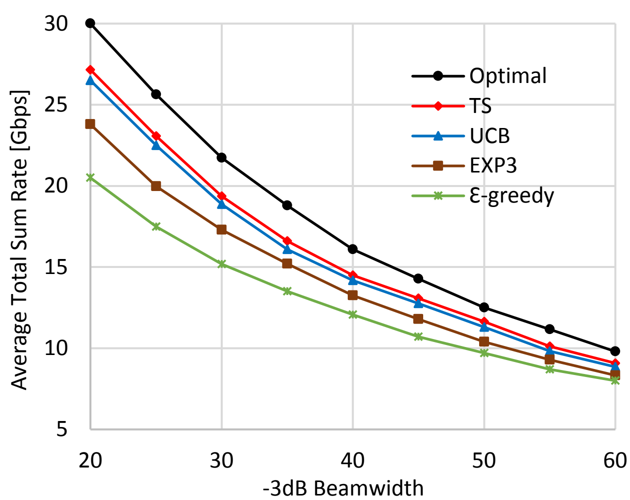


**Figure 6:** Average total system rate against the -3dB beamwidth using 4 concurrent links and LoS blockage probability of 0

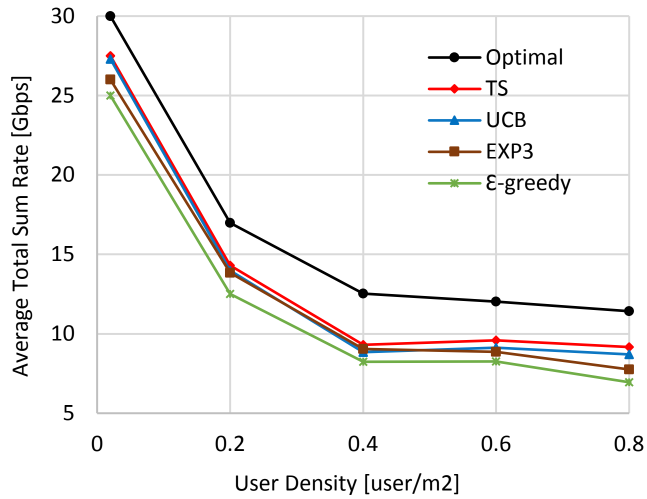


**Figure 7:** Average total system rate against user density using 4 concurrent, LoS blockage probability of 0 and -3 dB beamwidth of 50°

## 7 Conclusion

In this work, we investigated mmWave concurrent BF and formulated it as an optimization problem. An online learning approach based on multiplayer MAB is proposed to address the topic. In this formulation, WiGig APs are acting as the players, the beam space as the arms of the bandits, and the achievable data rates as the rewards. Due to the fully decentralized setting of the problem, WiGig APs play the game selfishly without any knowledge about other APs actions and or observations. During the learning process, a WiGig AP selects its beam ID only based on its successive interaction with the environment and its own reward observations. In this paper, four MAB algorithms are adopted by each WiGig AP to interact with the environment and select their beam IDs, namely, $\epsilon$-greedy, UCB, TS, and EXP3, and their performances are compared. Despite the selfish behavior of the WiGig APs, they learn to take actions that enhance the overall performance of the WiGig WLAN. This is done by selecting beam IDs that avoid mutual interference with the other concurrent links during the learning process. Extensive numerical analysis proved the potency of the proposed MAB based schemes compared to the optimal performance coming from exhaustively searching all available concurrent beams combinations in terms of average total sum rate and spatial reuse factor. The obtained results of this study open the doors for more applications to MAB algorithms to address several mmWave challenges.

**Funding Statement:** The paper is fully funded by RIKEN-AIP.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

**References**

**Ahmed, I.; Khammari, H.; Shahid, A.; Musa, A.; Kim, K. S. et al.** (2018): A survey on hybrid beamforming techniques in 5G: architecture and system model perspectives. *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3060-3097.

**Alkhateeb, A.; Alex, S.; Varkey, P.; Li, Y.; Qu, Q. et al.** (2018): Deep learning coordinated beamforming for highly-mobile millimeter wave systems. *IEEE Access*, vol. 6, pp. 37328-37348.

**Aykin, I.; Akgun, B.; Feng, M.; Krunz, M.** (2020): MAMBA: a multiarmed bandit framework for beam tracking in millimeter-wave systems. *IEEE INFOCOM Conference on Computer Communications*, Toronto, Canada.

**Booth, M. B.; Suresh, V.; Michelusi, N.; Love D. J.** (2019): Multiarmed bandit beam alignment and tracking for mobile millimeter wave communications. *IEEE Communications Letters*, vol. 23, no. 7, pp. 1244-1248.

**Cayamcela, M. M. E.; Lee, H.; Lim, W.** (2019): Machine learning for 5G/B5G mobile and wireless communications: potential, limitations, and future directions. *IEEE Access*, vol. 7, pp. 137184-137206.

**Chen, Y. T.; Xiong, J.; Xu, W. H.; Zuo, J. W.** (2019): A novel online incremental and decremental learning algorithm based on variable support vector machine. *Cluster Computing*, vol. 22, no. 3, pp. 7435-7445.

**Gatzianas, M.; Kalfas, G.; Vagionas, C.; Mesodiakaki, A.** (2019): Downlink coordinated beamforming policies for 5G millimeter wave dense networks. *European Conference on Networks and Communications*, Valencia, Spain, pp. 342-346.

**Ghasempour, Y.; Silva, C. R.; Cordeiro, C.; Knightly, E. W.** (2017): IEEE 802.11 ay: Next-generation 60 GHz communication for 100 Gb/s Wi-Fi. *IEEE Communications Magazine*, vol. 55, pp. 186-192.

