Machine Learning for Robust Beam Tracking in Mobile Millimeter-Wave Systems

Sopan Sarkar*, Marwan Krunz*, Irmak Aykin* and David Manzi†
*Dept. of Electrical and Computer Engineering, University of Arizona, United States
†Raytheon Corporation, United States
Email: sopansarkar@email.arizona.edu, krunz@ece.arizona.edu, aykin@email.arizona.edu, dgmanzi@rtx.com

***Abstract*—Narrow beams in millimeter-wave (mmWave) communication introduce significant beam misalignment challenges. In this paper, we introduce MAMBA-X, an enhanced version of the MAMBA beam tracking scheme. Basically, MAMBA uses a restless multi-armed bandit framework to capture the dynamics of mmWave links by discounting the relevance of past observations using a *"forgetting factor"* (γ_1) and increases the weight of recent observations via a *"boost factor"* (γ_2). Because the original MAMBA uses fixed values for γ_1 and γ_2, it cannot quickly adapt to variations in user mobility. Moreover, if the time between consecutive beam selection instances is large compared to channel dynamics, past observations become obsolete. To tackle these issues, we first use the concept of beam coherence time to establish a bound on the beam selection intervals. Secondly, we show that the performance of MAMBA depends primarily on the value of γ_1 which, in turn, depends on UE mobility. We develop a Long Short-Term Memory (LSTM) model to dynamically predict and update the optimal value of γ_1. Through extensive simulations at 28 GHz and using publicly available 5G NR experimental dataset, we evaluate MAMBA-X. Our results indicate that the total delivered traffic is improved by up to 46.8% relative to the original MAMBA and 142% compared to the default beam management scheme in 5G NR.**

***Index Terms*—Millimeter-wave, beam tracking, LSTM, reinforcement learning, beam coherence time.**

I. Introduction

With the increasing demand for high data rates, wireless systems are increasingly shifting to new spectrum frontiers in the millimeter-wave (mmWave) band. MmWave spectrum is a key aspect of next-generation wireless systems such as 5G NR and WiGig. One major drawback of transmissions at mmWave frequencies is that the signal suffers from very high attenuation compared to sub-6 GHz bands. None-the-less, by utilizing high-dimensional phased antenna arrays with high directivity (and gain) it is possible to achieve higher data rates even in a harsh mmWave channel environment [1].

Although analog beamforming provides high gains, it is quite difficult to establish and maintain a directional link. During initial access (IA) or cell discovery, the BS establishes a communication link with a new UE and updates the links of already connected ones. However, the IA process incurs significant overhead [2]. Moreover, due to UE mobility and environmental changes, significant beam misalignment can occur resulting in reduced data rate and communication outage [3]. Therefore, it is critical to track mobile UEs efficiently so as to maintain the quality of service of the communication link.

Several techniques were proposed in the literature to address this issue [4]–[6]. MAMBA [7] is a restless multi-armed bandit (MAB) framework for beam tracking in mmWave systems. In MAMBA, the BS acts as the agent and interacts with each beam to learn the changes in beam quality over time. MAMBA utilizes a technique called adaptive Thompson sampling (ATS) to determine the best beam for transmission/reception based on past observations of the channel. Moreover, to address the nonstationarity of the channel, MAMBA uses a *"forgetting factor"* (γ_1) and a *"boost factor"* (γ_2) to control the impact of its past observations on the decision of best beam selection.

The performance of MAMBA depends on how well it can adapt to link/channel variations. In mobile scenarios, link dynamics are strongly dependent on UE speed and distance from the base station. Thus, using fixed values for γ_1 and γ_2 cannot properly capture the channel variations. Moreover, if the time between two beam selection instances is large compared to the rate at which the channel changes, past observations become obsolete [8]. Therefore, making beam selection decisions based on such outdated observations will lead to errors and necessities establishing time limit within which the beam selection must be performed.

In this paper, we propose an extension of MAMBA, called MAMBA-X, that addresses the aforementioned problems. Specifically, our main contributions are as follows:

- We determine an upper bound on the optimal time for beam selection in MAMBA. This time is derived based on the beam coherence time, which reflects the impact of beam misalignment due to UE mobility, and dictates how frequently the channel should be observed and the decision to select a new beam. The derived bound ensures the optimal performance for MAMBA, even in the extreme scenarios where the mmWave channel is changing rapidly due to mobility.
- We show through simulations that the performance of MAMBA is particularly sensitive to the value of the *forgetting factor* γ_1. Furthermore, we show how the throughput-optimal value of γ_1 depends on UE mobility (i.e., distance and velocity), relative to the base station.

This research was supported in part by NSF (grants CNS-1563655, CNS-1731164, and IIP-1822071) and by the Broadband Wireless Access & Applications Center (BWAC). Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of NSF.

