



# Reinforcement Learning of Millimeter Wave Beamforming Tracking over COSMOS Platform

Imtiaz Nasim

inasi001@fiu.edu

Florida International University

Miami, FL, USA

Panagiotis Skrimponis

ps3857@nyu.edu

New York University

Brooklyn, NY, USA

Ahmed S. Ibrahim

abrahim@fiu.edu

Florida International University

Miami, FL, USA

Sundeep Rangan

srangan@nyu.edu

New York University

Brooklyn, NY, USA

Ivan Seskar

seskar@winlab.rutgers.edu

WINLAB, Rutgers University

North Brunswick, NJ, USA

## ABSTRACT

Communication over large-bandwidth millimeter wave (mmWave) spectrum bands can provide high data rate, through utilizing high-gain beamforming vectors (briefly, beams). Real-time tracking of such beams, which is needed for supporting mobile users, can be accomplished through developing machine learning (ML) models. While computer simulations were used to show the success of such ML models, experimental results are still limited. Consequently in this paper, we verify the effectiveness of mmWave beam tracking over the open-source COSMOS testbed. We particularly utilize a multi-armed bandit (MAB) scheme, which follows reinforcement learning (RL) approach. In our MAB-based beam tracking model, the beam selection is modeled as an action, while the reward of the algorithm is modeled through the link throughput. Experimental results, conducted over the 60-GHz COSMOS-based mobile platform, show that the MAB-based beam tracking learning model can achieve almost 92% throughput compared to the Genie-aided beams after a few learning samples.

## CCS CONCEPTS

- Networks → Network experimentation; Wireless access networks;
- Hardware → Radio frequency and wireless circuits.

## KEYWORDS

Beamforming tracking; COSMOS testbed; millimeter wave; multi-armed bandit; reinforcement learning; wireless experimentation.

### ACM Reference Format:

Imtiaz Nasim, Panagiotis Skrimponis, Ahmed S. Ibrahim, Sundeep Rangan, and Ivan Seskar. 2022. Reinforcement Learning of Millimeter Wave Beamforming Tracking over COSMOS Platform. In *16th ACM Workshop on Wireless Network Testbeds, Experimental evaluation & Characterization (WiNTECH'22), October 17, 2022, Sydney, NSW, Australia*. , 5 pages. <https://doi.org/10.1145/3556564.3558242>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

WiNTECH'22, October 17, 2022, Sydney, NSW, Australia

© 2022

ACM ISBN 978-1-4503-9527-4/22/10...\$15.00

<https://doi.org/10.1145/3556564.3558242>

## 1 INTRODUCTION

Communication over Millimeter wave (mmWave) spectrum bands can provide high data rate, thanks to their abundant bandwidth [1–3]. Such high data rate is crucial for many emerging applications, over static or mobile networks, such as extended reality and self-driving cars. However mmWave communication, with its frequency bands of 30 GHz or more, suffers from the inherent large path-loss at such high frequencies. To compensate for such large path-loss, high-dimensional phased antenna arrays have been employed. The directional alignment of these phased antenna arrays at both transmitter and receiver can produce high power gain to compensate for large path-loss [4–6]. While such alignment of beamforming vectors is possible for static communication nodes (i.e., transmitter, receiver), it is much harder to be maintained for mobile users including vehicles [7]. In particular, beamforming requires accurate knowledge for pointing beams towards the receiver, which comes at the cost of large overhead [8]. This is achievable for static or slowly-changing environments but is difficult for fast-changing environments, such as those in the vehicular context where the UEs are dynamic [9]. Therefore, there is a need for real-time tracking of mmWave beamforming vectors (briefly, beams) in mobile communication, and this is the *scope* of this paper.

Recently, mmWave beam tracking was addressed in the literature via either analytical approaches or machine learning (ML) models. For example, beam tracking approaches using extended Kalman filtering were proposed in [9, 10]. Furthermore, data detection was jointly considered with mmWave beam tracking in [11]. Beyond model-based analytical approaches, multiple ML models have been proposed for mmWave beam tracking. For example, a supervised deep learning (DL) model was employed in [12] for mmWave mobile systems. Furthermore, a recurrent neural network (RNN) was proposed in [13] to track the angle of arrival (AoA) at the user equipment (UE). Moreover, a deep neural network (DNN) was introduced in [14] to predict a user's temporal channel behavior using a long short term memory (LSTM) model.

