A Beam-Switching Scheme for Resilient mm-Wave Communications With Dynamic Link Blockages

Aniq Ur Rahman[1] and Gourab Ghatak[2]

[1]National Institute of Technology Durgapur, India. Email: aur.20150002@btech.nitdgp.ac.in
[2]Indraprastha Institute of Information Technology Delhi, India. Email: gourab.ghatak@iiitd.ac.in

***Abstract*—In this paper we propose a beam-switching scheme to increase the reliability of the millimeter wave (mm-wave) communication links which are intermittently blocked. We assume that the blockage statistics between the UE equipments (UEs) and the base stations (BSs) change over time. As a result, the data-rates at the UE from different BSs depend not only upon their proximity but also are governed by the blockage dynamics. We model the scenario as a two-armed bandit problem, where playing an arm is analogous to the UE switching the boresight direction of its mm-wave directional beam towards one of the possible BS. In this two-armed bandit problem, we employ a Thompson sampling (TS) algorithm with parameter reset, to enable the UE to select the serving BS. The rewards considered in the algorithm are the spatio-temporal average data-rate coverage probability experienced by the UE, obtained using tools from stochastic geometry. We show that the proposed algorithm outperforms the classical received signal-strength indicator (RSSI) based association scheme in terms of the average rate-coverage probability. The proposed scheme is able to track the changing vehicular blockage dynamics and opportunistically switch between the BS links to maximize the rate-coverage.**

***Index Terms*—Thompson sampling, stochastic geometry, mm-wave, beam-switching, reinforcement learning, multi-armed bandits.**

I. Introduction

Future wireless applications anticipate an explosion in the plethora of use-cases and services, which cannot be sustained by incremental improvements on the existing communication schemes [1]. Specifically, to address the high-data rate applications, exploiting mm-wave spectrum is gaining popularity. In addition to the large available bandwidths, mm-wave communications employ directional antennas which reduces co-channel interference, thereby improving the performance at the UEs [2].

On the downside, mm-wave transmissions suffer from detrimental path-losses and high sensitivity to blockages [3]. For example, a vehicle located between a BS and a pedestrian UE may block the signal and induce a temporary service outage [4]. Similarly, in the indoor mm-wave and sub-mm-wave systems, human blockages can prove to be detrimental to the system performance [5], [6]. To overcome this, the first-generation mm-wave networks will necessarily need the sub-6GHz architecture for their functionality [7]. Specifically, the sub-6GHz link can act as a backup communication resource when the less reliable mm-wave link suffers from outage. However, as the network needs, volume of UEs, and use-cases grow, there will be a call for reliable communications using standalone mm-wave access points [8].

A. Related Work

The field of link-reliability in mm-wave communications is a new, albeit an active area of investigation. Recently, Peng *et al.* [9] have studied statistical characteristics of human blockages experimentally at mm-wave and sub-mm-wave frequencies, using a realistic mobility model for indoor movements of humans. The authors suggested installation of distributed antennas to maintain the signal coverage. An *et al.* [10] have studied a beam switching technique where the mm-wave beam-direction is altered towards an non line-of-sight (NLOS) link to decrease the impact of the blockages in the line of sight (LOS) link. Genc *et al.* [11] proposed a technique in the indoor mm-wave scenario, that consists of using reflected links to avert the blockages in the LOS link. However, none of these studies propose any algorithm to facilitate the beam-switching towards NLOS or reflected links.

Macrodiversity techniques have also been investigated wherein a UE connects to multiple BSs thereby decreasing the overall service outage by facilitating coordinated transmissions from the associated BSs [12], [13]. Recently, Gupta *et al.* [14] have investigated the performance of microdiversity by considering random blockages. Lee *et al.* [15] have studied the performance of coordinated beamforming with dynamic clusters of access points in terms of the spectral efficiency. They have provided a statistical evaluation of the system using stochastic geometry. Coordinated multipoint (CoMP) [16] was introduced in the Release-11 of LTE to augment the reliability of transmissions. One technique to facilitate CoMP i.e., dynamic point selection was studied by Agarwal *et al.* [17]. In dynamic point selection, one access point is dynamically selected with the objective of maximizing the system capacity. Several simulation-based studies exist that characterize the performance of CoMP, e.g., see Morozov *et al.* [18].

