Angle of Arrival Estimation in Dynamic Indoor THz Channels with Bayesian Filter and Reinforcement Learning

Bile Peng, Qi Jiao, Thomas Kürner

Technische Universität Braunschweig
Schleinitzstraße 22, 38106 Braunschweig, Germany

Abstract – This paper presents a novel algorithm to estimate the Angle of Arrival (AoA) in a dynamic indoor Terahertz channel. In a realistic application, the user equipment is often moved by the user during the data transmission and the AoA must be estimated periodically, such that the adaptive directional antenna can be adjusted to realize a high antenna gain. The Bayesian filter is applied to exploit continuity and smoothness of the channel dynamics for the AoA estimation. Reinforcement learning is introduced to adapt the prior transition probabilities between system states, in order to fit the variation of application scenarios and personal habits. The algorithm is validated using the ray launching channel simulator and realistic human movement models.

Index Terms – Terahertz communication, angle of arrival estimation, dynamic channel, Bayesian filter, reinforcement learning

I. Introduction

Terahertz (THz) communication utilizes the frequency spectrum around 300 - 350 GHz and is expected to realize a data rate of several tens of Gbit/s [1]. The semiconductor technology advances and channel characterization works in recent years have made THz communication a promising solution for future high speed short range communications [1–4], [5] (pp. 495-526).

One of the greatest challenges of THz communications is the extremely high path loss due to the high carrier frequency. In order to realize a reasonable received signal strength, a high gain antenna can be used to compensate for the high path loss. Since the high gain antenna is also highly directional, it is crucial for the receiver to estimate the Angle of Arrival (AoA) of the incoming signal in order to align the main lobe direction for a high antenna gain. The AoA estimation is especially difficult in a dynamic scenario, i.e. the user equipment is moving during the data transmission. Therefore, the AoA is time-variant and needs to be estimated periodically.

For a stationary scenario or a single snapshot of the dynamic case, the AoA can be estimated with methods presented in [6–8]. In the dynamic scenario, the Bayesian filter can be applied to improve the estimation accuracy [9–14]. In our previous works, [15] presents a Bayesian inference based three dimensional AoA estimation algorithm for indoor THz communications. Since the spherical coordinate system is not linear, the system state cannot be represented as parameters and their rate of change (this point will be explained later in detail). Instead, we use the finite memory of parameters to describe the AoA evolvement. In this algorithm, the prior transition probabilities between system states are crucial to the performance. In this paper, we further discuss the possibility to obtain transition probabilities by means of reinforcement learning.

The remaining part of this paper is organized as follows: Section II describes the underlying channel and signal models. Section III derives the Bayesian filter for the AoA estimation. Section IV explains how to train the prior transition probabilities using the reinforced learning. The simulation results are presented in Section V. Section VI concludes the whole paper.

II. Channel and Signal Models

The THz channel is characterized by its high path loss and specular spatial distribution (i.e. the propagation paths are concentrated in several certain directions) [16]. While many papers focus on the azimuth estimation alone and leave the antenna pattern in the vertical direction omnidirectional, the three dimensional AoA estimation can exploit the antenna directivity in both azimuth and elevation directions and a higher antenna gain can be expected. In this paper, we use path loss, elevation and azimuth in the spherical coordinate system to characterize one propagation path. The switched-beam system [17] is assumed for a realizable hardware complexity, which predefines discrete antenna main lobe directions rather than continuous main lobe adjustment. If we assume that the angle between 2 adjacent main lobe directions is $6°$, the main lobe directions are distributed as Fig. 1 depicts, where elevation $\theta = 0°$ is vertically upward whereas $\theta = 90°$ is the horizontal direction.