In addition to supervised DL, multiple reinforcement learning (RL) models were also proposed beam tracking in mmWave communications. For example, mmWave beam tracking for single-user was considered in [15]. Going beyond a single user, we were able to develop in [16] an RL model for simultaneously tracking the mmWave beams for multiple UEs. While the aforementioned learning models introduce novel beam tracking solutions, they were mostly verified

via computer simulations or proprietary testbeds (e.g., in [15]). On the contrary, there is a need to validate such mmWave beam tracking solutions on *open-source* and *large-scale* testbed, and this is the *experimentation gap* addressed in this paper.

The National Science Foundation (NSF), through its Platforms for Advanced Wireless Research (PAWR) program, is currently supporting a few open-source testbeds. Of particular interest in this paper is the Cloud Enhanced Open Software Defined Mobile Wireless Testbed for City-Scale Deployment (COSMOS) platform [17, 18]. The COSMOS platform supports mmWave communication and can emulate multiple mobility scenarios, which can be remotely controlled. Therefore, the *goal* of this paper is to design and validate an RL model for mmWave beam tracking, using the COSMOS platform.

In this paper, we propose an RL-based *multi-armed bandit (MAB)* learning model to track the mmWave beams in the indoor COSMOS platform, which is located in Rutgers University, New Jersey, USA. In the COSMOS platform, two software-defined radio (SDR) kits are used to represent two communication nodes, namely, a base station and a UE. The UE is placed over a remotely-controlled XY-table, which is used to randomly move the UE within the lab area. The proposed MAB model aims to continuously find an adequate beamforming vector for the moving UE, as it traverses the COSMOS lab area. In doing so, the base station works as the MAB agent to select the best beam.

The MAB formulation is well suited algorithm for the exploration-exploitation learning nature of RL models [19], as it can track the changes in a mobile environment and adapt to them accordingly. Moreover, Thompson Sampling (TS) [19] is considered for selecting the best arms (i.e., beams) of the proposed MAB model. Throughout the learning process, one arm is played by the MAB agent at each time slot and an associated reward is observed. We model the reward as the achievable throughput at the UE, which is computed using the information-theoretic Shannon capacity formula taking into account the received signal-to-noise ratio (SNR). Our experimental results show that the proposed MAB model can successfully track the best beams after a few time instants, and achieve 92% of the genie-aided throughput, which always picks the optimal beam.

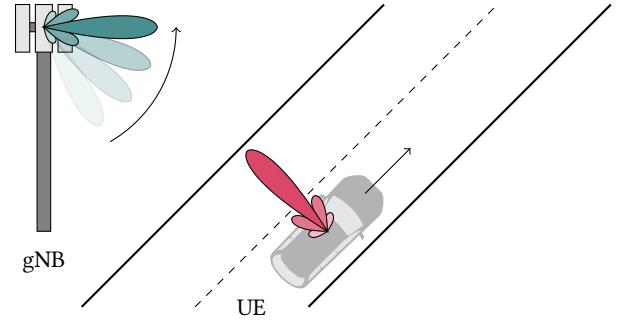
The rest of the paper is organized as follows. Section 2 introduces the system model and describes how we formulate the beam selection problem into an MAB one. Section 3 explains the proposed MAB solution. The system setup and the experimental results are discussed in Section 4. Finally, Section 5 concludes the paper.

## 2 SYSTEM MODEL AND PROBLEM FORMULATION

In this section, we first describe the system model for beamforming selection in mmWave mobile communications over the COSMOS testbed. Then, we show the problem formulation of the beam tracking and its mapping into an MAB one.

### 2.1 System Model

We consider a downlink system with a base station sending data packets to a mobile UE. This consideration can be applied to a vehicular system, as illustrated in Fig. 1. While we fix the position of the base station, the mobile UE moves along a random trajectory within



**Figure 1: An illustration of the considered network with a base station serving a mobile UE.**

the COSMOS testbed over the COSMOS XY-table. We assume that the base station has  $M$  antennas while the UE has a single antenna. The  $M \times 1$  complex channel vector between the base station and the UE at a specific time slot  $t$  is denoted as  $\mathbf{h}_t$ . We choose the beamforming vectors from a fixed codebook  $\mathbf{F} = \{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ , where  $m$  indicates the maximum number of available beams in the codebook. We assume a beamforming vector  $\mathbf{f}_t$  is selected at time slot  $t$  from the codebook  $\mathbf{F}$ .