- We develop a Long Short-Term Memory (LSTM) based recurrent neural network (RNN) for predicting the optimal value of γ_1. The LSTM network is trained offline using a dataset collected through extensive simulations and predictions are performed online based on estimated UE mobility with respect to the BS.
- We conduct simulations and also utilize available 5G experimental dataset to verify the efficiency of MAMBA-X in terms of total delivered traffic, average data rate, instantaneous data rate, and outage duration in outdoor scenarios. The results indicate that with the proposed modifications, the total delivered traffic improves by up to 46.8% relative to original MAMBA and 142% compared to the default beam management scheme proposed for 5G.

II. System Model

A. MmWave Mobility Scenario at 28 GHz

Without loss of generality, we consider an outdoor scenario in which a BS is tracking a mobile UE. [1] Two UE mobility models are considered: random waypoint model (RWM) and random circular-motion model (RCM). In RWM, the UE selects a random destination and a random speed. It then moves to this destination at that speed. It pauses at the first destination for a fixed period of time before randomly selecting another destination and speed. RCM is a constrained version of RWM, where the UE moves along an arc that is centered at the BS, i.e., the BS-UE distance is fixed. We use RCM to study the impact of $\gamma 1$ and $\gamma 2$ on the of performance MAMBA and their dependence on distance and velocity. On the other hand, RWM is used to simulate real UE movement in the mmWave environment.

Beamforming is applied at both the BS and the UE. Relatively wide beams are used at the UE given its small form-factor and few antenna elements. Therefore, we consider beam tracking only from the BS side.

B. Codebook-based Beamforming

Let the BS and UE be equipped with uniform planner arrays (UPAs). The total number of antenna elements at the BS and the UE are denoted by A_{BS} and A_{UE}, respectively. Tx and Rx beamforming vectors (a.k.a beamers) depend on $A_{\text{UE}} \times A_{\text{BS}}$ the complex channel matrix $\mathbf{H}$ between the BS and UE. Denote the codebooks for the BS beamformer as $\mathcal{M} = \{\mathbf{m}_1, \mathbf{m}_2, \cdots, \mathbf{m}_{M_{\text{BS}}}\}$ and for the UE beamformer as $\mathcal{N} = \{\mathbf{n}_1, \mathbf{n}_2, \cdots, \mathbf{n}_{N_{\text{UE}}}\}$, where M_{BS} and N_{UE} are the total numbers of narrow beams that can be generated at the BS and the UE, respectively. We assume that after IA, both the BS and the UE decide on a directional link. Suppose that the BS uses a Tx beamforming vector $\mathbf{m}_k \in \mathbb{C}^{A_{\text{BS}} \times 1}$, and the UE uses an Rx beamforming vector $\mathbf{n}_l \in \mathbb{C}^{A_{\text{UE}} \times 1}$ (k and l are the indices of the Tx/Rx beamforming vectors in their respective codebooks). Let $s(t)$ be the transmitted signal at any time t. Then the received signal $y_{kl}(t)$ at t can then be expressed as:

$$y_{kl}(t) = \mathbf{n}_l^H \mathbf{H} \mathbf{m}_k s(t) + \mathbf{n}_l^H \mathbf{z}(t) \tag{1}$$

where $\mathbf{z} \in \mathbb{C}^{A_{\text{UE}} \times 1}$ is a vector representing complex circularly-symmetric white Gaussian noise. Each $(\mathbf{m}_k, \mathbf{n}_l)$ pair achieves a certain Rx power $P_{kl}(t)$ at time t, where $P_{kl}(t) = |y_{kl}|^2$. The distribution of $P_{kl}(t)$ is non-stationary, as $\mathbf{H}$ varies with time.

C. MAMBA Beam Tracking Scheme

1) Problem Formulation: MAMBA is modeled as a single-state Markov decision process, where the BS acts as an agent interacting with directional beams to learn the changes in beam qualities over time. Beam quality impacts the best modulation and coding scheme (MCS) that the beam can support at a given instance. MAMBA utilizes an RL algorithm, called adaptive Thompson sampling (ATS), to select the best beam/MCS pair. This selection is updated periodically at fixed time intervals. The choice of this time interval does not take into account the mobility scenario and is defined arbitrarily (1 ms in [7]).