It is to note that the spherical coordinate system is not linear. For example, an azimuth difference of $30°$ results in the biggest Euclidean distance if the elevation is horizontal ($\theta = 90°$) while the Euclidean difference is 0 in the vertical directions ($\theta = 0°$ or $\theta = 180°$). Therefore, the predefined main lobe directions seem "denser" around the horizontal

978-0-9928-6265-7/16/$31.00

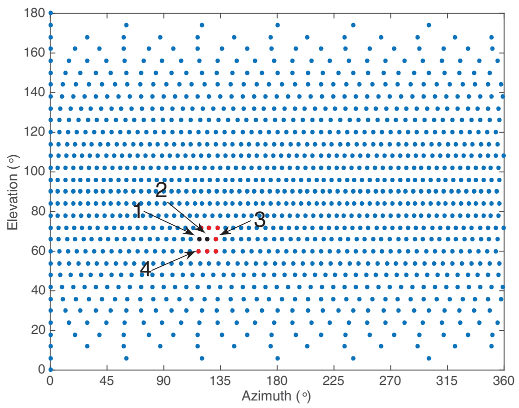


Fig. 1: Predefined antenna main lobe directions in the spherical coordinate system

direction and "sparser" in the vertical directions although they are actually uniformly distributed.

While many Bayesian filters apply the rate of change to the Bayesian inference, it is impossible for the three dimensional AoA estimation because of the nonlinearity of the spherical coordinate system. In this paper, we use the finite memory instead of the rate of change for the Bayesian inference. For example, if we assume the optimal main lobe direction in time instant t is 1 and in time instant $t+1$ is 2 in Fig. 1, the optimal main lobe direction in time instant $t+2$ must be an adjacent direction of 2 (i.e. the black and red dots in Fig. 1) under the condition that the time between 2 time instants is short enough, because the human movement speed is limited. Furthermore, direction 3 should have a higher *prior* probability than direction 4 because the movement $1 \rightarrow 2 \rightarrow 3$ seems more natural than the movement $1 \rightarrow 2 \rightarrow 4$. We define a *state* as the memory of optimal main lobe directions of length 2. In every time instant, a new optimal main lobe direction updates the state. For example, direction 3 updates state (1,2) to (2,3) and direction 4 updates state (1,2) to (2,4) (left is early). The *transition probability* is defined as the prior probability between 2 states, e.g. $p[(2,3)|(1,2)]$ is the probability of state (2,3) given the previous state (1,2). In the next sections, we will describe how to obtain the posterior probability of states and directions using the Bayesian filter and how to obtain the prior transition probability.

III. The Bayesian Filter

The Bayesian filter uses both previous estimates and current measurement for the current estimation. For inference from the previous estimate, if the optimal state at time instant $i-1$ is s_{i-1}, the probability that the optimal state at time instant i is s_i, is the transition probability $p(s_i|s_{i-1})$.

For inference from the current measurement at time instant i, we let the antenna main lobe direction scan over all predefined direction in Fig. 1 and record the received signal power z_i. For a proposed THz communication system, the SNR is expected significantly higher than 0 dB [18]. Therefore, if a high received power is detected in a certain direction, it is safe to assume that it is due to the signal power rather than thermal noise and the probability that this main lobe direction is an optimal direction is proportional to its received power. If this direction is the current direction of a state s_i, we also assume that the probability of the state is proportional to the received power. The probability $p(s_i|z_i)$ is denoted as the *likelihood*.

We denote the received powers from time instant 1 to i as a vector $\mathbf{z}_{1:i}$, where the subscripts indicate the first and last time instants, and define the *posterior* probability $\alpha_i = p(\mathbf{z}_{1:i}, s_i)$ as the probability of a certain state at time instant i and the received powers from the first to the current time instant. According to the forward algorithm, we have

$$\begin{aligned} \alpha_i =& p(\mathbf{z}_{1:i}, s_i) \\ =& p(\mathbf{z}_{1:i-1}, z_i, s_i) \\ =& \sum_{s_{i-1}} p(\mathbf{z}_{1:i-1}, z_i, s_{i-1}, s_i) \\ =& \sum_{s_{i-1}} p(s_i, z_i|\mathbf{z}_{1:i-1}, s_{i-1}) p(\mathbf{z}_{1:i-1}, s_{i-1}) \\ =& \sum_{s_{i-1}} p(s_i, z_i|s_{i-1}) p(\mathbf{z}_{1:i-1}, s_{i-1}) \\ =& \sum_{s_{i-1}} p(z_i|s_i) p(s_i|s_{i-1}) \alpha_{i-1} \end{aligned} \tag{1}$$

where α_i is the posterior probability at time instant i, $p(z_i|s_i)$ is the likelihood, $p(s_i|s_{i-1})$ is the transition probability and α_{i-1} is the posterior probability at time instant $i-1$.

