Indoor Mobility Prediction for mmWave Communications using Markov Chain

Aysenur Turkmen*, Shuja Ansari*, Paulo Valente Klaine*, Lei Zhang*, Muhammad Ali Imran*
*James Watt School of Engineering, University of Glasgow, UK
a.turkmen.1@research.gla.ac.uk

***Abstract*—Millimeter-wave (mm-wave) communication, which has already been a part of the fifth generation of mobile communication networks (5G), would result in ultra dense small cell deployments due to its limited coverage characteristics. To enable seamless handovers between indoor and outdoor environments, a mobility prediction of an indoor user is studied by deploying Markov chains. Based on the effect of external factors on the user's mobility, a simulation scenario is created to model the trajectory of an indoor user w.r.t the most visited areas before leaving the indoor environment. Based on that, a method for initializing the transition matrix of Markov chains is proposed, via Q-learning. The proposed solution is compared to a standard online learning Markov chain model in terms of different mobility models and learning rates. Results show that the proposed solution is always able to outperform the standard method in terms of prediction accuracy.**

***Index Terms*—mm-wave 5G, Markov Chain, Q-Learning, indoor mobility, user trajectory, predictive handover, femtocells.**

I. Introduction

5G is already being deployed in major cities of China and the UK with the promise of achieving high data rate communications with minimum latency [1]. Millimeter wave frequencies, ranging in the spectrum from 3 to 300 GHz, are being exploited in 5G to meet the increased wireless data traffic coming from a range of smart connected devices. In spite of having a promising feature for wireless communication with a higher frequency range, mm-waves come along with a serious issue of penetration losses, due to the physical nature of high frequency radio communications with shorter wavelengths. A simple solution to this issue is the deployment of small cells (SCs), where mm-wave frequency driven BSs are placed closer, in order to provide a robust coverage in mm-wave frequencies [2]. However, without a careful management and regulation, the dense deployment of SCs may cause frequent handovers (HO). Switching from one BS's coverage to another BS based only on received signal strength (RSS), may degrade the Quality of Service (QoS) and Quality of Experience (QoE) of users. Additionally, the studies by 3GPP verify that the HO failure rate in a macro-pico heterogeneous network is high as 60%, which is doubled compared in a macro-only network [3]. Moreover, interference management, spectrum sharing, resource management, energy efficiency, user association, and the economics of this ultra dense network are some of the challenging areas that still need to be addressed.

In this respect, predictive mobility management is a promising candidate to provide seamless connectivity by predicting future locations and anticipating user equipment (UE) HOs as well as triggering the network BSs to make them ready for incoming HO requests. Predictive mobility management requires some input from the environment or the network itself, which can be fed, for example, into machine learning algorithms. Due to its importance, numerous studies have investigated mobility prediction in wireless communications. In [4] the authors performed mobility prediction via Markov chains, to show the influence of the transition probability matrix, which is created based on their assumptions on user movement. The authors in [5], expanded their work in [4] by using user's mobility history as an input to a transition probability matrix, to discover the most frequently visited base station. In [6], a machine learning based mobility management scheme for 4G X2 HO process is proposed to predict future HOs in order to reduce the HO delay. The authors introduced the concept of 3D transition matrix to address the path dependency problem of classical Markov chain, which occurs when users perform the HOs to the same cell. In [7], a mobility prediction model was developed where they assigned Markov Chains to the data plane network in 4G. They proposed a trajectory dependency parameter that can control their proposed model's reaction to random and less frequent movements.

