

Machine Intelligence Technique for Blockage Effects in Next-Generation Heterogeneous Networks

Samuel AMALORPAVA MARY RAJEE, Arulraj MERLINE

Dept. of Electronics and Communication Engineering, Sethu Institute of Technology, Pulloor, Kariapatti 626 115, India

maryrajeece@sethu.ac.in, merlineece@sethu.ac.in

Submitted April 11, 2020 / Accepted June 19, 2020

Abstract. Millimeter wave (mmWave) links such as 28 GHz and 60 GHz propose high data rates and capacity needed in 5G Heterogeneous network (Hetnet) real-time system. The key factors in network planning of Hetnet are the locations and links of base stations, and their coverage, transmitted power, antenna angle, orientation etc. However, large-scale blockages like static buildings, human etc. affect the performance of urban Hetnets especially at mmWave frequencies. A mathematical framework to model dynamic blockages is adapted and their impact on cellular network performance is analyzed. A machine learning approach based on Q-learning with Epsilon-Greedy algorithm is proposed to solve the blockage problem in such complex networks. The proposed results are evident and show the positive effect of increasing the base station density linearly with the blockage density to maintain the network connectivity. The performance of the proposed Epsilon-Greedy algorithm is compared with Epsilon-Soft algorithm. The performances of above said mmWave links are compared in terms of their coverage probability and throughput. The results show that an Epsilon-Greedy algorithm outperforms an Epsilon-Soft algorithm.

Keywords

Heterogeneous network, millimeter wave, dynamic blockage, Q-learning, Epsilon-Greedy algorithm

1. Introduction

Recent advancements in Machine Type Communication applications such as autonomous driving, eHealth, Augmented/Virtual Reality (AR/VR), and tactile Internet are posturing tremendous challenges regarding capacity, reliability, latency, and scalability. In order to support such mission critical services, it requires more spectral bandwidth, to overcome the problem of capacity shortage. Most promising millimeter wave frequency bands such as 28 GHz and 60 GHz are contributing over 1 GHz of bandwidth [1]. However, the mmWave propagation characteristics are qualitatively different compared to those in the lower frequencies.

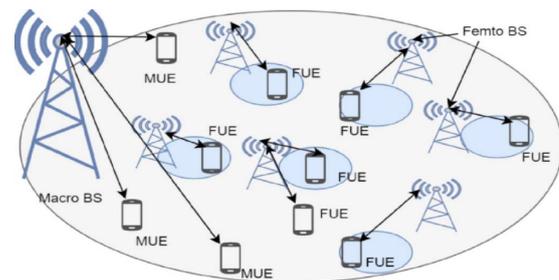


Fig. 1. Example of a heterogeneous network.

First, mmWave are much more vulnerable to blockage and do not penetrate common walls and tinted glasses well. Second, the path loss is much higher in these frequencies, which limits the communication distance. The environment affects many network properties; particularly in fifth-generation (5G) networks with possible millimeter wave signals, this environmental influence is increased.

The sharp difference between Non Line-of-Sight (NLOS) and Line-of-Sight (LOS) channel models is a result of blockage and it is correlated to the city geometry. In general, blockage modeling is one of the necessary steps to analyze the millimeter wave networks. An illustration of k-tiers of Base Station (BS) is shown in Fig. 1, A Hetnet is a collection of k-tiers of Base Station (BS), distinguished by their transmit powers and densities. The two-tier Hetnet structure constitutes primary and secondary tier. The primary tier includes high power macro base station that serves macro user equipments (MUEs), while the second tier comprises of low power femto base station that serves femto user equipments (FUEs). All users are assumed to be operated in the same mmWave spectrum.

2. Related Work

Because of NLOS limitations of mmWaves, blockage modeling is more important in 5G networks. One can use the probability of LOS function to find out whether a link is LOS or not, like the multi-ball model introduced by [2]. One other option is to use stochastic models to generate building footprints and to place base stations (BSs). The authors in [3] placed random-sized and orientation rectangles to model the buildings, which is a good fit in rural

areas. BSs are deployed according to a Poisson Point Process (PPP) to analyze the network properties, In many situations, these models are the preferred approach, as they can be produced easily. However, they might not be accurate enough. By using Boolean scheme modeling, BS and building locations are distributed, the authors in [4] analyzed the possible rates and coverage in mmWave networks.

A mmWave Hetnet link may suffer with three kinds of blockages, namely, static, dynamic, and self-blockage. Static blockage is due to trees, buildings and permanent structures. It may cause permanent blockage to the LOS link. However, the impact of static blockage is smaller when compared to dynamic blockage for an open area; it is due to mobile blockers such as moving humans and vehicles. It may cause frequent interruptions to the LOS link. Self-blockage also plays a vital role in interrupting mmWave performance. The authors of [5] analyzed self-blockage through simulation. Furthermore, in this paper, an urban setting is considered where static blockages may block the LOS paths between user equipments (UEs) and BSs, and at the same time provide NLOS paths between UE and BSs [6]. This paper provides an insight about dynamic blockage analysis and derives the blockage probability and duration considering the blockage of LOS as well as NLOS paths using M/M/ ∞ queuing model. In this queuing model the blockers are distributed according to Poisson process and the blockage duration of a single blocker is exponentially distributed with parameter μ .