This paper aims to investigate the performance of an algorithm that enables the UE to switch from one mm-wave link to the other in case the former is suffering from outage due to blockages. This beam-switching scenario is modeled as an exploration-exploitation trade-off problem where the UE opportunistically senses and/or transmits/receives along the available mm-wave links. The multi-armed bandits (MAB) problem is an important model for studying such exploration-exploitation trade-offs in reinforcement learning. Specifically, the scenario considered is an example of non-stationary MAB problem. Aurelien Garivier and Eric Moulines [19] considered a scenario where the distribution of the rewards remain constant over epochs and change at unknown time instants (i.e., abrupt changes). They analyzed the theoretical upper bounds of regret for the discounted upper confidence bound (UCB) and sliding window UCB. Another class of algorithms aimed to tackle the restless bandit problems are based on the classical TS approach. Gupta *et al.* [20], extending the idea to Bayesian methods, proposed Dynamic TS (Dynamic TS). O Besbes *et al.* [21], by assuming subtle structures in the variations of reward generating process, authors were successful in establishing bounds on minimal achievable regret against a dynamic oracle and developed a near optimal policy, REXP3. Hartland *et al.* [22] considered dynamic bandits with abrupt changes in the reward generation process, and proposed an algorithm called Adapt-EvE. In this paper we assume random (following a Poisson arrival process but unknown to the UE) points-of-change of the reward characteristics of the arms and study a TS based approach to facilitate beam-switching between the different BSs, in order to maximize the data-rate.

B. Contributions and Organization

The contributions made in this paper is summarized as follows:

- We characterize a mm-wave multi-connectivity framework installed in a UE, to facilitate switching of the mm-wave beam between two BSs. In particular, we model the beam-switching scheme as a MAB problem where the rewards are characterized as the rate coverage probability experienced by the UE. In order to take into account the random locations of the UE with respect to the BSs, we characterize the rewards as the spatio-temporal average of the rate coverage probability observed by a typical UE. This not only models the performance of a UE in different network configurations (e.g., different cities), but also in different locations within the same network.
- In this MAB framework for beam-switching, we study a TS based approach to address the exploration-exploitation dilemma. In case the blockage processes of the links are stationary, we show that the TS approach manages to choose the best arm (i.e., associate to the BS which provides the better downlink rate).
- Finally, we consider temporal dynamics in the blockages. In particular, the blockage probabilities of the links are assumed to change after an unknown and random time-period. For this case, we study a version of the TS scheme with parameter reset, where the effect of the previous experiences of rewards are eliminated. We show that the proposed algorithm actually tracks the dynamics blockage characteristics and as a result improves the resilience of the communication link.

The rest of the paper is organized as follows. In Section II we introduce the system model considered in this paper and outline the optimization objectives. In Section III we characterize the rewards of the MAB framework in terms of the rate coverage probability experienced by the UE. Then, in Section IV we present the TS based approach for facilitating beam-switching between the BSs. Numerical results to discuss the efficacy of our proposed scheme is given in Section V. Finally, the paper concludes in Section VI.

II. System Model and Optimization Objectives

A. Network Geometry

We consider a wireless network consisting of mm-wave BSs deployed to provide high-speed downlink connectivity and ad-hoc coverage to the UEs. The positions of the BSs are modeled as atoms of a Poisson point process (PPP) of intensity λ [23]. We evaluate the performance of the network from the perspective of a typical UE located at the origin.