The inference can be illustrated in Fig. 2. If both likelihood and transition probability are high, the state transition is likely to take place. The total posterior probability of a state is the sum of the probabilities from all possible previous states.

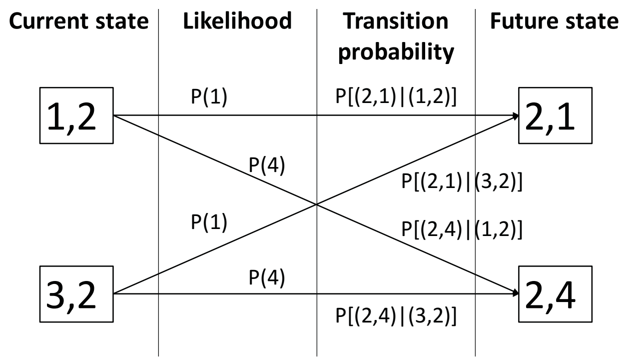


Fig. 2: Illustration of state transitions

In this way, we can iteratively estimate the AoA along time. By the initialization, the posterior probability α_1 can be set uniformly distributed for all states.

Another important point is, the estimation interval between 2 time instants must be small enough to restrict the AoA change within the resolution of the predefined main lobe directions. If we assume the access point is stationary, which is usually this case, the AoA change is the result of both displacement and rotation of the user equipment. The estimation interval must be determined this way, that even if the user equipment is displacing and rotating at its maximum speed and the effects of displacement and rotation are constructively overlapped, the AoA change can still be restricted within the resolution of the predefined main lobe directions. According to our measurement, we set the estimation interval to 0.04 s.

IV. Training of Transition Probabilities by Reinforcement Learning

Section III describes the algorithm to estimate the current AoA from both current measurement and previous estimates, leaving the important problem of obtaining the transition probability unsolved. We can either carry out a training before the application, or update the transition probabilities according to the feedbacks during the application. Reinforcement learning [19] is a promising solution for this concept. We assume that only an accurate AoA estimate can provide a reasonable SNR and therefore this is the necessary condition for the successful decoding of the transmitted data (note that it is not a sufficient condition because the successful decoding also depends on other factors, e.g. AoD estimate of the transmitter, phase noise of the local oscillator and IQ-imbalance). Therefore, the error-detection test (e.g. using the cyclic redundancy check) after the forward error correction can provide the information whether the AoA estimation is correct or unknown. In this paper, we assume a memory length of 2. If the AoA estimates in 3 continuous time instants are correct, a successful state transition has taken place, the state transition probability between these 2 states shall be increased as a reward. The updated transition probability shall be used for the future estimation, until the transition probabilities are optimized for the current application scenario. Fig. 3 demonstrates this idea.

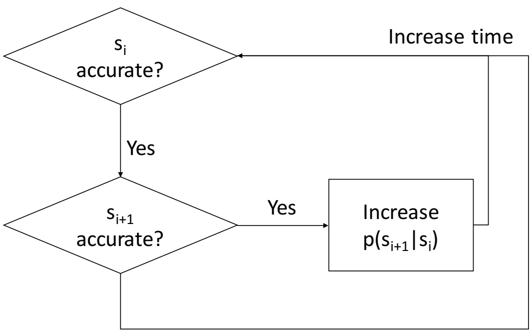


Fig. 3: Reinforcement learning flow chart

The transition probability $p(s_{i+1}|s_i)$ is calculated as

$$p(s_{i+1}|s_i) = \frac{n(s_{i+1}|s_i)}{n(s_i)} \tag{2}$$

where $n(s_{i+1}|s_i)$ is the number of successful estimates that the previous state is s_i and the current state is s_{i+1} and $n(s_i)$ is the number of successful estimates that the previous state is s_i. The former will increase by one after 3 successive precise estimates whereas the latter will increase by one after 2 successive precise estimates.