Reinforcement learning (RL) enables users to dynamically adapt to the varying environment in order to optimize an unknown system by interacting with it [7]. Apart from this signal degradation, certain other parameters such as blockage frequency and duration affect the performance. With this objective in mind, the aim of this paper is to apply the machine learning RL method to deploy a self-organizing dense Hetnet with femtocell based on the blocker density and analyze the performance of mmWave links in the presence of dynamic blockage.

3. System Model

A Hetnet is a collection of K-tiers of BSs, distinguished by their transmit powers and BS densities. A single tier heterogeneous cellular network is modeled via homogeneous PPP Φ_T of density λ_T . The BSs are operated in a mmWave frequency band, BSs are distinguished by their fixed power transmitted (PT) and blockage model parameters.

Consider a circular disc $B(0, R)$ of radius R where a typical user is located at the origin. Each BS in the disc is the serving source to the user. Thus the total number of base stations can be deployed in the area πR^2 following a PPP distribution with parameter $\lambda_T \pi R^2$,

$$P(n) = \frac{[\lambda_T \pi R^2]^n}{n!} \exp[-\lambda_T \pi R^2]. \quad (1)$$

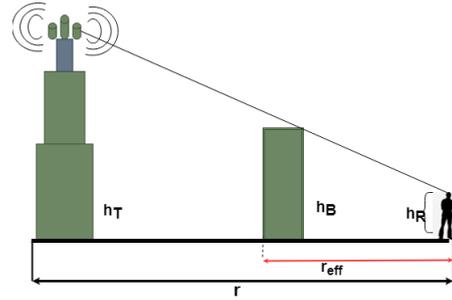


Fig. 2. Dynamic blockage model.

The BSs distances from the user $\{R\} \forall=1,2,3,\dots,n$ are independent and identically distributed (iid).

3.1 Dynamic Blockage Analysis

The mobile blockers are distributed with density parameter λ_B according to a homogeneous PPP. The Poisson arrival rate of the mobile blockers is α_i whilst crossing the i^{th} BS-User link. The blockage duration is exponentially distributed with parameter μ and it is independent of the random blocker arrival process.

Figure 2 portrays a random blocker when moving at an angle θ ranges from $[0, 2\pi]$ relative to the link. Here blocker's velocity V , frequency of blockage, height of the BS (transmitter) h_T , blocker height h_B , user height (receiver) h_R , distance r and the blocker arrival rate α are the blockage components involved.

The blockage rate α_i of the transmitter (BS) and user link (receiver) r_i is

$$\alpha_i = C r_i \quad (2)$$

in which C is a blockage component and is directly proportional to blocker density λ_B as follows

$$C = \frac{2}{\pi} \lambda_B V \frac{h_B - h_R}{h_T - h_R}. \quad (3)$$

The blockage duration is an exponential random variable with mean μ . Since the random arrival during blockage in user link is independent of other links, therefore the expected blockage duration of all n BS is calculated as

$$E[T_b] = \frac{1}{n\mu} \quad (4)$$

where $[T_b]$ is the duration of all n base station.

3.2 Blockage Probability

The blockage probability $P(B)$ is defined as the function of statistical properties of mean duration of mobile blockers and angle of movement.

$$P(B) = \exp[-ap\lambda_T \pi R^2] \quad (5)$$

where

$$a = 1 - \frac{2RC}{3\mu}, \quad p = 1 - \frac{\theta}{2\pi}.$$

For large ultra dense Hetnet, a and p are variables with regard to mobile blockers and BS density λ_T , the coverage probability $\mathbb{P}(c)$ should be 1. In order to meet the requirement $P(B)$ should be less than a predefined threshold

$$P(B) = \exp[-ap\lambda_T\pi R^2] \leq \Gamma. \quad (6)$$

A user is considered to be in LOS and mmW band is selected with the following LOS probability

$$p_{\text{LOS}}(r) = \exp[-P(B)]r. \quad (7)$$

Here, r is the direct distance between the user and the connected BS and $P(B)$ is the one-step transition blockage probability due to blockage parameter in the considered region. They are computed using statistical parameters of the buildings such as density, height and the average size of the blockages. In this blockage region, the NLOS probability of the link is expressed as

$$p_{\text{NLOS}}(r) = 1 - p_{\text{LOS}}(r). \quad (8)$$

The path loss model used in the network for the mmWave link, in dB, is modeled as,

$$L_{\text{mm}}(r) = \begin{cases} 20\log\left(\frac{4\pi}{\lambda}\right) + 10\alpha_L \log(r) + \chi_{\text{LOS}} & \text{if LOS} \\ 20\log\left(\frac{4\pi}{\lambda}\right) + 10\alpha_N \log(r) + \chi_{\text{NLOS}} & \text{if NLOS} \end{cases} \quad (9)$$

where $L_{\text{mm}}(r)$ denotes distance dependent path loss of mmWave link, r is the radial distance between the transmitter and the receiver, λ is the wavelength, which models the effects of χ_{LOS} shadow fading. The path loss exponents in LOS and NLOS mmWave links are denoted α_L and α_N respectively.