The received signal for a given time slot  $t$  is modeled as

$$y_t = \mathbf{h}_t^H \mathbf{f}_t x_t + w_t, \quad (1)$$

where  $x_t$  is the transmitted signal at time  $t$  with a transmission power  $P$ . The  $w_t$  term in (1) represents an additive white Gaussian noise (AWGN) with zero-mean and variance  $\sigma^2$ . The received SNR can be computed as,

$$\text{SNR}_t = \frac{P|\mathbf{h}_t^H \mathbf{f}_t|^2}{\sigma^2}. \quad (2)$$

Finally, we apply the Shannon's modified capacity formula to compute the throughput at time slot  $t$  as

$$C_t = B \frac{S_c}{S_{\text{tot}}} \log_2(1 + \text{SNR}_t), \quad (3)$$

where  $B$  is the transmission bandwidth, and  $S_c$  is the number of subcarriers, for orthogonal frequency division multiplexing (OFDM) technology, used out of the total  $S_{\text{tot}}$  subcarriers.

### 2.2 Problem Formulation

The problem can be formulated as maximizing the capacity given in (3), over the  $T$  time slots. Consequently, the optimum beamforming codeword,  $\mathbf{f}_t^*$ , is the one that maximizes the following optimization formula

$$\mathbf{f}_t^* = \arg \max_{\mathbf{f}_t \in \mathbf{F}} \frac{1}{T} \sum_{t=1}^T B \frac{S_c}{S_{\text{tot}}} \log_2 \left( 1 + \frac{P|\mathbf{h}_t^H \mathbf{f}_t|^2}{\sigma^2} \right). \quad (4)$$

Such problem is mapped into an MAB one by first modeling each beamforming vector in codebook  $\mathbf{F}$  as an *arm* of the MAB agent. Therefore, selecting an action  $\mathbf{a}_t$  for any time  $t$  refers to selecting the beamforming vector  $\mathbf{f}_t$  from codebook  $\mathbf{F}$  by the proposed MAB solution. Second, the observed reward, denoted as  $\mathbf{r}_t$ , for any selected action  $\mathbf{a}_t$  (i.e., a codebook vector) is equivalent to the throughput given in (3). These rewards are modeled as random

samples from the selected beams underlying reward distribution. Third, the proposed MAB solution based on TS algorithm, produces a set of selected arms  $a_t$  that are equivalent to the desired beams  $f_t$ , and their associated rewards  $r_t$  are observed till  $t = T$ . The goal of the proposed MAB model is to select the optimal beams that maximize the expected time-average reward i.e.,  $[\frac{1}{T} \sum_{t=1}^T r_t]$ .

### 3 PROPOSED MULTI-ARMED BANDIT LEARNING MODEL

The proposed MAB beam tracking scheme selects a beam based on its prior knowledge from the past information. One arm is played by the agent at each time slot  $t$  and an associated reward is observed. The algorithm uses this information about the reward to select the beam in the next time slot  $t + 1$ .

**Algorithm 1:** Thompson sampling based MAB

---

```

for  $t = 1$  to  $T$  do
  for any beam  $b$  in  $F$  do
    // Choose and apply action
    Select action  $a_t$  based on  $\{\alpha_1, \alpha_2, \dots, \alpha_m\}$ ;
    for action  $a_t$  observe reward  $r_t$  do
      // Update distributions
    for  $b \in F$  do
      if  $a_t = b$  then
         $\alpha_{b,t+1} \leftarrow \gamma_1 \alpha_{b,t} + \gamma_2 r_t$ 
      else if  $a_t \neq b$  and  $\max\{\gamma_1 \alpha_{b,t}\} > 1$  then
         $\alpha_{b,t+1} \leftarrow \gamma_1 \alpha_{b,t}$ 
      else
         $\alpha_{b,t+1} \leftarrow 1$ 
  
```

---

Since we consider a mobile environment, assuming fixed knowledge about the beams associated reward will result in lower capacity. Therefore, we need an update of each beams reward. We employ TS, which is a Bayesian inference algorithm [19], to update the knowledge about the beams associated reward according to [15, 16]. TS is a posterior sampling technique that requires a suitable prior to represent the knowledge of an arm's reward before making an observation. The main idea is to choose an arm based on its probability of being the best arm.