B. Communication Policy

We assume that during the association phase, the typical UE measures the downlink power in the control channel and associates with two BSs that provide the highest and the second highest downlink powers. This can take place, e.g., in the sub-6GHz band (see e.g., [7], [24]). Let us denote the two BSs by BS_1 and BS_2, respectively. As in our model all the BSs transmit with the same power P, and the sub-GHz transmissions are relatively transparent to the blockages, BS_1 and BS_2 are consequently the two nearest BSs from the typical UE. Accordingly, the distance distributions of BS_1 and BS_2 are given by [25]:

$$d_1 \sim f_{d_1}(x) = 2\pi\lambda x \exp\left(-\pi\lambda x^2\right); \quad x \in [0, \infty) \tag{1}$$

$$d_2 \sim f_{d_2}(y) = 2(\pi\lambda)^2 y^3 \exp\left(-\pi\lambda y^2\right); \quad y \in [d_1, \infty) \tag{2}$$

After the association, the UE and the BSs switch to the mm-wave band and the data transmission commences, where, the UE receives downlink data service from either BS_1 or BS_2. Let us introduce the blockage model for the mm-wave links.

C. Blockage and Propagation Model

We assume that the mm-wave communication links suffer from independent temporary blockages due to vehicles or humans. The blockage in the i^{th} channel, b_i (where $i \in \{1, 2\}$) is modelled as a Bernoulli process: $b_i \sim \mathcal{B}(p_i)$. Here $\mathcal{B}$ is the Bernoulli distribution and p_i is the probability of blockage of the i^{th} channel. In general, the temporal statistics of the blockages are relatively constant with respect to the duration of data-services of shorter duration, e.g., in case of mobile UE applications. Whereas for services such as fixed wireless access or mm-wave backhaul, the temporal variation of the blockages play a key role in the data-rate performance. In this paper, we analyze both the cases: static blockages, as well as dynamic blockages, where the parameters p_i change over time. For the latter case, we assume that the blockage characteristics change after an interval t_P, which follows a Poisson arrival process with parameter λ_T, unknown to the UE.

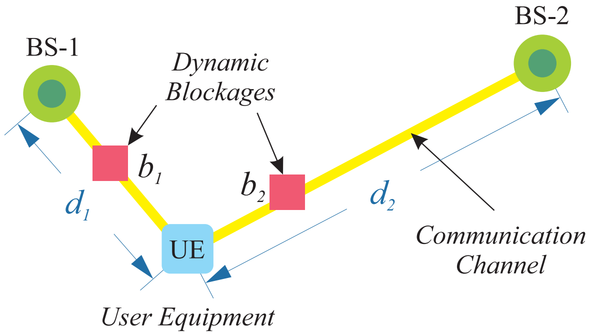


Fig. 1. System Illustration

We assume that the BSs are equipped with directional antennas of gain G. Let the path loss coefficient be given by K. In case the communication link between the transmitter and the receiver is in LOS state (i.e, devoid of blockages), the path loss exponent is given by α_L. On the other hand, in case the communication link is in NLOS state, the path loss exponent is given by α_N. Due to the low local scattering in mm-wave communications, we consider a Nakagami fading h, with parameter n_0 and variance equal to 1 [26]. Thus, in case the transmitter and the receiver are located at a distance r from each other, the instantaneous received power is given by $PGKhr^{-\alpha_L}$ if the link is not blocked, and $PGKhr^{-\alpha_N}$ otherwise. In order to avoid the singularity at the origin, we assume the path-loss model to be bounded [27]. As a result the received power averaged over fast-fading is given by:

$$P_R(r, j) = PGK \min\{1, r^{-\alpha_j}\}, \quad j \in \{L, N\} \tag{3}$$

We model the beam-switching framework as a two-armed bandit problem where the rewards are the downlink rate coverage probabilities from the two associated BSs. The characterization of the rewards parameterized on the distance of the BSs is given in the next section. After describing our TS based beam-switching algorithm in Section IV, we will evaluate the expected value of the rewards accumulated using the distributions of distance of the BSs. The system model is illustrated in Fig. 1.