The Signal to Noise plus Interference Ratio (SINR) coverage probability has been characterized as a performance metric [18].

$$\text{SINR}_{\text{FUE}_i} = \frac{P_f h_0 G_e L_{\text{mm}}(r)}{\sigma^2 + I_i} \quad (10)$$

in which P_f represents transmission power of mmWave BS and h_0 is the channel gain. $G_e = G_{\text{TX}} G_{\text{RX}}$ It represents the effective antenna gain of the transmitter and the receiver. $L_{\text{mm}}(r)$ represents the path loss experienced by the user. σ^2 denotes the additive white Gaussian noise (AWGN).

3.3 Coverage Probability

The coverage probability of an arbitrarily located user (User Equipment) in a Hetnet has been defined as it can connect at least one accessible tier, and the probability of coverage of user is the probability that the SINR from the candidate serving BSs of any tiers exceeds the corresponding threshold

$$\mathbb{P}_{\text{cov}}(i) = \mathbb{P}[\text{SINR}_{\text{FUE}_i} \geq \Gamma], \quad (11)$$

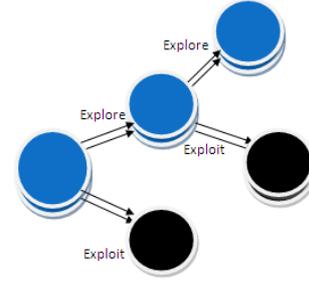


Fig. 3. Example of states with action for Epsilon-Greedy algorithm.

$$\mathbb{P}_{\text{cov}}(i) = p_{\text{LOS}}(r) \mathbb{P}_{\text{cov}}(\text{mm}) + p_{\text{NLOS}}(r) \mathbb{P}_{\text{cov}}(\text{mm}). \quad (12)$$

The capacities of the femtocell are calculated as:

$$C_{\text{FUE}_i} = \log_2(1 + \text{SINR}_{\text{FUE}_i}), i = 1, 2, 3, \dots, k. \quad (13)$$

4. Utilizing Machine Intelligence in Hetnet

In an online decision-making situation such as the reinforcement learning problem, an agent is faced with two choices - explore or exploit. The term exploration in an active learning system is defined as the process of deliberately taking a non-greedy action with the sole aim of gathering more information about the environment [8]. Alternatively exploitation is the act of taking the best possible action given the current information about the environment. A central challenge in reinforcement learning is to find the sweet spot between exploration and exploitation [9], i.e., to figure out when to explore and when to exploit. In order to maximize the coverage probability, threshold level is set for blockage density.

Figure 3 shows the Q-learning Epsilon-Greedy algorithm procedure with decision making choices of either to exploit or explore. Each femto BS pays out according to a different probability distribution and these distributions are unknown to user. As the femto BS is operated, the track of the average probability of each femto BS is taken into account. Then, the user selects the femto BS with the highest current average probability.

4.1 Contribution

A new Q-learning approach with Epsilon-Greedy algorithm is proposed to provide better coverage throughout the whole heterogeneous network. The following contributions are

- Cooperative Q-learning is achieved through Epsilon-Greedy algorithm for balancing exploration/exploitation assigned to femto base stations to carry out blockage threshold between them.
- Based on reward function, the required coverage for each femto user is achieved, so as to increase the BS density of the network.

- Dynamic blockage analysis is solved analytically and comparative results are shown to solve blockage problem in ultra dense Hetnet.

4.2 Q-Learning with Epsilon-Greedy Algorithm

The Q-learning problem consists of an environment and one or more agents. It is used to interact with the dynamic environment through trial-and-error interactions. Through interactions, decisions are made, which is modeled as an Epsilon-Greedy algorithm. The agent can move between different states by choosing different actions using state transition probability. The trace of actions that the agent takes determines the policy. The agent will receive a reward from the dynamic environment with each transition; a cumulative reward is stored as a consequence of action. The agent will continue its actions to maximize a reward. FUEs are assumed to be agents; and they can come across an optimal decision policy through trial and error interaction with the dynamic environment [10–13].

Here the objective is to maximize the coverage probability of the femto Base Station (FBS) with the limitation of blockage probability of the BS-user link. In the context of a mmWave Hetnet, femto BS acts as an intelligent agent in the Q-learning algorithm.

Q-learning can be defined as a function approximated in which the values are dependent on the state $x(t)$ and action a_t at time step t [6]. The one-step Q-learning approach form is defined as follows

$$Q(x_t, a_t) \leftarrow (1 - \alpha)Q(x_t, a_t) + \alpha \max R_t + (Q_{t+1}, a). \tag{14}$$

R_t is the reward at time step t and $0 < \gamma < 1$ is the discount factor. The optimization of coverage probability $P(C)$ can be formulated as

$$Q(x_t, a_t) = \max P(C)[R_t + \gamma(Q_{t+1}, a)]. \tag{15}$$

Epsilon-Greedy algorithm is a way of selecting random actions with uniform distribution from a set of available actions. Figure 4 explains the flow of actions performed using the Epsilon-Greedy algorithm. Using this, either we can select random action with epsilon probability or we can select an action with $(1 - \text{epsilon})$ probability that gives maximum reward in given state. In the dynamic environment there are basically two states for BS (LOS and NLOS), Epsilon-Greedy optimization is done through cooperative Q-learning and their rewards are drawn by choosing the threshold probability greater than the blockage probability, hence the overall coverage will be increased and the preferred Hetnet environment is created. All the past and critical values are stored and fed back to Q-learning for improvement.