The MAMBA framework is specified by the tuple $\langle \mathcal{A}, \mathcal{R} \rangle$, where $\mathcal{A} \triangleq \{\mathbf{m}_1, \cdots, \mathbf{m}_{M_{\text{BS}}}\}$ is the set of actions, i.e., possible beams at a given time, and $\mathcal{R}$ is the set of rewards, i.e., achievable rates or MCS indices associated with these actions, respectively. At time t, an action $a_t \in \mathcal{A}$ refers to the selection of beam, say i, which results in a reward $\mathbf{r}_t \in \mathcal{R} \sim \mathbf{\Theta}_{i,t}$ to be observed. Here, $\mathbf{r}_t$ is a random sample drawn from the selected beam's reward distribution $\mathbf{\Theta}_{i,t}$ at time t with the mean of $\mathbf{\Theta}_{i,t}$, i.e., $\mathbb{E}[\mathbf{\Theta}_{i,t}] = \boldsymbol{\theta}_{i,t}$ being unknown. After IA, MAMBA selects a policy $\boldsymbol{\mu}$ characterized by the actions to be taken at subsequent times $t = 1, 2, .., T$ so as to maximize the cumulative reward.

Through IA, the BS has prior belief about the reward distribution associated with each beam. Using Bayesian inference, BS continuously updates the posterior of mean reward ($\boldsymbol{\theta}$), given the observed data ($\mathbf{u}$), as follows:

$$\Pr(\boldsymbol{\theta}|\mathbf{u}) = \Pr(\mathbf{u}|\boldsymbol{\theta}) \Pr(\boldsymbol{\theta}) / \Pr(\mathbf{u}). \tag{2}$$

In MAMBA, the reward distribution is represented by a K-dimensional categorical random variable, $\mathbf{r}_t = [r_t^{(0)}, \cdots, r_t^{(K-1)}]$, where K is the number of MCS indices supported by the system plus one. Here, $\mathbf{r}_t$ is a vector of all zeros except one entry of value 1, representing the highest attainable MCS index k, $k \in \{1, \cdots, K-1\}$. The observed data rate at time t is given by $\mathbf{r}_t \mathbf{w}^T$, where $\mathbf{w} \triangleq [w_0, w_1, \cdots w_{K-1}]$ is the value vector whose entries correspond to the transmission rates associated with the corresponding MCS indices, with $\mathbf{w} \triangleq 0$ if the transmission is unsuccessful. For any given beam, the BS can communicate with the UE using one of the $K-1$ MCS indices. If the BS fails to establish communication with the UE, $w_0 = 0$ is selected, i.e., ${r_t}^{(0)} = 1$. The BS performs its computation based on the feedback (ACK/NACK) received from the UE. Specifically, the BS calculates the received signal strength (RSS) of the ACK/NACK packet and determines the highest possible MCS index for the reverse

[1] Multiple UEs can be served by the same BS beam. When a UE moves out of the beam, the BS switches the data flow to the new beam.

direction. The optimal policy, $\boldsymbol{\mu}^*$ aims at maximizing the expected throughput:

$$\underset{\boldsymbol{\mu}}{\text{maximize}} \quad \sum_{t=1}^{T} \boldsymbol{\theta}_{i,t}\mathbf{w}^T$$
$$\text{s.t.} \quad \sum_{k=0}^{K-1} \theta_{i,t}^{(k)} = 1, \theta_{i,t}^{(k)} \geq 0, \ \forall i,t,k. \tag{3}$$

Because the rewards vector is unknown and non-stationary in a dynamic mmWave scenario, (3) cannot be solved directly. Instead an ATS-based RL algorithm is used to learn the expected rewards and select the optimal policy.

2) Adaptive Thompson Sampling (ATS): As discussed before, the average reward for each possible beam is modeled as a categorical random variable. So, MAMBA models the prior of the expected rewards using a Dirichlet distribution with parameter $\boldsymbol{\alpha}_{i,t}$, Dir($\boldsymbol{\alpha}_{i,t}$). This is because the Dirichlet distribution is the conjugate prior of the categorical distribution. According to (2), the posterior computed at each round will also follow a Dirichlet distribution. When the prior is the conjugate distribution of the likelihood, the update rule is simpler. At time t, action a_t is taken and reward $\mathbf{r}_t$ is observed. To capture the dynamic nature of the directional mmWave channels under mobility, MAMBA implements a *"forgetting factor"*, γ_1 that reduces the effect of past observations and a *"boost factor"*, γ_2 that emphasizes on the most recent observations. For beam $i \in \mathcal{A}$, the updated posterior under these factors is as follows:

$$\boldsymbol{\alpha}_{i,t+1} = \begin{cases} \gamma_1\boldsymbol{\alpha}_{i,t} + \gamma_2\mathbf{r}_t, & \text{if } a_t = i \\ \gamma_1\boldsymbol{\alpha}_{i,t}, & \text{if } a_t \neq i \text{ and } \gamma_1\left\|\boldsymbol{\alpha}_{i,t}\right\|_1 > 1 \\ \mathbf{1}, & \text{otherwise.} \end{cases}$$

where, $\gamma_1 \in (0,1]$, $\gamma_2 \geq 1$, and $\|.\|_1$ is the 1-norm of a vector (maximum absolute value of its elements).