III. Characterization of Rewards

In case the UE receives services from BS_1, the useful power is given by $PGKh_1d_1^{-\alpha_N}$ or $PGKh_1d_1^{-\alpha_L}$ depending on whether the mm-wave link to BS_1 is in NLOS or LOS, respectively. Similarly, the useful power in case of service from BS_2 is given by $PGKh_2d_2^{-\alpha_N}$ or $PGKh_2d_2^{-\alpha_L}$, respectively. As a result, the corresponding signal to noise ratio (SNR) experienced at the UE is given by:

$$S_i = \frac{PGKh_id_i^{-\alpha_j}}{N_0} = \frac{P_R(d_i, j)h_i}{N_0}, \tag{4}$$

where $i \in \{1, 2\}$ and $j \in \{L, N\}$. The SNR coverage probability at a threshold τ is defined as the probability that the UE receives an SNR of higher than τ. It is evaluated as:

$$\begin{aligned} P_{C i,j}(\tau; d_i, j) &= \mathbb{P}(S_i \geq \tau) = \mathbb{P}\left(\frac{P_R(d_i, j)h_i}{N_0} \geq \tau\right) \\ &= \mathbb{P}\left(h_i \geq \frac{N_0\tau}{P_R(d_i, j)}\right) \\ &\overset{(a)}{=} \sum_{n=1}^{n_0} (-1)^{n+1}\binom{n_0}{n}\exp\left(-\frac{n N_0 n_0 (n_0!)^{-\frac{1}{n_0}}\tau}{P_R(d_i, j)}\right). \end{aligned} \tag{5}$$

Here step (a) follows as in [28]. Based on the SNR coverage probability, the rate coverage probability is at a threshold of r_0 bps is defined as the probability that the UE experiences a downlink data-rate of r_0 bps. Mathematically, it can be calculated as [28]:

$$R_{j\,i}(r_0, d_i) = P_{C\,i}\left(2^{\frac{r_0}{B}} - 1; d_i, j\right) \tag{6}$$

The formulation of the SNR coverage probability is thus parameterized on d_is. After introducing the TS algorithm in the next section, we will average out the rate coverage probability by using the distance distributions of d_is from (1) and (2).

IV. TS Based Beam-Switching

The aim of the work is to design beam-switching scheme that takes into consideration jointly the effect of the distance of the two BSs and also the vehicular blockage characteristics in both of these links. We convert this problem of exploration-exploitation dilemma as a two-armed bandit problem [29] and define the reward R_i associated with the i^{th} channel at time t as:

$$R_i(d_i, t) = \begin{cases} R_{L\,i}(r_0, d_i), & \text{with probability } 1 - p_i(t) \\ R_{N\,i}(r_0, d_i), & \text{with probability } p_i(t) \end{cases} \tag{7}$$

where $R_{L\,i}$ and $R_{N\,i}$ are the LOS and NLOS rewards of the i^{th} channel respectively given by (6). Recall that $p_i(t)$ is the probability of blockage of the i^{th} channel at time t. Thus, in case the boresight direction of the UE is towards the BS_1 at time t, it receives a reward of $R_{L\,1}(r_0, d_1)$ or $R_{N\,1}(r_0, d_1)$ with probabilities $1 - p_1(t)$ or $p_1(t)$, respectively.

Next, we describe the TS algorithm applied to this problem. The goal is to balance between exploiting (i.e., beamforming towards) the BS link that is *currently* known to maximize performance and investing to accumulate new information (e.g., exploring the reward/blockage conditions by connecting to the other BS link) that may improve future performance. In the general scenario, i.e., with dynamic link blockages, the reward distribution evolves over time for the two links. To keep up with this change, we reset our algorithm after a fixed interval of time t_W, which we refer to as reset duration.