1) Actions: $A = a_1, a_2, \dots, a_n$ Among these actions, Femto BS chooses its coverage probability, which is the coverage space between transmitter and receiver. In an unknown environment, the Femto BS chooses actions

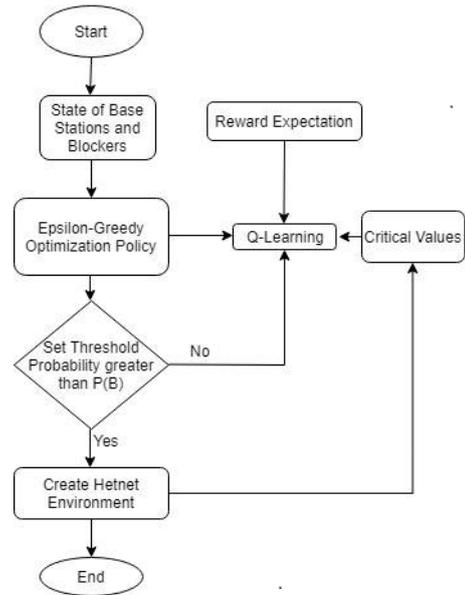


Fig. 4. Flow diagram for Epsilon-Greedy algorithm.

with the different probability. Therefore, states are chosen between $P(C)_{\min}$ and $P(C)_{\max}$.

2) States: States are chosen based on the vicinity of the FBS to the user equipment. The two parameters are defined as state of each FBS

- 1) blockage probability $P(B) \leq \bar{P}$,
- 2) coverage probability $P(C) > \bar{P}$.

Threshold probability is \bar{P} , The state of the FBS I at time step t is defined as $S_t(I)$. For each FBS, a Q-table is constructed with all possible states as its rows and actions as its columns. In the proposed system model, the state remains constant as long as its location based on its coverage probability is fixed.

3) Reward: The reward function targets the objective of the Q-learning method.

Here it is possible to observe the need to explore new optimization mechanism to set the blockage density threshold, because the problem of associating blocker density to base station density is hard and not computable even for small sized Hetnets. The objective of the optimization problem is to maximize the capacity of the network by maximizing the coverage probability with regard to blockage probability, so a higher capacity for user results in a higher reward. One reasonable solution is to act greedy with probability $1 - \epsilon$ (exploiting) and act randomly with probability ϵ (exploring). There is a trade-off between exploitation and exploration for different values of ϵ . The fairness performance to explore and exploit using epsilon-greedy algorithm in terms of Jain's fairness index is defined as follows

$$f(x_1, x_2, \dots, x_n) = \frac{\left(\sum_{i=1}^J x_i \right)^2}{J \sum_{i=1}^J x_i} \tag{16}$$

A fairness value equal to 1 is achieved when all FUEs have the same capacity. The fairness index is consistent and reaches the value of 1.

5. Numerical Evaluation and Results

The analytical results are compared with MATLAB simulation where a rectangular area of 200 m × 200 m is considered, and mobile blockers are located uniformly in this area. Our area of interest is the circular disc $B(0, R)$ of radius $R = 100$ m, the mobile blockers randomly chose a direction, and move in that direction for a blockage duration $t = 20$ ms, the probability of blockage and expected blockage frequency interruptions are $1e-5$. The simulation parameters are tabulated in Tab. 1.

The derived expressions for blockage probability and blockage frequency as a function of the density of the BS and blockers are verified and shown in Fig. 5. The minimum BS density required to guarantee in the midst of critical dynamic blockage is derived and shown.

The results in Fig. 5 shows for a blocker density of 0.1 bl/m², a BS density of 100/km² can decrease the interruptions to once in ten seconds, 200/km² can decrease the frequency of interruptions to once in 100 seconds, and 300/km² can decrease the same to once in 1000 seconds. It is crucial to reduce frequency of interruptions particularly for mission critical applications.

It is observed that the mmWave Hetnets are blockage limited instead of capacity limited. So as to optimize the value of blockage probability threshold as 0.35, Q-learning based on ϵ -greedy algorithm is proposed and the simulation results are derived using MATLAB.

The values of learning rate and discount factor are set to be $\alpha = 0.5$ and $\gamma = 0.5$ respectively. The first 80% of iteration is performed by using epsilon-greedy algorithm with random epsilon = 0.9. The total number of iterations for Q-learning is set to 2000. The agents are randomly added to the network.