**Gui, Y.; Zeng, G.** (2020): Joint learning of visual and spatial features for edit propagation from a single image. *The Visual Computer*, vol. 36, no. 3, pp. 469-482.

**Hashemi, M.; Sabharwal, A.; Koksal, C. E.; Shroff, N. B.** (2018): Efficient beam alignment in millimeter wave systems using contextual bandits. *IEEE Conference on Computer Communications*, Honolulu, HI, pp. 2393-2401.

**Hashima, S.; Hatano, K.; Takimoto E.; Mohamed, E. M.** (2020): Neighbor discovery and selection in millimeter wave D2D networks using stochastic MAB. *IEEE Communications Letters*, Early Access, doi: 10.1109/LCOMM.2020.2991535.

**Huo, Y.; Dong, X.; Xu, W.; Yuen, M.** (2019): Enabling multi-functional 5G and beyond user equipment: a survey and tutorial. *IEEE Access*, vol. 7, pp. 116975-117008.

**Kumar, D.; Saloranta, J.; Kaleva, J.; Destino, G.; Tölli, A.** (2018): Reliable positioning and mmWave communication via multi-point connectivity. *Sensors (Basel)*, vol. 18, no. 11.

**Kutty, S.; Sen, D.** (2016): Beamforming for millimeter wave communications: an inclusive survey. *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 949-973.

**Li, H. X.; Li, W. J.; Zhang, S. G.; Wang, H. D.; Pan, Y. et al.** (2019): Page-sharing-based virtual machine packing with multi-resource constraints to reduce network traffic in migration for clouds. *Future Generation Computer Systems*, vol. 96, pp. 462-471.

**Liu, K. Y.; Peng, J.; Wang, J. G.; Yu, B. Y.; Liao, Z. F. et al.** (2019): A learning-based data placement framework for low latency in data center networks. *IEEE Transactions on Cloud Computing*. doi:10.1109/TCC.2019.2940953.

**Mohamed, E. M.; Elhalawany, B. M.; Khallaf, H. S.; Zareei, M.; Zeb, A. et al.** (2020): Relay probing for millimeter wave multi-hop D2D networks. *IEEE Access*, vol. 8, pp. 30560-30574.

**Mohamed, E. M.; Sakaguchi, K.; Sampei, S.** (2017): Wi-Fi coordinated WiGig concurrent transmissions in random access scenarios. *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10357-10371.

**Qiao, J.; Shen, X.; Mark, J. W.; He, Y.** (2015): MAC-layer concurrent beamforming protocol for indoor millimeter wave networks. *IEEE Transactions on Vehicular Technology*, vol. 64, no. 1, pp. 327-338.

**Rappaport, T. S.; Sun, S.; Mayzus, R.; Zhao, H.; Azar, Y. et al.** (2013): Millimeter wave mobile communications for 5G cellular: it will work! *IEEE Access*, vol. 1, pp. 335-349.

**Rappaport, T. S.; Xing, Y.; MacCartney, G. R.; Molisch, A. F.; Mellios, E. et al.** (2017): Overview of millimeter wave communications for fifth-generation (5G) wireless networks-with a focus on propagation models. *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6213-6230.

**Sim, M. S.; Lim, Y.; Park, S. H.; Dai, L.; Chae, C.** (2020): Deep learning-based mmWave beam selection for 5G NR/6G with sub-6 GHz channel information: algorithms and prototype validation. *IEEE Access*, vol. 8, pp. 51634-51646.

**Song, Y.; Yang, G. B.; Xie, H. T.; Zhang, D. Y.; Sun, X. M.** (2017): Residual domain dictionary learning for compressed sensing video recovery. *Multimedia Tools and Applications*, vol. 76, no. 7, pp. 10083-10096.

**Wang, J.; Cheng, M.; Wang, J.; Lin, M.; Wu, Y. et al.** (2018): Bandit inspired beam searching scheme for mmWave high-speed train communications. arXiv, abs/1810.06150.

**Wang, J.; Jiang, C.; Zhang, H.; Ren, Y.; Chen K. Ch. et al.** (2020): Thirty years of machine learning: the road to pareto-optimal wireless networks. *IEEE Communications Surveys & Tutorials*, Early Access, doi: 10.1109/COMST.2020.2965856.

**Wang, X.; Kong, L.; Kong, F.; Qiu, F.; Xia, M. et al.** (2018): Millimeter wave communication: a comprehensive survey. *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 1616-1653.

**Wilhelmi, F.; Cano, C.; Neu, G.; Bellalta, B.; Jonsson, A. et al.** (2019): Collaborative spatial reuse in wireless networks via selfish multiarmed bandits. *Ad Hoc Networks*, vol. 88, pp. 129-141.

**Wu, W.; Cheng, N.; Zhang, N.; Yang, P.; Zhuang, W. et al.** (2019): Fast mmWave beam alignment via correlated bandit learning. *IEEE Transactions on Wireless Communications*, vol. 18, no. 12, pp. 5894-5908.

**Zeng, Z.; Cremaschi. S,** (2018): A relaxed Knapsack problem-based decomposition heuristic for large scale multistage stochastic programs. *28th European symposium on Computer Aided process engineering*, vol. 43, pp. 519-524.

**Zhang, Z.; Li, Y. B.; Wang, C.; Wang, M. Y.; Tu, Y. et al.** (2018): An ensemble learning method for wireless multimedia device identification. *Security and Communication Networks*. doi:10.1155/2018/5264526.