After the distributions are updated, the BS selects the beam for the next round based on random samples taken from the current posterior distributions. Specifically, at each time t, the BS samples from each beam's updated distribution to obtain $\mathbf{s}_{i,t} \sim \text{Dir}(\boldsymbol{\alpha}_{i,t})$, $\forall i \in \mathcal{A}$ (where $|\mathbf{s}_{i,t}| = 1$, $\forall i \in \mathcal{A}$ and $\forall t \in \{1, \cdots, T\}$), and selects the action according to Thompson sampling as follows:

$$a_t = \arg\max_{i \in \mathcal{A}} \mathbf{s}_{i,t}\mathbf{w}^T. \tag{4}$$

III. Optimizing Beam Selection Times

MAMBA is effective as long as channel characteristics change relatively slowly. This is because, if the mobility and beam misalignment are too fast relative to the time between two beam selection rounds, the algorithm is slow to learn the dynamic environment, as the available observations become obsolete. To address this issue we define a bound between beam selection instances (T_s). Within this bound the channel must be measured and decision to rerun MAMBA is made.

Consider a scenario in which the UE moves within the coverage area of the BS, as shown in Fig. 1. At point A, the BS runs IA to discover and connect with the UE. The distance between the BS and UE at A is d. Right after IA, we assume that the BS and UE beams are perfectly aligned. Suppose that the UE starts moving at a fixed speed v towards point B and reaches B after a short duration τ. Due to the motion of the UE from A to B, beam misalignment may occur and the signal-to-noise ratio (SNR) of the received signal drops. The angular change in the AoA of the line-of-sight signal at the UE between times t and $t+\tau$ is defined as the beam pointing error and is represented by $\Delta\phi$ in Fig. 1. From the sine law of triangle, we get

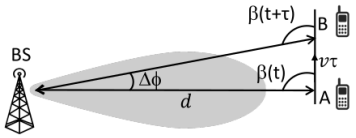


Fig. 1. Misalignment caused by UE mobility.

$$\frac{d}{\sin(\pi - (\beta(t) + \Delta\phi))} = \frac{v\tau}{\sin(\Delta\phi)}.$$

Here, $\beta(t)$ is the angle between the antenna boresight of the BS and the direction of travel of the UE at time t. For a small value[1] of $\Delta\phi$, $\sin(\Delta\phi) \approx \Delta\phi$ and $\sin(\pi - (\beta(t) + \Delta\phi)) = \sin(\beta(t))$. Then the beam pointing error $\Delta\phi$ can be expressed as:

$$\Delta\phi = \frac{v\tau}{d}\sin(\beta(t)). \tag{5}$$

To determine the beam selection time, we rely on the concept of beam coherence time, T_B [9]. Specifically, we define T_B as the time when the received SNR of the beam drops below a certain threshold, ξ, from the peak:

$$T_B = \inf_{t'>0}\{t' | \frac{\text{SNR}(t+t')}{\text{SNR}(t)} < \xi\}. \tag{6}$$

Suppose that the transmit power and antenna gain of the UE are fixed, and the noise power does not vary drastically over the duration T_B. Since the received SNR at any time t is proportional to the antenna gain, G of the BS, we can express ξ as $\xi = \frac{G(\Delta\phi)}{G(0^\circ)}$. Here, $G(\phi)$ is the antenna gain at an angle ϕ from the boresight. For UPAs, we can express $G(\phi)$ as[10]:

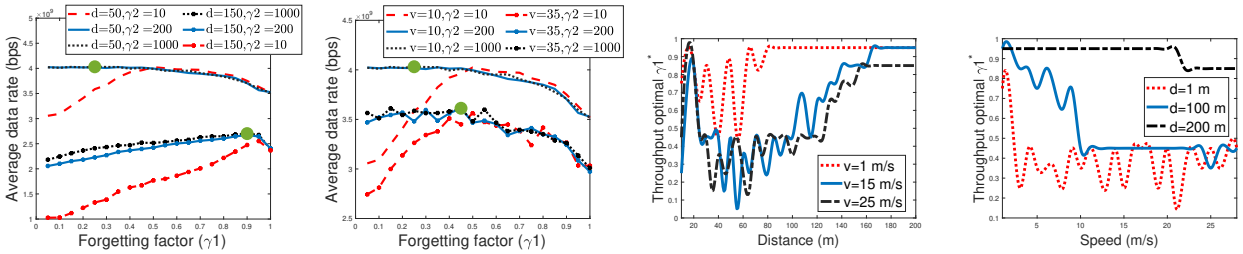
$$G(\phi) = G_{\text{max}} 10^{-\frac{3}{10}(\frac{2\phi}{\phi_\omega})^2}, \text{ for } |\phi| \leqslant \frac{\phi_m}{2} \tag{7}$$