Symbol	Parameters	Values
R	Radius	100 m
V	Velocity of blockers	1 m/sec
h_b	Height of blockers	2 m
h_r	Height of user equipment	1.4 m
h_T	Height of transmitter (FBS)	5 m
$E(T_B)$	Expected blockage duration	0.5 sec
P_f	Femto BS power	20 dBm
G_e	Antenna gain	10, -10 dBi
B_{mm}	Bandwidth	100 MHz
f_{mm}	Carrier frequency	28 GHz, 60 GHz
χ_{LOS}	LOS mmWave Log normal shadowing	$\mu = 0, \sigma = 5.2$ dB
χ_{NLOS}	NLOS mmWave Log normal shadowing	$\mu = 0, \sigma = 7.38$ dB
σ^2	Noise power	$174 + 10\log_{10}B_{mm} + 10$ dBm

Tab. 1. Simulation parameters.

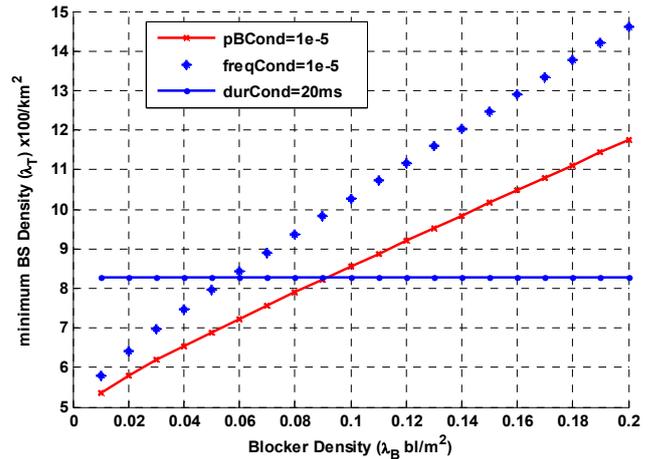


Fig. 5. Expected blocker density on coverage event vs. Base station (BS)

For each number of agents, the algorithm performs through all iterations. Finite number of iterations and fixed ϵ for a single Femto BS network (single agent network) are considered in this Q-learning approach, ϵ -greedy policy with S_i as the size of the state-space where in this analysis $S_i = 10$ is fixed for all the iterations. The result in Fig. 6 shows that the learning of the environment takes up to 15 iterations. Then it captures the performance parameter and achieves reward as 1. The results of this paper show the application of machine learning to address network planning based on its blocker density in dense Hetnets. In order to maintain the network connectivity, the optimized base station density should scale super linearly with the blockage density. The epsilon-greedy algorithm is mainly driven by randomness. In addition, they are stateless, i.e. the exploration is driven without knowing which areas of the state space have been explored.

The performance of Epsilon-Greedy algorithm is compared with so called Epsilon-Soft algorithm and shown in Fig. 7. It is stochastic deterministic policy but finite, probability of selecting any possible action. Having a policy which has some chance of selecting any action is important theoretically when rewards and state transitions are stochastic.

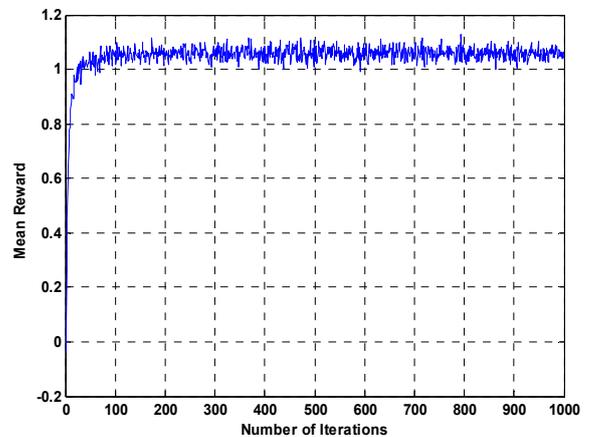


Fig. 6. Performance of the reward function over iterations.

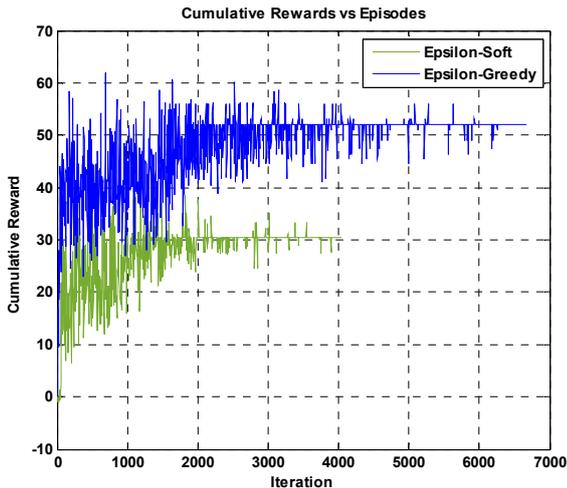


Fig. 7. Comparative performance of the epsilon-greedy and soft algorithms.