However, most studies consider mobility prediction only in outdoor environments, ignoring the fact that almost 80% of

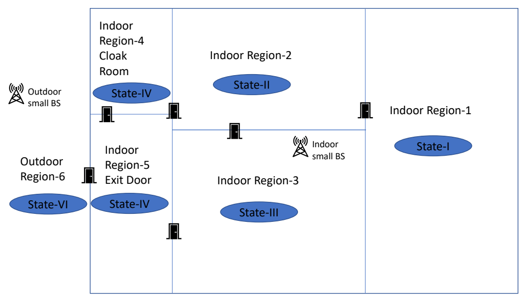


Fig. 1: System Model showing Indoor Regions as Markov Chain states.

mobile traffic is generated by indoor users [8]. Thus, given that 5G networks are expected to be much more dense, with the deployment of mm-wave SCs occurring both indoors and outdoors, it is essential that mobility prediction capabilities are provided, especially when users are transitioning from indoors to outdoors as high penetration loss prone mm-wave driven indoor environment expected to become isolated from the outdoor [9]. Based on that, in this article we propose a user-centric indoor mobility prediction algorithm for preemptive handovers based on Markov chains. Given that indoor user mobility is not completely random, we propose a novel method for indoor user mobility prediction and compare it with traditional Markov chains from [7].

The remainder of this paper is as follows: Section II describes the methodology for indoor mobility prediction, Section III presents the proposed solution, Section IV discusses simulation results, and conclusions are drawn in Section V.

II. Methodology

A. Markov Chain for Mobility

Since almost 80% of mobile traffic is generated by indoor users, the simulations in this study are based on tracking the user's mobility in an indoor environment [8]. In addition, it is well known that user movements have some pattern and are not completely spontaneous, but rather target oriented, such as going to the train station or heading to the kitchen from the living room [10]–[12]. In [12], user mobility is modelled in a non-random manner, inspired by the process of human decision making in [13]. The authors in [13] looked that process by taking into account internal and external factors where the former is represented by individual characteristics and the latter is indicated by environment stimulus and group behaviour. Motivated by the above mentioned studies, our hypothesis is that a user has more regularities in their movement within an indoor environment, where degrees of freedoms are lower as compared to an outdoor scenario. Considering these regularities, we designate a special area in our indoor model, called the cloak room, where users usually visit to take his/her coat, shoes, keys, umbrella etc., before going outdoor, or vice-versa. Therefore, we model an indoor environment segmenting indoor regions (IR) into Markov Chain states as shown in Fig. 1, and prediction algorithms are implemented to track the probabilities of user following the given scenario trajectory. Moreover, we also consider in this environment that a single BS is located in the outdoor environment, positioned at a distance of d from the left wall of the user's building. Regarding the indoor environment, a building with an area of A is considered, with a single small cell providing coverage for the entire region.

B. Markov Chain for Mobility Prediction

Markov chain is a stochastic process and is referred as memory-less, since the next state relies on the current state rather than the previous state [5]. A Markov chain consists of a set of states, which in our scenario are S={IR_1, IR_2,..IR_n}, where being the states'indices $\mathbb{I} = \{1, 2, ..., n\}$ and transitions, $\mathrm{t_{i,j}}$, represent the movement probabilities from one state to another, as illustrated in Fig. 2. Markov chains are mainly used for predictions in a randomly changing system, and mathematically models the probabilities of transitions to the next states, as:

$$\mathbb{P}(S_{\text{n+1}} = s_{\text{n+1}} | S_{\text{n}} = s_{\text{n}}, ..., S_1 = s_1) = \mathbb{P}(S_{\text{n+1}} = s_{\text{n+1}} | S_{\text{n}} = s_{\text{n}}). \quad (1)$$

Contrary to the studies mentioned in Section I where the states are defined as base stations; our proposed scheme defines the Markov chain states as indoor regions (IR) within an indoor environment. Received signal strength (RSS) approach which is one of the simplest and broadly used techniques for indoor localization [14] is utilized to determine in which state UE is.