where $G_{\text{max}} = G(0^\circ)$ is the maximum gain, ϕ_ω is the HPBW, and ϕ_m is the main lobe beamwidth. For the case $\Delta\phi > \phi_m$, we consider the situation as beam-alignment failure rather than beam misalignment. From (5), (6) and (7), we can calculate the value of T_B as,

$$T_B = \frac{d}{v\sin(\beta(t))}\frac{\phi_w}{2}\sqrt{\frac{-10}{3}\log\xi}. \tag{8}$$

Note that T_B is defined for a fixed UE speed and a given BS-UE distance. In practice, the motion of the UE can vary

[1]UPAs used in 5G mmWave BS's often have beam sweeping granularity below 5° and half-power beamwidth (HPBW) as low as $7^\circ - 10^\circ$



(a) (b) (c) (d)

Fig. 2. Impact of MAMBA parameters on performance: (a) average data rate vs. γ_1 for various distances; (b) average data rate vs. γ_1 for various speeds; (c) γ_1^* vs. distance; (d) γ_1^* vs. speed. The circles in (a) and (b) represents γ_1^*.

over time, and so T_B will also change. If T_B is used as the basis for determining T_s and is initially set to a large value due to low UE speed, but shortly after that the UE increased its speed, the beam misalignment will likely occur before the next beam selection time. To overcome this situation, we can set T_s equal to the minimum value of T_B, i.e.,$T_{B_{\min}}$. According to (8), this minimum value occurs when the UE is moving at $\beta(t) = 90^\circ$ at the maximum speed $v_{\max}$ supported by the BS and at a minimum distance $d_{\min}$ from it. We can always choose beam selection instances to be less than $T_{B_{\min}}$, which will provide more context regarding the channel dynamics. But in such a case, MAMBA needs to run more frequently resulting in increased communication overhead between the BS and UE which will reduce the overall system efficiency.

IV. Predicting the Optimal γ_1

To observe the impact of γ_1 and γ_2 on MAMBA performance, and generate dataset for training the LSTM network, we rely on computer simulations. In the simulation setup, a BS is placed at the center and the UE moves according to RCM. To model the communication link between the BS and UE, we set the number of antennas at the BS to A_{BS} = 16 and at the UE to A_{UE} = 2, and Tx power to P_{Tx} = 30 dBm. To generate large-scale effects, we run our simulations in the 28 GHz band using the model described in [11]. For small-scale effects, we randomly place 3 scatters on the ellipsoid between the BS and UE. The Rx beam at UE is kept the same, and BS does azimuth beam tracking with a scanning resolution of 5°. Beam scanning range is from $\pm 30^\circ$ degree from the broadside. The BS-UE separation is varied between 10 m to 200 m and the UE speed is varied between 1 m/s to 35 m/s, respectively. We run the simulation for different values of γ_1 and γ_2, and $\xi = 0.5$ (3dB change) and average the results over 500 runs.

A. Dependence of MAMBA on γ_1 and γ_2

Fig. 2(a) depicts the average data rate when the UE moves at a fixed speed of 10 m/s and the BS-UE distance are 50 m and 150 m respectively. Similarly, Fig. 2(b) shows the impact of various UE speeds (10 m/s and 35 m/s) on the data rate when the BS-UE distance is fixed to 50 m. From the two figures, it is evident that the average data rate varies significantly with γ_1 and less with γ_2. After a certain value of γ_2, the effect of γ_2 is more or less the same. Moreover, the throughput-optimal value of γ_1 (γ_1^*) depends on BS-UE distance and UE speed. For a fixed speed, if the UE is closer to the BS, a small displacement will cause a large angular deviation. As a result, the BS is more likely to switch between beams. On the other hand, if the UE is far away from the BS, it can be served by the same beam for a longer period of time. The same intuition is applicable to when the UE is at a fixed distance from the BS but is moving with variable speeds. Based on this fact, we perform simulations to observe the effect of UE mobility on γ_1^* for a fixed beamwidth which is evident from Fig. 2(c) and Fig. 2(d).