The first observation is that algorithms are affected differently by variations in terms of their average and cumulative rewards. The outcome of an Epsilon-Greedy algorithm in terms of their cumulative rewards is twofold times higher. The second observation is their convergence behavior. Due to greedy learning Epsilon-Greedy algorithm works well for optimization, hence converges a little late compared to an Epsilon-Soft algorithm

The LOS probability for mmWave is given in (7) and (8). Assume the value of pathloss exponent is 2 for making a mmWave link to be LOS. The average cell radius in a network is defined as a ball of a radius $R = 100$ m and that has the size of an average cell.

The operating frequency of mmWave network are assumed at 28 GHz and 60 GHz and the bandwidth assigned to each user is $B_{mm} = 100$ MHz. The path loss exponents for LOS and NLOS links α_L and α_N are 2 and 4, respectively. The performance of next generation cellular networks and their various enablers need to be evaluated in order to assess their effectiveness towards improvement of various performance metrics. The effect of shadowing and pathloss for the two different mmWave links 28 GHz and 60 GHz are presented in Fig. 8 and Fig. 9.

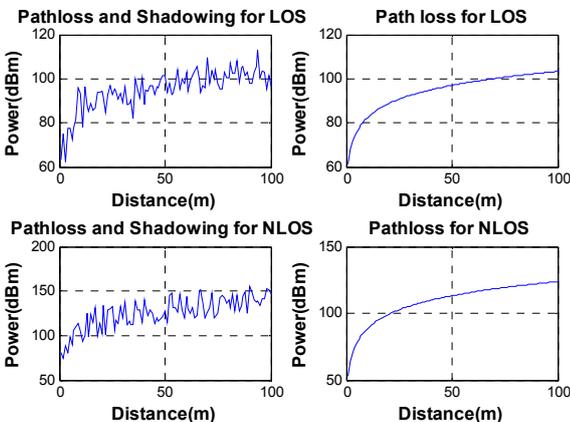


Fig. 8. Pathloss and shadowing for 28 GHz mmWave link as a function over distance.

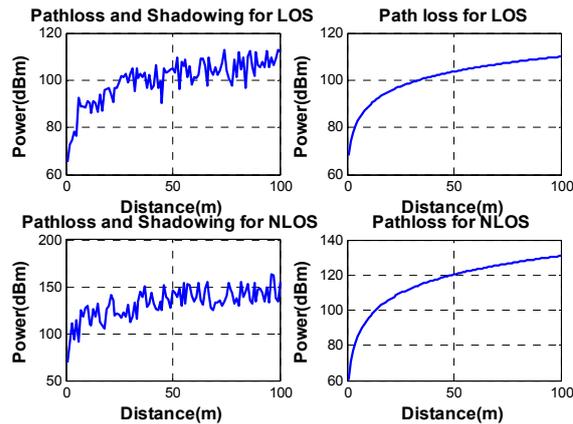


Fig. 9. Pathloss and shadowing for 60 GHz mmWave link as a function over distance.

The performance comparison is made for mmWave links and it is presented in Fig. 10. It is good enough to achieve a high data rate. 60 GHz mmWave communication SINR coverage is shown to outperform 28 GHz coverage. This is achieved, due to higher mmWave association with the indoor and it admits greater bandwidth. However, there are very less possibilities for LOS links, due to greater density of building blockages and more multi-paths [14–16] Hence, the indoor users can have an association with the 28 GHz links and it highly depends on the distance and the selection of optimal bias. The distribution mechanism is extremely important to achieve useful gains and to satisfy QoS requirements.

The training phase of an agent is set to 10000 iterations to compare the throughput performance of the mmW links. Figure 11 shows the average throughput accuracy and the average training times for each considered epsilon-greedy algorithm.

The training phase of an agent (FBS) is shown in Fig. 12 which means that the iteration algorithm employed more iterations to find optimal value. Because of the stochastic nature of the environment samples, the algorithm

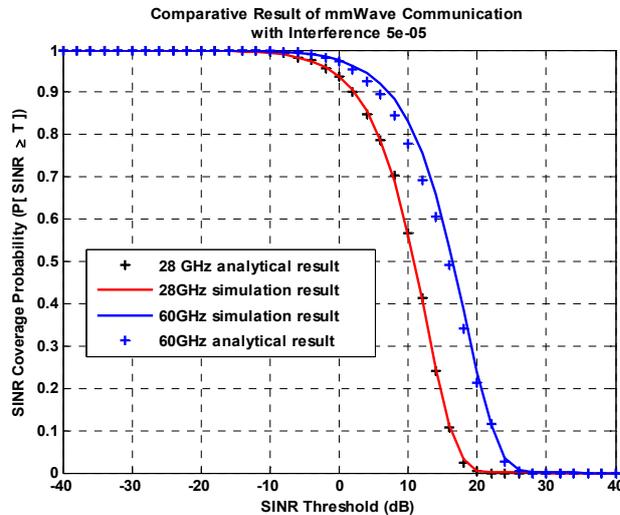


Fig. 10. Performance of the mmWave links as a function of SINR coverage probability.