We apply Dirichlet distribution [20] to model the Bayesian prior of the expected reward for each selected arm. The base station selects a beam based on the reward distribution of each arm from a knowledge distribution set  $\{\alpha_1, \alpha_2, \dots, \alpha_m\}$ , and observes an associated reward from the reward distribution set  $\{\theta_1, \theta_2, \dots, \theta_m\}$ . We assume a feedback link is available that passes the information about the observed reward to the base station. The base station continues to apply the Bayesian inference based on the feedback to update its knowledge of each arm's mean reward until time  $t = T$ . In other words, the base station continuously explores for the best beam and exploits based on its past learning from the feedback, thus balancing the exploration/exploitation trade-off. The optimal reward for any time slot  $t$  is achieved when the selected action  $a_t^*$  is equivalent to the optimal codeword  $f_t^*$  in (4).

**Table 1: Parameters**

Parameter	Value
No. of codebook beams	16
Tx Antenna M layout ( $\lambda/2$ )	$16 \times 4$
Operating Frequency	60 GHz
Bandwidth (BW)	983 MHz
Used subcarriers ( $S_c$ )	800
Total subcarriers ( $S_{\text{tot}}$ )	1024
Noise figure	7 dB
No. of time slots	1024
Forget factor ( $\gamma_1$ )	0.3
Boost factor ( $\gamma_2$ )	12

Furthermore, we need to keep in mind that we want the algorithm to always keep exploring to meet the dynamic vehicular demand. To accomplish this, we introduce a "forget" factor  $\gamma_1$ , that ignores the relevance from past occurrences and a "boost" factor  $\gamma_2$ , that increases the impact of the most recent observations to account for the non-stationary behaviour [15]. This makes our proposed MAB algorithm *adaptive* that keeps track of the changes in a mobile environment. The proposed TS based MAB algorithm for beam selection considering a mobile user is given in Algorithm 1.

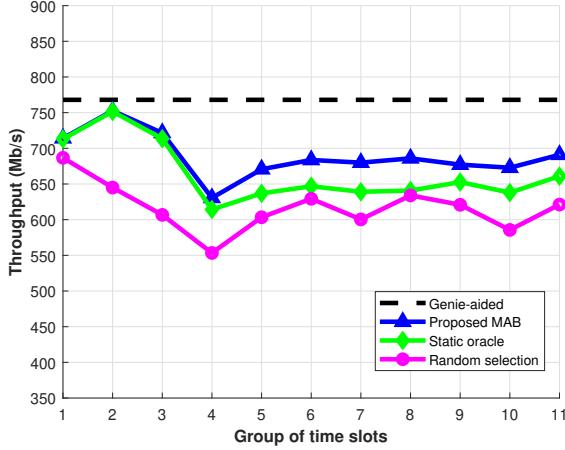
## 4 PERFORMANCE EVALUATION

In this section, we describe the experimental results of our proposed MAB model for beam selection in COSMOS testbed.

### 4.1 Experimental Setup

In [21], the authors describe the mmWave capabilities of the COSMOS testbed and the available open-source tutorials [22]. In this work, we use the resources of the benchtop mmWave setup of COSMOS-sb1. More specifically, we use a Xilinx RFSoC evaluation board connected to a Sivers IMA 57–64 GHz transceiver. Each array is on top of an XY table, enabling movement along the X and Y-axis and rotation around Z-axis. At the transmitter, we select 16 beams evenly spread from -45 to 45 degrees, while on the receiver, we select the beam pointing at 0 degrees. The UE moves at a random location on the X-Y plane at each time step, while the base station remains at a specific location throughout the experiment. The movement of the considered UE is limited by the XY table, which is 1.3m long. The UE requires around 3.6 seconds to cover this length. The distance between the transmitter and the UE is approximately 20m, according to the COSMOS lab setup. To calibrate the arrays, we follow the techniques developed by the authors in [23, 24]. To estimate the received signal strength, we use a frequency-domain channel sounder. The transmitter repeats a sequence with cyclic repeat while the receiver performs correlation for each sequence to collect the power delay profile from every direction. We provide the implementation as open-source in [25].