The probability distribution is derived from:

$$\mathbf{p}_k = \mathbf{p}_0 \mathbf{T}^k, \quad (2)$$

$$\mathbf{T} = \begin{bmatrix} t_{1,1} & t_{1,2} & \cdots & t_{1,n} \\ t_{2,1} & t_{2,2} & \cdots & t_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ t_{n,1} & t_{n,2} & \cdots & t_{n,n} \end{bmatrix}, \quad (3)$$

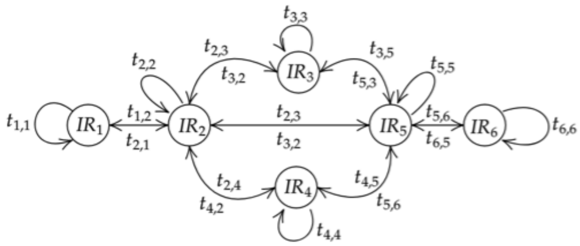


Fig. 2: Discrete-time Markov Chain with 6 finite state spaces (i.e., IRs).

where $\mathbf{p}_k$ is the k$^{\text{th}}$ transition probability vector, $\mathbf{p}_0$ is the initial distribution vector and $\mathbf{T}$ is the transition probability matrix.

III. Proposed Solution

Given that user movements are goal oriented, in this paper we propose a novel concept for initialising the transition matrix of a Markov chain, and evaluate its impact in indoor mobility prediction. However, before presenting the proposed solution, it is important to give an overview on how Markov chains can be used for mobility prediction.

1) Online Learning Transition Matrix: There are some steps needed to be set before initializing the transition matrix such as: 1) A transition from any state to itself is prohibited, making the transition matrix hollow, such that $t_{i,i} = 0, \forall i \in \mathbb{I}$; and 2) The transition matrix should be a right stochastic matrix, satisfying the condition $\sum_{j=1}^{n} t_{i,j} = 1, \forall i \in \mathbb{I}$.

Since the UE initiates the HO transition, we assume that the transition matrix $\mathbf{T}$ is updated according to the user's tracked movement. The idea is to assign higher probabilities to the most common routes followed by the user as compared to the other routes. A trajectory dependency parameter R_d is used to control the model's learning rate and reaction to random or less frequent movements as proposed by [7], where $0 \leq R_d \leq 1$. Consequently, small values of R_d update the transition matrix more slowly, giving more weight to the overall path of a user (minimizing the randomness). In the case of $R_d = 0$, $\mathbf{T}$ is not updated making the prediction independent of the past movement, whereas, in the case of $R_d = 1$, the prediction is biased towards the most recent trajectory.

To further explain the update procedure of $\mathbf{T}$, let us consider an example where a user follows the path: $IR_1 \longrightarrow IR_2 \longrightarrow IR_3$. For each movement between a region, e.g., from IR_1 to IR_2 the UE will update the probabilities of outbound movements from IR_1 to all neighbouring IRs in a game scheme of several stages. In the first stage, the outbound movement probability of UE from IR_1 to IR_2 is increased by an amount controlled by R_d, while the probabilities of direct movement of UE from IR_1 towards all playing IRs are decreased. This is expressed as:

$$t_{1,2} = t_{1,2} + \sum_{j} t_{1,j} R_d, j \in \mathbb{N}_{IR_1}, \tag{4}$$

$$t_{1,j} = t_{1,j} - \frac{\sum_j t_{1,j} R_d}{|\mathbb{N}_{IR_1}| - 1}, j \in \mathbb{N}_{IR_1}, \tag{5}$$

where $|\mathbb{N}_{IR_1}|$ is the cardinality of the set of neighbouring IRs for IR_1 which are taking part in the game. To satisfy the condition of inclusivity ($0 \leq t_{i,j} \leq 1$), a lower bound of 0 and an upper bound of 1 is set for each entry in $\mathbf{T}$. This brings in the challenge of satisfying the condition of right stochastic matrix. This is solved by adding additional stages that approach equilibrium without violating the conditions of transition matrix [7].