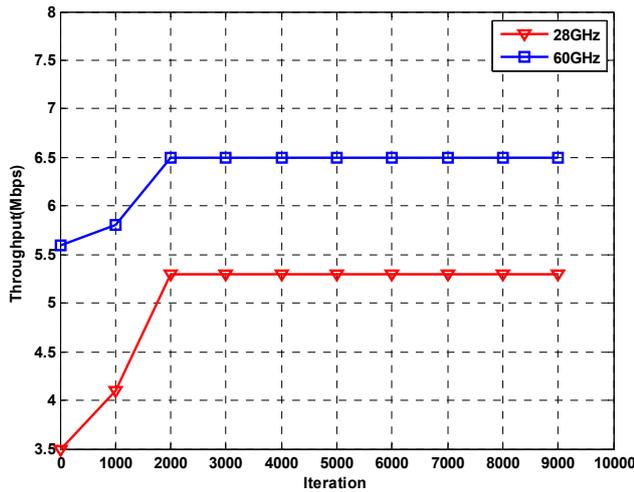


Fig. 11. Throughput performance of the mmWave links over iterations.

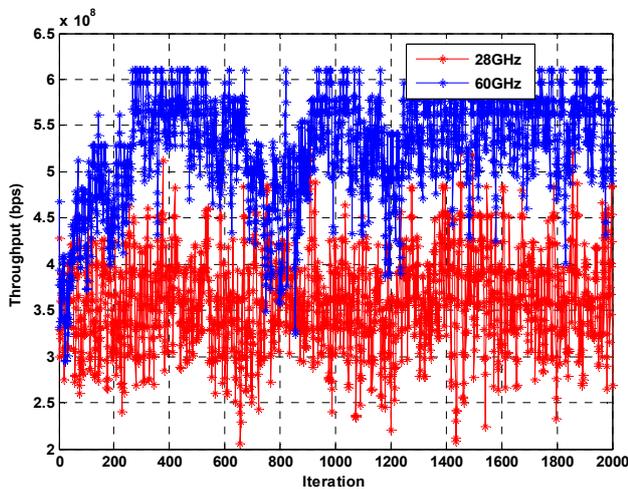


Fig. 12. Performance of the throughput function during training phase over the first 2000 iterations.

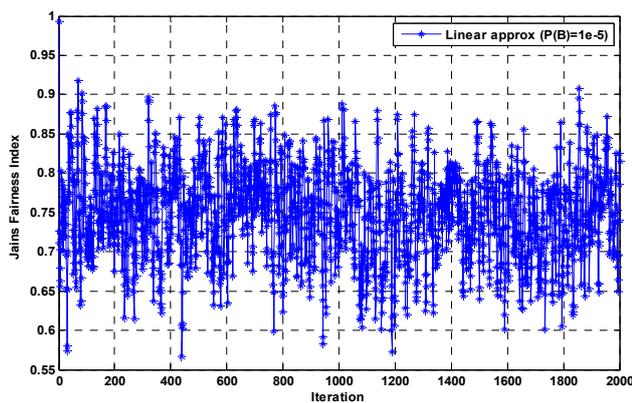


Fig. 13. Jains Fairness Index function over iterations.

might have dealt with tough values to determine optimal value. However, we can see that in those cases, although more time was needed, the goal of improving the throughput was achieved. The convergence of the proposed epsilon-greedy algorithm is examined through Jains fairness index. For every state and action pair $Q(s, a)$, the value of

$Q(s, a)$ is the change between the Q-value before and after the action is implemented at each iteration.

Figure 13 shows that the throughput fairness increases quickly with the increase of iteration. So from the figure we can see the epsilon-greedy algorithm can receive good convergence performance, which let the heterogeneous network system be always in a stable state. It reaches to 1 after its training phase.

6. Conclusion

In this paper, an epsilon-greedy algorithm has been proposed to solve dynamic blockage problem in mmWave enabled heterogeneous networks. The blockage problems in two different mmWave cellular networks (28 GHz and 60 GHz) have been analyzed. On one side, the existence of femto BS leads to unequal distribution, on the other side; it is a vital issue to provide an efficient backhaul transmission for these small cells. Extensive overviews of the existing blockage problems are discussed. The dynamic blockage is approached in different way and the possible solutions are provided to enhance its performance. The proposed epsilon-greedy algorithm is compared with epsilon-soft algorithm in terms of their cumulative rewards and convergence speed. The proposed epsilon-greedy algorithm achieved lower blockage probability performance and could accommodate more users while guaranteeing the Quality of Service (QoS) of users. Due to greedy learning the proposed algorithm works well for optimization. Future work will explore different methods of achieving interference control to optimize the network density.