The algorithm is described in Algorithm 1. The expected reward θ_i of arm i is drawn from a beta distribution having shape parameters M_i, N_i (step 4). The shape parameters are initialized to zero, to serve as a uniform distribution which signifies a zero-knowledge situation (step 1). The algorithm chooses to play the arm j with a higher value of expected reward (step 5). As we play the arms (step 6) and update the shape parameter values according to the rewards (i.e., rate-coverage), our choices become more and more tuned, towards the optimal choice (steps 7-8).

The link blockage characteristics might change over time based on various factors such as blockage density, signal strength etc. In this scenario, the algorithm will give poor results, being biased by the previous knowledge about a particular link. Therefore, we refresh the shape parameters to zero after an interval of time t_W (steps 10-12).

Algorithm 1 TS with parameter reset.

```
1:  N_i ← 0; M_i ← 0; ∀i ∈ {1,2}            ▷ initializing parameters
2:  t ← 0
3:  while t < T do
4:      θ_i ∼ β(M_i + 1, N_i − M_i + 1)      ▷ draw from beta dist.
5:      j ← i | θ_j = max(θ_i)               ▷ choosing the better arm
6:      R(t) ← R_j(t)                        ▷ playing the chosen arm
7:      N_j ← N_j + R(t)                     ▷ updating the beta distribution
8:      M_j ← M_j + R(t)                     ▷ updating the beta distribution
9:      t ← t + 1
10:     if (t mod t_W) = 0 then              ▷ refreshing parameters
11:         N_i ← 0; M_i ← 0; ∀i ∈ {1,2}
12:     end if
13: end while
```

A. Performance Metric: Average Rate Coverage Probability

We define the average rate coverage probability (ARCP) at a time instant t as the mean of rewards accumulated up to time t for all possible locations of BS_1 and BS_2. We denote the various ARCPs as BEST (Best choice), TS (TS), CH-1 (Channel 1), CH-2 (Channel 2) and RAND (Random choice). Mathematically they are defined as spatio-temporal averages as follows:

$$\text{TS}(\tau) = \mathbb{E}_{d_1,d_2}\left[\frac{1}{\tau}\sum_{t=0}^{\tau} R_{TS}(d_1,d_2,t)\right] \tag{8a}$$

$$\text{BEST}(\tau) = \mathbb{E}_{d_1,d_2}\left[\frac{1}{\tau}\sum_{t=0}^{\tau} \max\{R_1(d_1,t), R_2(d_2,t)\}\right] \tag{8b}$$

$$\text{CH-1}(\tau) = \mathbb{E}_{d_1}\left[\frac{1}{\tau}\sum_{t=0}^{\tau} R_1(d_1,t)\right] \tag{8c}$$

$$\text{CH-2}(\tau) = \mathbb{E}_{d_2}\left[\frac{1}{\tau}\sum_{t=0}^{\tau} R_2(d_2,t)\right] \tag{8d}$$

$$\text{RAND}(\tau) = \mathbb{E}_{d_1,d_2}\left[\frac{1}{\tau}\sum_{t=0}^{\tau} U\{R_1(d_1,t), R_2(d_2,t)\}\right] \tag{8e}$$

where $R(t)$ is reward at time t using the TS approach. The BEST strategy chooses the better arm in hindsight. Whereas, the RAND strategy chooses one arm at random, i.e., $U\{a,b\}$ implies that a and b are equi-probable choices. The spatial averages are taken with respect to the distance distributions as given by (1) and (2).

V. Numerical Results and Discussion

In this section, we discuss the salient features of the proposed beam-switching algorithm based on numerical simulations. The list and the values of the parameters are shown in Table I. The path loss exponents are taken to be 2 for LOS and 4 for NLOS regimes of propagation, respectively. The path-loss coefficient K and the noise-density are obtained from the 3GPP specifications [30].