2) Q-Learning Initialization of the Transition Matrix: Based on the model proposed by [7] and the fact that user mobility is not totally random, but rather goal oriented, in this paper we propose to initialize the transition matrix $\mathbf{T}$ according to a Q-learning algorithm. Q-learning, is a reinforcement learning technique that learns an action-value function that gives the expected utility of taking a given action in a given state and following a fixed policy thereafter [15]. Since reinforcement learning algorithms are goal oriented by nature, it is deemed as a suitable fit for this problem.

Considering our Markov chain model with finite state spaces represented by S, a finite set of possible actions $U(i)$ where $i \in S$ and transition probabilities represented by $t_{i,j}$ such that $\sum_{i,j} t_{i,j} = 1$ for all $j \in S$. It is assumed that before using the Markov chain for mobility prediction, the user would gather some data based on its movement. As such, in this context we have trained a Q-learning model according to the scenario from Fig. 1, where a user could start in any state and its goal was to reach the outside region (state 6). Based on that, Q-learning is able to update its function according to

$$\begin{aligned} Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \\ \phi \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] , \end{aligned} \tag{6}$$

where $Q(s_t, a_t)$ is the current action-value function, α, is the learning rate, r_{t+1} is the expected reward at the next time step, ϕ is the discount factor and $\max_a Q(s_{t+1}, a)$ is an estimate of the optimal future action-value function at the next state over all possible actions.

Based on this model, given any starting state, Q-learning learns the next action of a user in order for it to reach the outside (the goal). Thus, by training this model and counting how many times each state action pair were visited, a transition matrix can be built, given by $t_{i,j} = N_i / \sum_j N_j$.

IV. Simulation Results

The simulation environment contains six Markov states, the states from one to five belongs in the indoor environment and state six represents the outdoor environment, as illustrated in Fig. 1. The transition between the states are evaluated to examine the accuracy of the initial values of transition matrix. The system checks the probabilities for a sequence of 100 days with 6 transitions each day. Four different mobility scenarios are applied for each day for a single user: 0% of random data, where users follow predefined routes every day; 10%, 20% and 50% random data, in which random routes are followed with the the given percentages and evenly distributed across the 100 day period. The proposed solution with Q-learning initialization is compared to the solution in [7] in terms of prediction accuracy, identified as the ratio between the number of correct and total number of predictions. For the Q-learning, a learning rate of $\alpha = 1$, a discount factor of $\phi = 0.8$ and an ϵ-greedy policy with ϵ decaying from 1 to 0.3 are assumed. A total of 500 episodes are simulated, with varying number of iterations (the algorithm would stop when the outside region is reached). In terms of the reward, a reward of 0 is assumed for every step the user would take, except in states 5 (door) and 6 (outside), where a reward of 25 and 100 is given.

Fig.3 illustrates the average prediction accuracy values w.r.t different R_d values for the solution from [7] and our proposed Q-learning method. In the first scenario, shown in the solid line, since the transition matrix is initialized with equiprobable values over all possible states, it can be seen that for low values of R_d, the prediction accuracy is very low, with accuracy ranging from 25% to 40% for the different mobility models when $R_d = 0$. This occurs because the Markov chain model does not assume any prior knowledge of user mobility, which results in an initial learning time, in which the algorithm makes wrong predictions more often than correct ones. However, as R_d increases, we can see that the prediction accuracy reaches values of 98 - 99% when no randomness is considered and declines for higher values of R_d for the other mobility models. This occurs because when R_d is higher the transition matrix

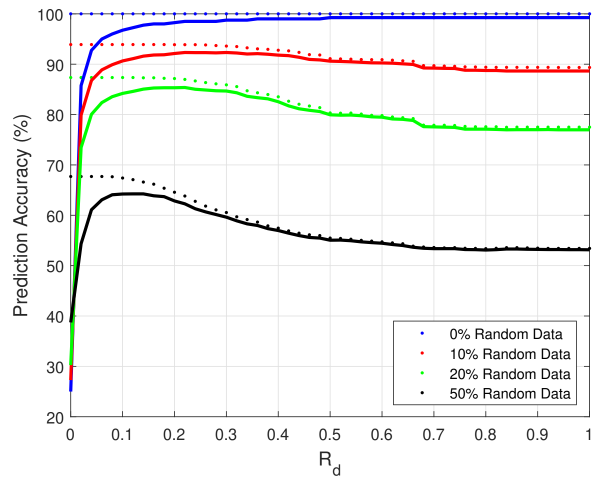


Fig. 3: Accuracy for different values of R_d for both methods.