V. Simulation Results

In this section, we demonstrate the effect of the reinforcement learning with the Bayesian filter by means of simulations. A ray tracing simulator [20] is applied to generate channel models for the simulation. The scenario is a small office, as shown in Fig. 4. The access point is in the middle of the room and under the roof, as depicted with the red point. The user

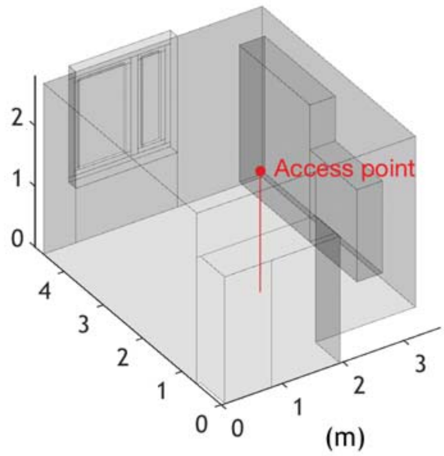
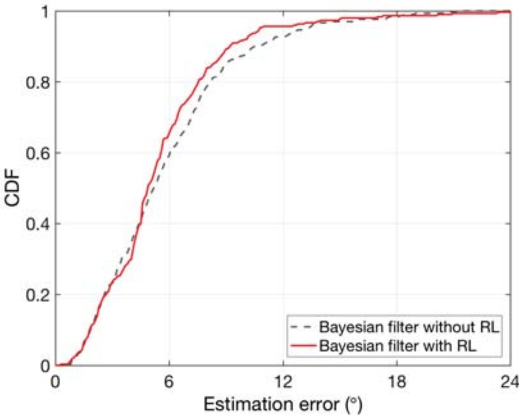


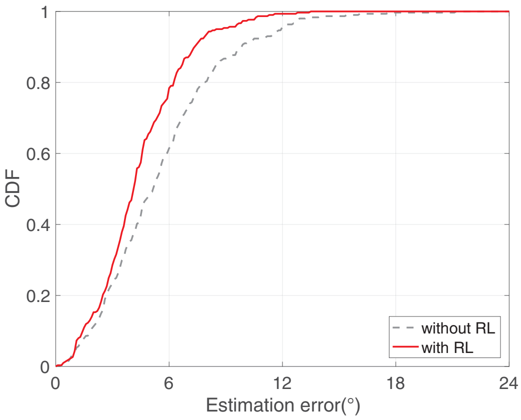
Fig. 4: The office scenario

is moving according to a realistic movement model [15]. The transmission power is assumed as 1 mW. The bandwidth is 50 GHz (300 - 350 GHz). In order to simplify the problem without loss of generality, the antenna gain of the access point is assumed to be constantly 5 dB as a conservative assumption and the antenna of the user equipment is assumed to have a Gaussian antenna diagram [21] with a half power beam width of 12° (such that 2 adjacent main lobe directions realize an antenna gain difference of 3 dB) and a maximum antenna gain of 23.3 dB.

At the initialization, we allocate the same transition probability to all the possible state transitions. The transition probabilities are updated according to the principle in Fig. 3 after every time instant. We first consider the cumulative distribution function (CDF) of the estimate errors. Fig. 5 shows the estimate error CDF after 2000 and 80000 time instants. As expected, the estimate errors with reinforcement learning are smaller than without reinforcement learning and the improvement with 80000 training steps is bigger than with 2000 training steps. Since the resolution of the predefined main lobe directions is 6°, we consider an estimate with an error less than 6° to be correct. We can observe that the probability of a correct estimate with reinforcement learning after 80000 steps is 0.78 while the probability without reinforcement learning is merely 0.61.