Several benchmarks are considered for comparison to our proposed solution. The Genie-aided solution is equivalent to an optimal scheme that always selects the best beam resulting in highest capacity. This scheme is considered as the upper bound for our MAB-based RL model. The static oracle is the other considered benchmark that conducts an exhaustive search to find the optimal



**Figure 2: Achievable throughput versus group of time slots (16 × 4 phased array).**

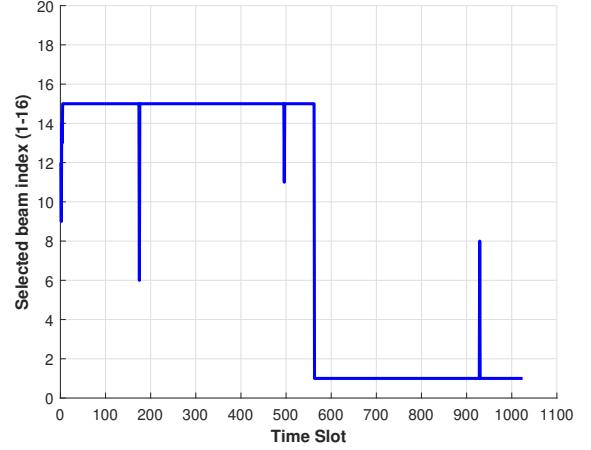
codeword at the beginning of each phase and continues to pick that particular codeword until the next scan. This scheme is equivalent to what is proposed for beam tracking in 5G NR, i.e., using the same beam until the next scan [15]. Lastly, the random selection is considered as the lower bound that has no prior knowledge of the beam quality and arbitrarily selects one at each time instance. The parameters considered for our experiment are given in Table 1.

## 4.2 Experimental Results

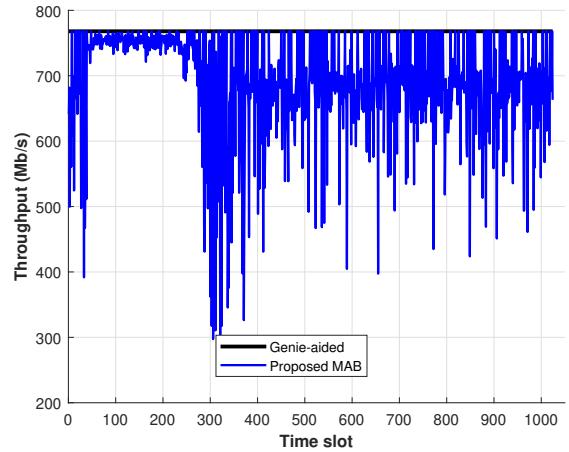
Fig. 2 shows the performance of the proposed MAB beam selection model with comparison to the other benchmarks. We consider 1024 time slots, each of which are approximately 3.5 ms long. We partition the available time slots into small groups of 100 instances, and average over each of them to evaluate the throughput performance. As shown in Fig. 2, the proposed MAB model can achieve a higher throughput than the static oracle and random selection benchmarks, thanks to its accurate learning capability of the best beams. This is because, unlike our MAB beam selection model, the static oracle or the random selection benchmark cannot track the optimal beams at every time slot and hence, achieves lower throughput. In addition, our proposed MAB beam selection scheme achieves a throughput of 700 Mb/s, which is 92% compared to the capacity of 760 Mb/s achieved by the Genie aided solution.

Fig. 2 also shows that it takes a few iterations (i.e., 4 groups of time slots) before the accomplished throughput starts to be monotonically increasing. This is an indication that the MAB-based learning model takes a few iterations to learn the dynamic environment and allocates the mmWave beamforming vectors accordingly. The random selection lower bound performs worst as it has no prior knowledge of the good beams and arbitrarily picks one at each time instance.

Fig. 3 shows the selected beam index by our proposed MAB solution approach for the given time slots. As shown in Fig. 3, the 15-th beam index is selected more frequently until the 550-th time slot due to its high associated reward. However, as the UE moves



**Figure 3: Selected beam index for proposed Thompson sampling based MAB scheme.**



**Figure 4: Achievable throughput at each time slot (16 × 4 phased array).**

and the reward associated with 15-th beam index deteriorates, the algorithm responds to the environmental change by selecting beam index 1 for most of the remaining time slots. Also, there are some instances when the proposed scheme selects some other beams, such as 6-th beam index at 184-th time slot, 9-th beam index at 500-th time slot and 8-th beam index at 920-th time slot. This is a key feature of the considered TS method, demonstrating that while beams with currently high predicted rewards are more likely to be chosen, other beams may also be explored at some instances, even if they have a lower associated reward i.e., exploration versus exploitation.