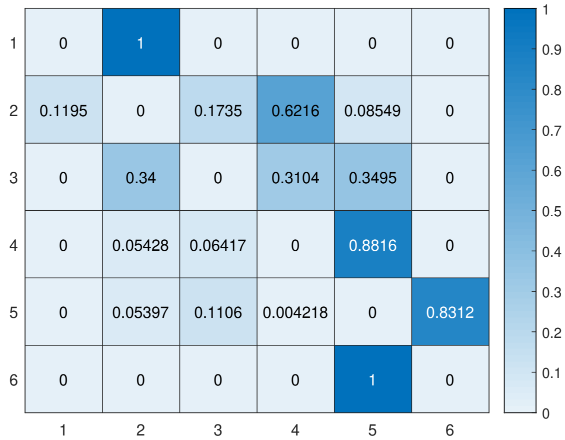
updates faster, thus when randomness is introduced, it is less reliable [7].

For the proposed scenario, shown in the dashed line, it can be seen that the prediction accuracy without any randomness, has the maximum accuracy of 100% as the system already had the initial transition matrix value, which is derived from the Q-Learning method. In addition, this method gives more robust estimate of the accuracy with respect to lower R_d values. However, the increment in the random data, reduces the accuracy comparatively with the higher values in R_d, for the same reasons as mentioned above.

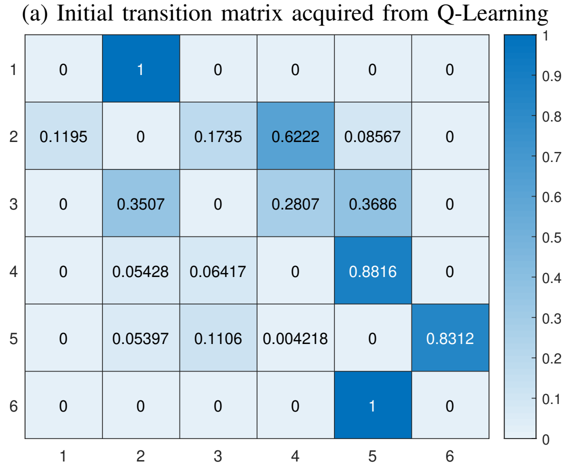
TABLE I: Accuracy gain in percentage.

R_d Values	0	0.2	0.4	0.6	0.8	1
0% Random Data	300	1.78	1.01	0.75	0.75	0.75
10% Random Data	244	1.89	1.06	0.72	0.73	0.78
20% Random Data	188	2.13	1.18	0.40	0.61	0.68
50% Random Data	74	2.74	0.83	0.50	0.51	0.51

Table I shows the gain in terms of accuracy between the proposed solution and the solution in [7], for different values of R_d. It can be seen that, initializing the Markov transition matrix with values from Q-learning yields higher gains when R_d is smaller, with gains over 1% for values of $R_d \geq 0.4$ whereas when R_d is larger, the two solutions converge to each other. This occurs when R_d is larger as the values in transition matrix are updated more quickly, therefore only the most recent paths are considered important than previous one. Therefore, the initialization is not as effective as when smaller values of R_d are considered. Lastly, Fig. 4a demonstrates a heatmap of the path that the user follows according to the Q-learning mobility pattern, which is then used to initialize the



(a) Initial transition matrix acquired from Q-Learning



(b) Transition matrix after learning, when $R_d = 0.2$.