B. LSTM-based Prediction of the Optimal γ_1

From Section IV-A, we observed that γ_1^* depends on UE mobility. We use LSTM-based RNN to model this dependency. Let $t_1, t_2, \ldots,$ be the time instances at which MAMBA-X is executed. At any of these instances, say t_n, let the input vector be denoted by $X^{(t_n)} = X(d_{t_n}, v_{t_n}, t_n)$, corresponding to the vector containing the distance, d_{t_n} and speed, v_{t_n} of the UE at t_n. Let $\gamma_1^*(t_n)$ denote the value of γ_1^* at any arbitrary time t_n. The prediction of $\gamma_1^*(t_n)$ can be formulated as:

$$\gamma_1^*(t_n) = \mathcal{F}(X^{(t_n)}) \tag{9}$$

where $\mathcal{F}(.)$ defines the mapping from the input $X^{(t_n)}$ to the output $\gamma_1^*(t_n)$, which needs to be learned.

LSTM network is a special kind of RNN that is capable of learning long-term dependencies through a number of hidden variables [12]. A typical LSTM cell architecture is shown in Fig 3. Basically, an LSTM cell has three gate structures: forget gate ($f^{(t_n)}$), input gate ($i^{(t_n)}$) and output gate ($o^{(t_n)}$). These gates are composed of a sigmoid (σ) layer along with a point-wise multiplication operation structure. Each gate outputs a number between 0 and 1, where 0 indicates that the gate is totally blocked and 1 indicates that all information of the input is kept in the cell. The operations performed by the LSTM cell on the inputs $(X^{(t_n)}, h^{(t_n-1)})$ are as follows:

$$f^{(t_n)} = \sigma\left(W^{(fX)}X^{(t_n)} + V^{(fh)}h^{(t_n-1)} + b^{(f)}\right)$$
$$i^{(t_n)} = \sigma\left(W^{(iX)}X^{(t_n)} + V^{ih}h^{(t_n-1)} + b^{(i)}\right)$$
$$C^{(\tilde{t}_n)} = \tanh\left(W^{(cX)}X^{(t_n)} + V^{(ch)}h^{(t_n-1)} + b^{(o)}\right)$$
$$C^{(t_n)} = f^{(t_n)}C^{(t_n-1)} + i^{(t_n)}\tilde{C}^{(t_n)}$$
$$o^{(t_n)} = \sigma\left(W^{(oX)}X^{(t_n)} + V^{(oh)}h^{(t_n-1)} + b^{(o)}\right)$$

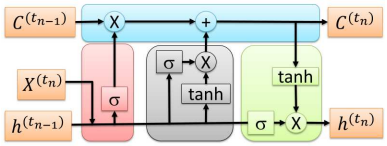


Fig. 3. LSTM cell architecture.

$$h^{(t_n)} = o^{(t_n)} \tanh(C^{(t_n)}).$$

Here, $W^{(fX)}$, $W^{(iX)}$, and $W^{(oX)}$ are the weights of the forget, input, and output gates, respectively. Also, $b^{(f)}$, $b^{(i)}$, and $b^{(o)}$ are the biases that corresponds to these gates. Finally, $V^{(fh)}$, $V^{(ih)}$, and $V^{(oh)}$ are their corresponding recurrent weights.

In this paper, we construct the LSTM network with 5 LSTM layers each with 32 hidden units and a dense output layer with tanh activation function. Moreover, use an adam optimizer and select root mean squared error (RMSE) as the loss function. We train the LSTM network offline by the data collected using RCM, while the predictions are performed online using RWM. The training is completed when RMSE reaches the minimum value. The predicted value of γ_1^* using our LSTM network during testing phase is shown in Fig. 4(a). The consistency between the actual and predicted value of γ_1^* at each time slot t_n is clearly observed from the figure. The developed LSTM network achieves an RMSE value of 0.082 and 0.116 during training and testing phase respectively. We incorporate all the above modifications in our MAMBA-X framework.

V. Performance Evaluation

The goal of beam tracking is to extend the period between two IA cycles as much as possible by reducing the number of outages, and increasing the overall throughput. To evaluate the performance of the MAMBA-X framework, we perform extensive simulations and run our algorithm on publicly available 5G NR experimental dataset.

A. Simulation Results

For simulation, the parameters used for both MAMBA and MAMBA-X are similar to the one described in Section IV. However, in this case, we use RWM rather than RCM to characterize UE mobility. For MAMBA, we let the fixed value of γ_1 and γ_2 to be 0.2 and 20 respectively as defined in [7], and we set (i) T_s to vary based on T_B as calculated using (8) (let us denote it by MAMBA(var)), (ii) $T_s = T_{B_{\min}}$ (15 ms according to simulation parameters and denote it as MAMBA(15ms)), and (iii) $T_s < T_{B_{\min}}$ (let $T_s = 1$ ms and denote it as MAMBA(1ms)), respectively. On the other hand, for MAMBA-X we let γ_1 vary based on the UE mobility and set γ_2 to a large value such as 500, and $T_s = T_{B_{\min}} = 15$ ms. The slot duration equals to the corresponding T_s.