The impact of selecting such beams with lower associated reward is demonstrated in Fig. 4, where we show the throughput achieved at every time slot by the proposed MAB model. The lower spikes produced by our MAB approach at some instances indicate that

a less significant throughput was achieved due to the selection of beams which are not optimal for those locations. Nonetheless, over 90% of occurrences indicate a high throughput, implying that the proposed beam tracking model can select the good beams more often due to its improved learning accuracy.

## 5 CONCLUSIONS AND FUTURE WORK

In this paper, we have proposed a multi-armed bandit learning model for tracking of mmWave beamforming vectors in mobile environment. The proposed model was experimentally evaluated using the mmWave COSMOS lab. The proposed MAB beam tracking model applies Thompson sampling to select the best beam that maximizes the capacity for a mobile user. Experimental results validated the efficacy of the proposed model with comparison to other benchmarks. We have shown that the proposed learning model provides 92% of the data rate achieved by a Genie-aided benchmark within a few learning iterations. The solution was applied for a single-user scenario considering two SDRs, where one acted as the base station and the other as the UE. Considering multiple dynamic UEs for beam tracking in mmWave COSMOS environment is a future work of this paper.

## ACKNOWLEDGMENTS

The work of Imtiaz Nasim and Ahmed S. Ibrahim is supported in part by the National Science Foundation under the supplemental Platforms for Advanced Wireless Research (PAWR) award number CNS-1816112.

Panagiotis Skrimponis and Sundeep Rangan are with NYU Wireless, Tandon School of Engineering, New York University, Brooklyn, NY. They are supported under NSF grants 1952180, 1925079, 1564142, 1547332, the Semiconductor Research Corporation (SRC) and the industrial affiliates of NYU Wireless.