Table I: Simulation Parameters

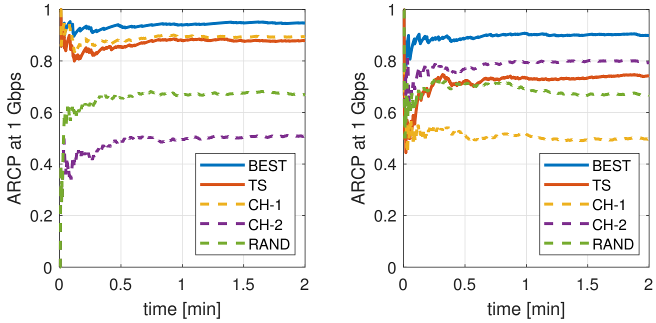
Symbol	Description	Value
G	Antenna Gain	10 dB
P	Transmitter Power	0 dB
α_L	Path Loss exponent (LOS)	2
α_N	Path Loss exponent (NLOS)	4
K	Path Loss Coefficient	$\left(\frac{\text{carrier wavelength}}{4\pi}\right)^2$
N_0	Noise density	-174 dBm/Hz $+\ 10\log_{10}(B)$ + 10 dB
B	Bandwidth	1 GHz
r_0	Data rate	1 Gbps

A. Static Blockages

First, we study the performance of our algorithm when the blockage characteristics of the channels are temporally static. In Fig. 2 we plot the performance of the TS based beam-switching and the other baseline schemes described in (8). Particularly, we investigate two cases: (a) BS_1 has a lower blockage probability than BS_2, and (b) BS_1 has a higher blockage probability than BS_2. In both cases, we observe that the TS based beam-switching scheme selects the better arm in terms of ACRP at the end of the decision interval. Indeed, the TS algorithm is able to jointly take into account the distance as well as the blockage characteristics of the channels for BS selection. More interestingly, in the case (b) we see that due to higher probability of blockage in the mm-wave channel to BS_1, the overall ACRP of BS_1 is significantly lower (20% less as compared to BS_2). In case of an RSSI based association scheme, with probability 0.6 (i.e., 1 - p_1), the UE would have selected the BS_1 for mm-wave service which could lead to a dramatic degradation in the performance. Moreover, the proposed TS based BS selection gives a more accurate idea of the radio environment as compared to schemes where the BS association takes place initially in the sub-6GHz band and then the UE connects to mm-wave band in the selected BS (e.g., see [7]). This is precisely because in the later scheme the UE would likely associate to BS_1 due to its close proximity to the UE.

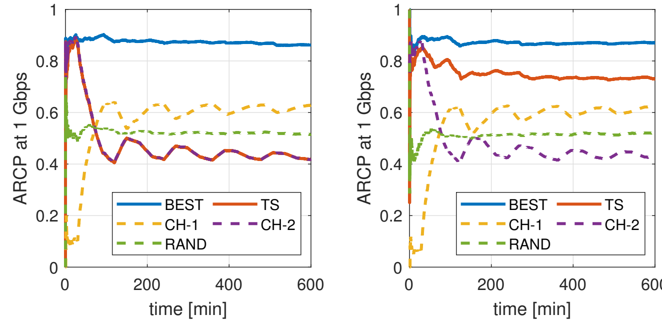
B. Impact of Blockage Dynamics

Next, we analyze the performance of the proposed scheme when the blockage characteristics of the channels are temporally varying. This is particularly important for application such as fixed wireless access or mm-wave backhaul where the



(a) $p_1 = 0.1$, $p_2 = 0.5$ (b) $p_1 = 0.4$, $p_2 = 0.2$

Fig. 2. Performance of Beam-switching scheme when blockage characteristics are static.