Fig. 4(b), depicts the total delivered traffic achieved by the BS-UE link when running MAMBA-X and compares it with the dynamic oracle, the static oracle, and various settings of MAMBA respectively. Here, the dynamic oracle represents the situation when the best beam is always selected. The static oracle represents the beam management scheme proposed in the 5G standard [13]. As seen from the figure, MAMBA-X outperforms the static oracle, MAMBA(15ms), and MAMBA(var) scheme and performs reasonably close to the dynamic oracle. Though MAMBA(1ms) has better throughput than MAMBA-X, it runs more frequently increasing overall overhead. The total delivered traffic achieved by MAMBA-X is 18.8% higher than that of the MAMBA(15ms) and 172.5% higher than the static oracle.

We compute the CDF of the link outage duration for the beam tracking frameworks and demonstrate the results in Fig. 4(c). Outage occurs when either the wrong beam is selected or the selected MCS index is not supported. The figure shows for how long the frameworks stay in outage after it loses communication. We see that with probability exceeding 0.8 the outage will be approximately 2 slots for MAMBA-X, where as, that for MAMBA(1ms), MAMBA(15ms) and MAMBA(var) is more than 7 slots. This indicates that MAMBA-X can quickly realign the beam in case of link outage which demonstrates its robustness. Moreover, there will be fewer IA cycles in MAMBA-X compared to MAMBA, thus extending the time between consecutive IA cycle.

Fig. 4(d) depicts the total outage probability for the beam tracking frameworks. From the figure, we observe that MAMBA-X has the least amount of outages compared to the other schemes. The outage probability decreases by almost 21% and 6% compared to MAMBA(15ms) and MAMBA(1ms), respectively.

B. Experimental Results

To verify the effectiveness of MAMBA-X in real world scenario, we rely on Lumos5G dataset [14] - a publicly available dataset from the University of Minnesota (UMN), USA. The dataset consists of traces that were collected around the U.S. Bank Stadium in Minneapolis downtown area which is a 1300 meter loop. It consists of information such as UE locations, speeds, 5G BS tower IDs, 5G synchronization signal (SS) measurements such as SS-RSRP, SS-RSRQ, throughput etc. that were sampled and logged every second. The UE speeds reported in the dataset varied from 0 m/s to 14 m/s.

From the dataset, we select the BS with "tower_id" 16 because of the large number of traces and calculate the BS-UE distance from the provided UE and BS locations. Moreover, we calculate the SS-RSSI from the SS-RSRP and SS-RSRQ for a bandwidth of 400MHz, and number of resource blocks equal to 275. Since no information regarding the characteristics of the antenna array or beamwidth was reported, we divide the provided UE locations with a resolution of 6° such that at any given time the UE is within the coverage of at max two beams. This represents an extreme case where the beam is either very narrow or the UE is moving very close to the highest speed supported by the system. Taking all these into consideration we run various beam tracking algorithm using the dataset. Here, T_s is set equal to the sampling time, i.e., 1 second.

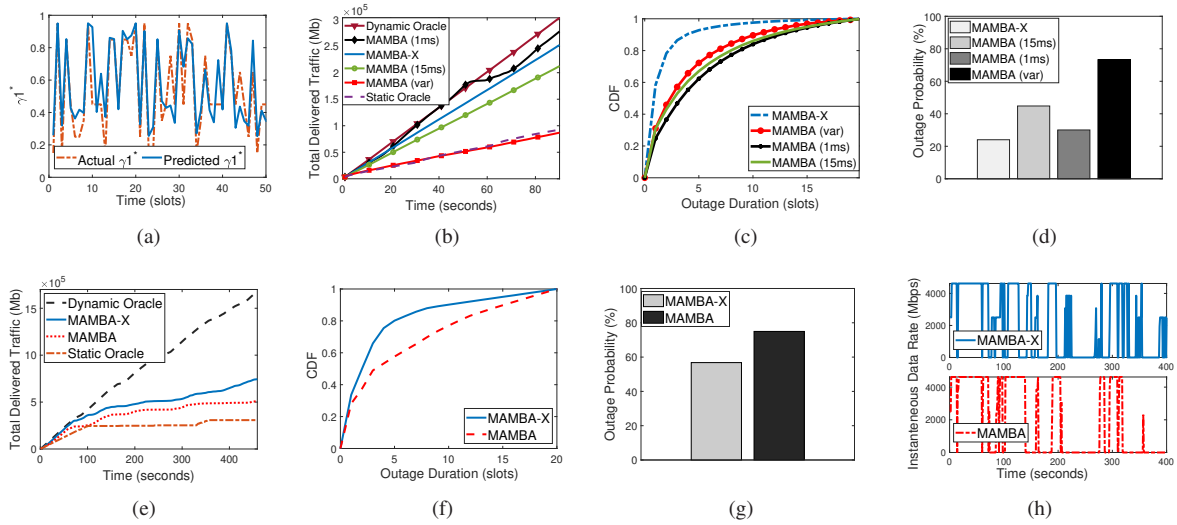


Fig. 4. (a) Predicted values of γ_1^* using LSTM at different time slots; Performance of various beam tracking algorithms. Simulation results: (b) total delivered traffic over time (c) CDF of link outage duration; (d) link outage probability; Experimental results: (e) total delivered traffic over time (f) CDF of link outage duration; (g) link outage probability; (h) instantaneous data rate.