Fig. 4: Markov Chain Transition Matrix.

transition matrix in the Markov chain algorithm. On the other hand, Fig. 4b shows a heatmap of what the proposed solution has learned, for a value of $R_d = 0.2$ and a 50% randomness in user mobility.

V. Conclusions

In this paper, we proposed a user-based indoor mobility predictions via Markov chain with an initial transition matrix, acquired from Q-learning. Results show that, the model with using an initial transition matrix has slightly higher accuracy however, this model would come at the price of more complexity as system needs to be trained based on some data. Therefore, we propose the online learning method for the transition matrix when there is no data available about the user's movement. Based on this acquired knowledge of the user's movement pattern in the indoor environment, among other functionalities, preemptive handovers for mmWave communications can be applied to reduce handover latency in the next generation densely deployed small cell networks.

Acknowledgement

This work was supported by EPSRC Global Challenges Research Fund the DARE Project under Grant EP/P028764/1. The first author was supported by the Republic of Turkey Ministry of National Education (MoNE-1416/YLSY).

References

[1] BBC News, "China rolls out 'one of the world's largest' 5G networks," https://www.bbc.co.uk/news/business-50258287, 2019, Accessed: 2020-06-24.

[2] N. Al-Falahy and O. Y. Alani, "Millimetre wave frequency band as a candidate spectrum for 5G network architecture: A survey," *Physical Communication*, vol. 32, pp. 120–144, feb 2019.

[3] J. G. Andrews, "Seven ways that hetnets are a cellular paradigm shift," *IEEE Communications Magazine*, vol. 51, no. 3, pp. 136–144, 2013.

[4] N. A. Amirrudin, S. H. S. Ariffin, N. N. N. A. Malik, and N. E. Ghazali, "User's mobility history-based mobility prediction in LTE femtocells network," in *2013 IEEE International RF and Microwave Conference (RFM)*, 2013, pp. 105–110.

[5] ——, "User's mobility history-based mobility prediction in LTE femtocells network," in *2013 IEEE International RF and Microwave Conference (RFM)*, 2013, pp. 105–110.

[6] M. Ozturk, P. V. Klaine, and M. A. Imran, "3D transition matrix solution for a path dependency problem of markov chains-based prediction in cellular networks," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017, pp. 1–5.

[7] A. Mohamed, O. Onireti, S. A. Hoseinitabatabaei, M. Imran, A. Imran, and R. Tafazolli, "Mobility prediction for handover management in cellular networks with control/data separation," in *2015 IEEE International Conference on Communications (ICC)*, 2015, pp. 3939–3944.

[8] E. Perez, K.-J. Friederichs, A. Lobinger, B. Wegmann, and I. Viering, "Erratum to: Utilization of licensed shared access resources in indoor small cells scenarios," in *Cognitive Radio Oriented Wireless Networks*, D. Noguet, K. Moessner, and J. Palicot, Eds. Cham: Springer International Publishing, 2016, pp. E1–E1.

[9] M. A. Imran, A. Turkmen, M. Ozturk, J. Nadas, and Q. H. Abbasi, "Seamless indoor/outdoor coverage in 5G," pp. 1–23, may 2020.

[10] P. S. Prasad and P. Agrawal, "Mobility prediction for wireless network resource management," in *2009 41st Southeastern Symposium on System Theory*, 2009, pp. 98–102.

[11] D. M. Rose, T. Jansen, S. Hahn, and T. Kürner, "Impact of realistic indoor mobility modelling in the context of propagation modelling on the user and network experience," in *2013 7th European Conference on Antennas and Propagation (EuCAP)*, 2013, pp. 3979–3983.

[12] B. Tang, C. Jiang, H. He, and Y. Guo, "Probabilistic human mobility model in indoor environment," in *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016, pp. 1601–1608.

[13] D. Helbing and P. Molnar, "Social force model for pedestrian dynamics," *Physical review E*, vol. 51, no. 5, p. 4282, 1995.

[14] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568–2599, 2019.

[15] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.