(a) After 2000 time instants



(b) After 80000 time instants

Fig. 5: Cumulative distributions of estimation errors

The effective antenna gain is defined as the antenna gain in the direction of the true AoA when the antenna main lobe is pointing at the estimated optimal main lobe. A precise estimate can realize a high SNR and a very low effective antenna gain will be realized and the THz connection will be broken if the estimate error is big, which results in serious performance degradation due to the time required for the reconnection and synchronization. Fig. 6 shows the effective antenna gain with and without reinforcement learning after 80000 time instants. It can be observed that the effective antenna gain with reinforcement learning is not only higher but also more stable than without reinforcement learning, which indicates that reinforcement learning can improve the signal quality and the system stability.

We use the level crossing rate (LCR) to evaluate the stability of the effective antenna gain in a more general sense. The LCR is defined as the number of times that the effective antenna gain is below a certain threshold per second. Fig. 7 depicts the LCR at different thresholds. If we want that the effective antenna gain is higher than 15 dB, we can realize an LCR of 1.75 s^{-1} whereas the LCR is increased to 3.67 s^{-1} without reinforcement learning. These results validate the reinforcement learning as a measure to be adapted to the application scenario in order to improve the system performance and stability.

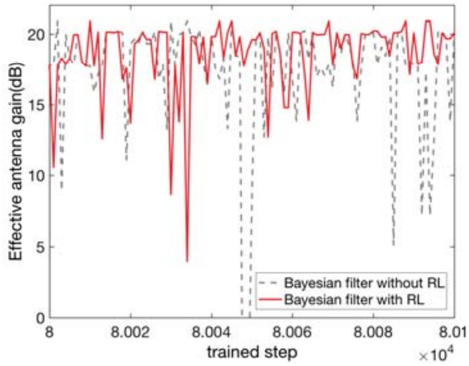


Fig. 6: Realized antenna gain after 80000 time instants

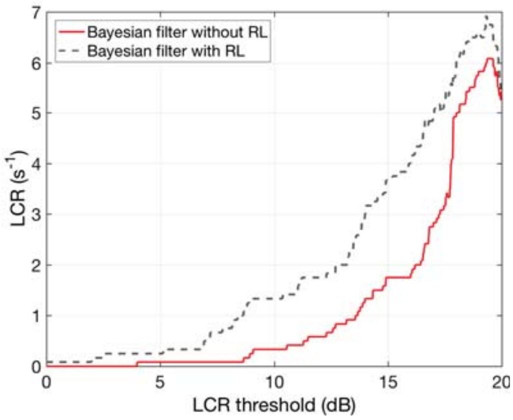


Fig. 7: Level crossing rate after 80000 time instants

VI. Conclusion

This paper introduces a Bayesian filter for the angle of arrival estimation in a dynamic indoor THz wireless channel, where the user equipment is moving during the data transmission and the angle of arrival must be estimated periodically. The Bayesian filter combines the prior probability from the previous estimates and the likelihood from the current measurement to obtain the posterior probability. The transition probability plays a crucial role in the prior probability calculation, which depends on the application scenario and user habit. The reinforcement learning is applied to optimize the state transition probabilities from the feedbacks of the previous estimates and hence improves the algorithm performance. We use a ray tracing simulator and a realistic human movement model to generate a series of channel models for validation of the algorithm. The simulation results show that the reinforcement learning can improve the algorithm performance significantly in both effective antenna gain and its stability. Therefore, a more stable THz link can be ensured during the user equipment movement.

Acknowledgement

The research presented in this paper has been kindly funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung – BMBF) in the framework of the project "TERAPAN" (03V0411) and the EU Framework Programme for Research and Innovation Horizon 2020 (project iBROW, Grant Agreement 645369).

References

[1] T. Kürner, "Towards future THz communications systems," *Int. J. THz Sci. Technol*, vol. 5, no. 1, pp. 11–17, 2012.

[2] T. Kleine-Ostmann and T. Nagatsuma, "A review on Terahertz communications research," *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 32, no. 2, pp. 143–171, 2011.

[3] J. Federici and L. Moeller, "Review of terahertz and subterahertz wireless communications," *Journal of Applied Physics*, vol. 107, no. 11, p. 111101, 2010.