Fig. 4(e) depict the total delivered traffic for different beam tracking algorithms. We observe that MAMBA-X performs better and the total delivered traffic increases by up to 46.8% and 142% compared to MAMBA and static oracle respectively. Moreover, from Fig. 4(f) we observe that the link outage duration for MAMBA-X is less than 5 time slots whereas that of MAMBA is less than 11 time slots for 80% of the time. Next, we compare the outage probability of both MAMBA and MAMBA-X in Fig. 4(g). As it can be seen from the figure, MAMBA-X has an outage probability of 56.7% which is 18.2% less than that of MAMBA (74.9%). Finally, Fig. 4(h) depicts the instantaneous data rate achieved by the MAMBA-X and MAMBA schemes. It is evident from the figure that MAMBA-X can quickly adopt to channel variations in mobility scenario and maintains the BS-UE link for a longer period of time.

VI. Conclusion

In this paper, we proposed MAMBA-X, an extension to MAMBA, a MAB-based beam tracking framework. Because MAMBA depends on past observations to predict the next beam/MCS pair to use for communication, we derived an upper bound for beam selection time. Moreover, we showed that the performance of MAMBA depends primarily on γ_1 and developed an LSTM-based RNN model to incorporate in MAMBA-X framework which dynamically predicted the value of γ_1. We ran MAMBA-X on both experimental and simulated dataset and observed that the total delivered traffic increased by up to 46.8% compared to the original MAMBA and by 142% compared to the beam tracking scheme proposed in 5G. Our future work will focus on extensive experimental evaluation of MAMBA-X scheme in different mmWave scenarios.

References

[1] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 173–196, 2018.

[2] T. Bai, A. Alkhateeb, and R. W. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Communications Magazine*, vol. 52, no. 9, pp. 70–77, 2014.

[3] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Transactions on Communications*, vol. 61, no. 10, pp. 4391–4403, 2013.

[4] V. Va, T. Shimizu, G. Bansal, and R. W. Heath, "Online learning for position-aided millimeter wave beam training," *IEEE Access*, vol. 7, pp. 30 507–30 526, 2019.

[5] J. Zhang, Y. Huang, Y. Zhou, and X. You, "Beam alignment and tracking for millimeter wave communications via bandit learning," *IEEE Transactions on Communications*, vol. 68, no. 9, pp. 5519–5533, 2020.

[6] R. Gupta, K. Lakshmanan, and A. K. Sah, "Beam alignment for mmwave using non-stationary bandits," *IEEE Communications Letters*, vol. 24, no. 11, pp. 2619–2622, 2020.

[7] I. Aykin, B. Akgun, M. Feng, and M. Krunz, "MAMBA: a multi-armed bandit framework for beam tracking in millimeter-wave systems," in *Proc. of the IEEE INFOCOM 2020 Conference*, Toronto, Canada, Jul. 2020.

[8] C. Vernade, A. Gyorgy, and T. Mann, "Non-stationary delayed bandits with intermediate observations," in *International Conference on Machine Learning*. PMLR, 2020, pp. 9722–9732.

[9] V. Va, J. Choi, and R. W. Heath, "The impact of beamwidth on temporal channel variation in vehicular channels and its implications," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 5014–5029, 2016.

[10] G. Yang, J. Du, and M. Xiao, "Analysis on 60 GHz wireless communications with beamwidth-dependent misalignment," *arXiv preprint arXiv:1611.07867*, 2016.

[11] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE JSAC*, vol. 32, no. 6, pp. 1164–1179, 2014.

[12] W. Zhang, M. Feng, M. Krunz, and H. Volos, "Latency prediction for delay-sensitive v2x applications in mobile cloud/edge computing systems," in *GLOBECOM 2020-2020 IEEE Global Communications Conference*. IEEE, 2020, pp. 1–6.

[13] Y.-N. R. Li, B. Gao, X. Zhang, and K. Huang, "Beam management in millimeter-wave communications for 5G and beyond," *IEEE Access*, vol. 8, pp. 13 282–13 293, 2020.

[14] 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.