References

- [1] KARJALAINEN, J., NEKOVEE, M., BENN, H., et al. Challenges and opportunities of mm-wave communication in 5G networks. In *The Proceedings of 9th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*. Oulu (Finland), 2014, p. 372–376. DOI: 10.4108/icst.crowncom.2014.255604
- [2] BAI, T., HEATH, R. W. Coverage analysis for millimeter wave cellular networks with blockage effects. In *The Proceedings of IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. Austin (TX, USA), 2013, p. 727–730. DOI: 10.1109/GlobalSIP.2013.6736994
- [3] LU, W., DI RENZO, M. Stochastic geometry modeling of mmWave cellular networks: Analysis and experimental validation. In *The Proceedings of IEEE International Workshop on Measurements & Networking*. Coimbra (Portugal), 2015, p. 1–4. DOI: 10.1109/IWMN.2015.7322991
- [4] FERAYDOONI, M., SABAEI, M., DEGHAN, M., et al. A mathematical framework to evaluate flexible outdoor user association in urban two-tier cellular networks. *IEEE Transactions on Wireless Communications*, 2017, vol. 17, no. 3, p. 1559–1573. DOI: 10.1109/TWC.2017.2780824
- [5] ABOUELSEUD, M., CHARLTON, G. The effect of human blockage on the performance of millimeter-wave access link for outdoor coverage. In *Proceedings of the IEEE 77th Vehicular*

- Technology Conference (VTC Spring)*. Dresden (Germany), 2013, p. 1–5. DOI: 10.1109/VTCSpring.2013.6692780
- [6] JAIN, I. K., KUMAR, R., PANWAR, S. Limited by capacity or blockage? A millimeter wave blockage analysis. In *Proceedings of the 30th International Teletraffic Congress (ITC 30)*. Vienna (Austria), 2018, p. 153–159. DOI: 10.1109/ITC30.2018.00032
- [7] AMIRI, R., ALMASI, M. A., ANDREWS, J. G., et al. Reinforcement learning for self organization and power control of two-tier heterogeneous networks. *IEEE Transactions on Wireless Communication*, 2019, vol. 18, no. 8, p. 3933–3947. DOI: 10.1109/TWC.2019.2919611
- [8] SASIKUMAR, S. N. Exploration in feature space for reinforcement learning. *Master Thesis*. Australian National University, May 2017, p. 1–65. [Online] Available at: <https://arxiv.org/abs/1710.02210>
- [9] ALQERM, I., SHIHADA, B. Cognitive aware interference mitigation scheme for LTE femtocells. In *International Conference on Cognitive Radio Oriented Wireless Networks*. Doha (Qatar), 2015, p. 607–619. DOI: 10.1007/978-3-319-24540-9_50
- [10] VAN OTTERLO, M., WIERING, M. Reinforcement learning and Markov decision processes. In Wiering, M., van Otterlo, M. (eds) *Reinforcement Learning. Adaptation, Learning, and Optimization*. Berlin, Heidelberg (Germany): Springer, 2012, vol. 12. ISBN: 978-3-642-27644-6
- [11] GALINDO-SERRANO, A., GIUPPONI, L. Self-organized femtocells: A fuzzy Q-learning approach. *Wireless Network*, 2014, vol. 20, no. 3, p. 441–455. DOI: 10.1007/s11276-013-0609-6
- [12] WHITEHEAD, S. D. A complexity analysis of cooperative mechanisms in reinforcement learning. In *AAAI Proceedings*. 1991, p. 607–613. [Online] Available at <https://www.aaai.org/Papers/AAAI/1991/AAAI91-095.pdf>
- [13] AMIRI, R., MEHRPOUYAN, H., FRIDMAN, L., et al. A machine learning approach for power allocation in HetNets considering QoS. In *The Proceedings of 2018 IEEE International Conference on Communications (ICC)*. Kansas City (MO, USA), 2018, p. 1–7. DOI: 10.1109/ICC.2018.8422864
- [14] PENG, M., LIANG, D., WEI, Y., et al. Self-configuration and self-optimization in LTE-advanced heterogeneous networks. *IEEE Communications Magazine*, 2013, vol. 51, no. 5, p. 36–45. DOI: 10.1109/MCOM.2013.6515045
- [15] BAHADORI, N., NAMVAR, N., KELLEY, B., et al. Device-to-device communications in the millimeter wave band: A novel distributed mechanism. In *2018 Wireless Telecommunications Symposium (WTS)*. Phoenix (AZ, USA), 2018, p. 1–6. DOI: 10.1109/WTS.2018.8363940
- [16] BHATTI, O. W., SUHAIL, H., AKBAR, U., et al. Performance analysis of decoupled cell association in multi-tier hybrid networks using real blockage environments. In *The proceedings of the 13th IEEE International Wireless Communications and Mobile Computing Conference (IWCMC)*. Valencia (Spain), 2017, p. 62–67. DOI: 10.1109/IWCMC.2017.7986263

About the Authors ...

AMALORPAVA MARY RAJEE Samuel (corresponding author) was born in 1977. She received her Bachelor of Engineering (Electronics and Communication Engineering) from Madurai Kamaraj University in 1998 and Master of Engineering (Communication Systems) from Anna University in 2010. She is pursuing Ph.D. in Wireless Networks in Anna University and working as an Assistant Professor in Sethu Institute of Technology. Her research interests include heterogeneous networks, optimization, small cells and machine learning.

MERLINE Arulraj was born in 1974. She received her Bachelor of Engineering (Electronics and Communication Engineering) from Thanthai Periyar Government Institute of Technology in 1995 and Master of Engineering (Communication Systems) from Thiagarajar College of Engineering in 2002 and the Ph.D. from Anna University, Chennai in 2014. She is a Professor in Sethu Institute of Technology. Her research interests include multiple-input-multiple-output radar and space time adaptive processing.