(a) Without Refresh (b) With Refresh

Fig. 3. Performance of TS with and without refreshing the learning parameters

receiver needs to switch beams to a different transmitter that provides the higher downlink rate or power in case the serving link is envisaged to remain blocked for a considerable amount of time. We assume that the dynamics of the blockage change after a random period of time which is Poisson distributed, and with the changes, the values of p_1 and p_2 cycle through the following values.

$$p_1 : 0.9 \to 0.2 \to 0.1 \to 0.3 \to 0.9 \ldots \quad (9)$$

$$p_2 : 0.1 \to 0.6 \to 0.9 \to 0.7 \to 0.1 \ldots \quad (10)$$

In case of temporally varying dynamics of the blockage characteristics, it is necessary for the TS algorithm to *unlearn* reward observations older than a certain time period t_W (which obviously depends on the frequency of the change in blockage statistics, governed by λ_T). To validate this, in Fig. 3 we plot the performance of the algorithms with and without steps 10-12 in the Algorithm 1. From Fig. 3a we see that in case the TS algorithm does not *reset* its parameters (i.e., does not unlearn the observations older than t_W), the performance of the TS is worse than that of BS_1, when the blockage characteristics of the channels change. Specifically, when the blockage in the channel 2 was lower ($p_2(t = 0) = 0.1$ and $p_1(t = 0) = 0.9$), naturally, the TS algorithm selects BS_2 as the serving BS. Then as the blockage characteristics evolve over time (e.g., $p_1(t = 200 \text{ min}) = 0.1$ and $p_2(t = 200 \text{ min}) = 0.9$), the previously collected rewards of the TS algorithm hinders efficient switching of the UE beam from BS_2 to BS_1. As a result of this, the UE camps in the BS_2 connection and continues to receive services from it, which is sub-optimal. Now we analyze the performance when we reset the TS parameters after a *refresh block* (t_W). In Fig. 3b, we employ the algorithm with $t_W = 12$ min. In this case, we observe that the proposed TS based beam-switching algorithm actually outperforms both BS_1 and BS_2 association. This is precisely due to the fact that the TS scheme efficiently manages to switch between the two BSs by observing the blockage dynamics of the channels. As a result, the UE opportunistically selects the BS which provides higher data rate.

C. Effect of Refresh Block Size

Naturally, the optimal value of t_W is dependent on λ_T, where the latter may be unknown to the network operator.

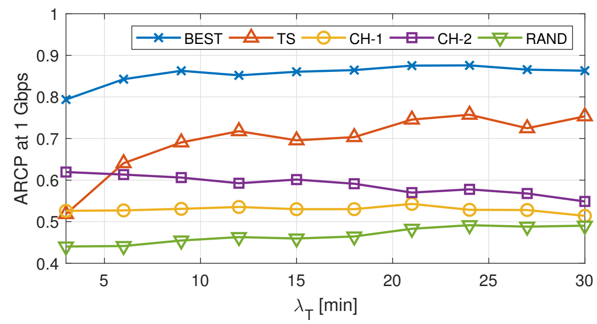


Fig. 4. Effect of varying λ_T on the performance of proposed beam-switching scheme, $t_W = 12$ min

In Fig. 4 we plot the performance of the TS based beam-switching scheme when for different values of λ_T, with $t_W = 12$ min. We observe that the performance of TS is significantly better than static association to either BS_1 or BS_2, as long as $t_W \ll \lambda_T$. It fares poorly in the case $\lambda_T \ll t_W$. This is due to the fact that when the blockage characteristics vary rapidly in comparison to the refresh block size, the TS algorithm is not able to keep up with the blockage dynamics. In such cases (which may occur during peak periods in city centers) we prescribe the operator to statically associate the UE with the best BS after a decision epoch. On the other hand, if the block size t_W becomes too small, the TS algorithm is not allowed to accumulate enough samples to learn from. This can be easily observed in Fig. 5a, where we have $\lambda_T = 30$ min and $t_W = 0.6$ min. Here the TS algorithm does not outperform BS_1 association, however, the performance is comparable with BS_1 association as the TS algorithm actually tracks the rewards of the BS_1. Even a slight increase in the value of t_W, improves the performance significantly (by 15%) as shown in Fig. 5b.

D. Discussions and Future Work

The results in this paper brings forth some open challenges. In a cellular network, a UE ideally would associate to multiple