[4] H.-J. Song and T. Nagatsuma, "Present and future of terahertz communications," *Terahertz Science and Technology, IEEE Transactions on*, vol. 1, no. 1, pp. 256–263, 2011.

[5] H. Song, T. Nagatsuma, S. Priebe, and T. Kürner, *Handbook of Terahertz Technologies: Devices and Applications*. Pan Stanford, 2015.

[6] B. Peng, S. Priebe, and T. Kürner, "Fast Beam Searching Concept for Indoor Terahertz Communications," in *Antennas and Propagation (EUCAP), 2014 8th European Conference on*, pp. 483–487, IEEE, 2014.

[7] M. Wax and T. Kailath, "Optimum localization of multiple sources by passive arrays," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 31, no. 5, pp. 1210–1217, 1983.

[8] B. Fleury, D. Dahlhaus, R. Heddergott, and M. Tschudin, "Wideband angle of arrival estimation using the SAGE algorithm," in *Spread Spectrum Techniques and Applications Proceedings, 1996, IEEE 4th International Symposium on*, vol. 1, pp. 79–85, IEEE, 1996.

[9] J. Salmi, A. Richter, and V. Koivunen, "Detection and tracking of mimo propagation path parameters using state-space approach," *Signal Processing, IEEE Transactions on*, vol. 57, no. 4, pp. 1538–1550, 2009.

[10] P. Closas, C. Fernandez-Prades, and J. A. Fernandez-Rubio, "A Bayesian approach to multipath mitigation in GNSS receivers," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 3, no. 4, pp. 695–706, 2009.

[11] X. Yin, G. Steinbock, G. E. Kirkelund, T. Pedersen, P. Blattnig, A. Jaquier, and B. H. Fleury, "Tracking of time-variant radio propagation paths using particle filtering," in *Communications, 2008. ICC'08. IEEE International Conference on*, pp. 920–924, IEEE, 2008.

[12] T. Jost, W. Wang, U.-C. Fiebig, and F. Perez-Fontan, "Detection and tracking of mobile propagation channel paths," *Antennas and Propagation, IEEE Transactions on*, vol. 60, no. 10, pp. 4875–4883, 2012.

[13] B. Krach, P. Robertson, and R. Weigel, "An efficient two-fold marginalized Bayesian filter for multipath estimation in satellite navigation receivers," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, p. 82, 2010.

[14] B. Peng, S. Priebe, S. Rey, and T. Kürner, "Angle of Arrival Estimation for Moving User Equipment with Application to Indoor Terahertz Communications Using Grid Based Bayesian Filter," in *Antennas and Propagation (EUCAP), 2015 9th European Conference on*, pp. 483–487, IEEE, 2015.

[15] B. Peng and T. Kürner, "Three Dimensional angle of arrival estimation in dynamic indoor Terahertz channels using forward-backward algorithm," *submitted to Vehicular Technology, IEEE Transactions on*.

[16] G. D. Durgin, *Space-time wireless channels*. Prentice Hall Professional, 2003.

[17] J. P. Maícas, *Recent Developments in Mobile Communications – A Multidisciplinary Approach*. InTech, December 2011.

[18] S. Priebe, S. Rey, and T. Kurner, "From broadband ray tracing propagation modeling to physical layer simulations of THz indoor communication systems," in *Radio and Wireless Symposium (RWS), 2013 IEEE*, pp. 142–144, IEEE, 2013.

[19] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *Journal of artificial intelligence research*, pp. 237–285, 1996.

[20] S. Priebe, M. Kannicht, M. Jacob, and T. Kurner, "Ultra broadband indoor channel measurements and calibrated ray tracing propagation modeling at thz frequencies," *Communications and Networks, Journal of*, vol. 15, no. 6, pp. 547–558, 2013.

[21] S. Priebe, M. Jacob, and T. Kürner, "Affection of THz indoor communication links by antenna misalignment," in *6th European Conference on Antennas and Propagation (EUCAP)*, pp. 483–487, IEEE, 2012.