## REFERENCES

- [1] Theodore S. Rappaport, Shu Sun, Rimma Mayzus, Hang Zhao, Yaniv Azar, Kevin Wang, George N. Wong, Jocelyn K. Schulz, Mathew Samimi, and Felix Gutierrez. Millimeter wave mobile communications for 5G cellular: It will work! *IEEE Access*, 1:335–349, 2013.
- [2] Sundeep Rangan, Theodore S Rappaport, and Elza Erkip. Millimeter-wave cellular wireless networks: Potentials and challenges. *Proc. of the IEEE*, 102(3):366–385, 2014.
- [3] Vutha Va, Takayuki Shimizu, Gaurav Bansal, and Robert W. Heath Jr. Millimeter wave vehicular communications: A survey. *Foundations and Trends in Networking*, 10(1):1–113, 2016. ISSN 1554-057X.
- [4] C. Nicolas Barati, S. Amir Hosseini, Sundeep Rangan, Pei Liu, Thanasis Korakis, Shivendra S. Panwar, and Theodore S. Rappaport. Directional cell discovery in millimeter wave cellular networks. *IEEE Transactions on Wireless Communications*, 14(12):6664–6678, 2015.
- [5] C. Nicolas Barati, Sourya Dutta, Sundeep Rangan, and Ashutosh Sabharwal. Energy and latency of beamforming architectures for initial access in mmWave wireless networks. *Journal of the Indian Institute of Science*, 100:281–302, Apr 2020. ISSN 0019-4964.
- [6] Morteza Hashemi, Ashutosh Sabharwal, C. Emre Koksal, and Ness B. Shroff. Efficient beam alignment in millimeter wave systems using contextual bandits. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, pages 2393–2401, 2018.
- [7] Jingxian Wu and Pingzhi Fan. A survey on high mobility wireless communications: Challenges, opportunities and solutions. *IEEE Access*, 4:450–476, 2016.
- [8] J. Wang, Z. Lan, C. Pyo, T. Baykas, C.-S. Sum, M. Rahman, J. Gao, R. Funada, F. Kojima, H. Harada, and S. Kato. Beam codebook based beamforming protocol for multi-gbps millimeter-wave wpan systems. *IEEE Journal on Selected Areas in Communications*, 27(8):1390–1399, 2009.
- [9] Vutha Va, Haris Vikalo, and Robert W. Heath. Beam tracking for mobile millimeter wave communication systems. In *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 743–747, 2016.
- [10] Suhanya Jayaprakasam, Xiaoxue Ma, Jun Won Choi, and Sunwoo Kim. Robust beam-tracking for mmWave mobile communications. *IEEE Communications Letters*, 21(12):2654–2657, 2017.
- [11] Shahram Shahsavari, Mohammad A. Amir Khojastepour, and Elza Erkip. Robust beam tracking and data communication in millimeter wave mobile networks. In *2019 International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT)*, pages 1–8, 2019.
- [12] Ahmed Alkhateeb, Sam Alex, Paul Varkey, Ying Li, Qi Qu, and Djordje Tujkovic. Deep learning coordinated beamforming for highly-mobile millimeter wave systems. *IEEE Access*, 6:37328–37348, 2018.
- [13] Daoud Burghal, Naveed A. Abbasi, and Andreas F. Molisch. A machine learning solution for beam tracking in mmWave systems. In *2019 53rd Asilomar Conference on Signals, Systems, and Computers*, pages 173–177, 2019.
- [14] Sun Hong Lim, Sunwoo Kim, Byonghyo Shim, and Jun Won Choi. Deep learning-based beam tracking for millimeter-wave communications under mobility. *IEEE Transactions on Communications*, 69(11):7458–7469, 2021.
- [15] Irmak Aykin, Berk Akgun, Mingjie Feng, and Marwan Krunz. MAMBA: A multi-armed bandit framework for beam tracking in millimeter-wave systems. In *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, pages 1469–1478, 2020.
- [16] Imtiaz Nasim, Ahmed S. Ibrahim, and Seungmo Kim. Learning-based beamforming for multi-user vehicular communications: A combinatorial multi-armed bandit approach. *IEEE Access*, 8:219891–219902, 2020.
- [17] Dipankar Raychaudhuri, Ivan Seskar, Gil Zussman, Thanasis Korakis, Dan Kilper, Tingjun Chen, Jakub Kolodziejski, Michael Sherman, Zoran Kostic, Xiaoxiong Gu, Harish Krishnaswamy, Sumit Maheshwari, Panagiotis Skrimponis, and Craig Gutterman. Challenge: COSMOS: A city-scale programmable testbed for experimentation with advanced wireless. In *Proc. ACM MobiCom'20*, 2020.
- [18] Cloud enhanced open software defined mobile wireless testbed for city-scale deployment (COSMOS). <https://cosmos-lab.org/>, 2022.
- [19] Shriya Agrawal and Navin Goyal. Analysis of thompson sampling for the multi-armed bandit problem. In *Proc. of the 25th Annual Conference on Learning Theory*, volume 23, pages 39.1–39.26. PMLR, 2012.
- [20] S. Liu. Dirichlet distribution: Motivating lda. <https://towardsdatascience.com/dirichlet-distribution-a82ab942a879>, 2019.
- [21] Tingjun Chen, Prasanthi Maddala, Panagiotis Skrimponis, Jakub Kolodziejski, Xiaoxiong Gu, Arun Paidimari, Sundeep Rangan, Gil Zussman, and Ivan Seskar. Programmable and open-access millimeter-wave radios in the PAWR COSMOS testbed. In *Proc. of the 15th ACM WiNTECH'21*, page 1–8, 2022. ISBN 9781450387033.
- [22] COSMOS tutorials. <https://wiki.cosmos-lab.org/wiki/tutorials/>, 2022.
- [23] Aditya Dhananjay, Kai Zheng, Jaakko Haarla, Lorenzo Iotti, Marco Mezzavilla, Dennis Shasha, and Sundeep Rangan. Calibrating a 4-channel fully-digital 60 GHz SDR. In *Proc. ACM WiNTECH'20*, page 40–47, 2020. ISBN 9781450380829.
- [24] Aditya Dhananjay, Kai Zheng, Marco Mezzavilla, Lorenzo Iotti, Dennis Shasha, and Sundeep Rangan. Pi-Radio v1: Calibration techniques to enable fully-digital beamforming at 60 GHz. *Computer Networks*, 196:108220, 2021. ISSN 1389-1286.
- [25] NYU WIRELESS. Experimentation with advanced SDR. <https://github.com/nyu-wireless/mmwsdr>, 2022.