Reinforcement Learning-based Receiver for Molecular Communication with Mobility

Lisa Y. Debus, Pit Hofmann, Jorge Torres Gómez, Frank H.P. Fitzek, and Falko Dressler

Abstract—Research in molecular communication (MC) is moving forward in big steps, enabling next-generation communication between nanosensors and presenting an alternative communication model for applications in life sciences and other industrial applications. While a lot of the current research in investigates the setup and en-/decoding process in these testbeds, few tackle the problem of inherently mobile structures with ever-changing channel characteristics and achieving symbol synchronization under these circumstances. In this paper, we employ reinforcement learning (RL) to present an approach to this problem. Using data from a real-world macroscale testbed, we train an RL agent to detect synchronization sequences via threshold adaption in a mobile setting. We comparatively evaluate our approach with the state of the art and report the RL agents ability to adapt to changing channel behavior produced by mobility, achieving a low probability of missed detection and small misalignment with the symbol time.

I. INTRODUCTION

Communication among nanosensors based on molecules still feels like fiction, but today's lab experiments already show that science is overcoming many of the hurdles. In particular, macroscale testbeds demonstrate the capabilities of molecular communication (MC) using information carriers [1], [2]. To decode the sent messages, synchronization mechanisms are critical to realizing digital modulation schemes. Symbol timing, frequency, and sampling clock offset result in the main synchronization errors preventing the decoding of binary information. In the case of MC, synchronizing the receiver and the emitter becomes particularly challenging due to random propagation delays and the inherent distortion produced by diffusion processes in MC channels [3]. Furthermore, nanonetworks are inherently mobile, as MC scenarios assume a fluid medium. Due to such mobility, synchronization mechanisms must operate under time-varying conditions, where the MC channel is hard to estimate, or such estimation is just impractical.

Leveraging on the extended use of machine learning (ML) models, in this work, we report implementing a synchronizer using reinforcement learning (RL). With the appropriate data, RL is able to train an agent especially adept at reacting to changing surroundings, which makes it perfect for use in mobile scenarios [5]. In this work, we target mobile scenarios where the transmitter moves according to Brownian motion while the receiver is static. As a baseline, we use a macroscale MC

Lisa Y. Debus, Jorge Torres Gómez and Falko Dressler are with the School for Electrical Engineering and Computer Science, TU Berlin, Berlin, Germany, email: debus@campus.tu-berlin.de, {torres-gomez, dressler}@ccs-labs.org.

Pit Hofmann and Frank H.P. Fitzek are with the Deutsche Telekom Chair of Communication Networks, Technische Universität Dresden, Dresden, Germany; F. Fitzek is also with the Centre for Tactile Internet with Human-in-the-Loop (CeTI), Dresden, Germany, email: {pit.hofmann, frank.fitzek}@tu-dresden.de.
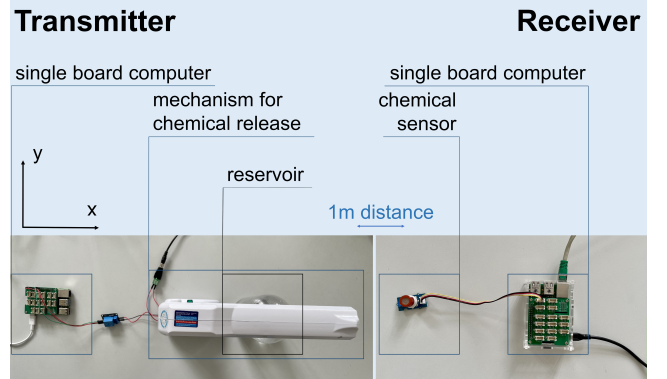
Fig. 1: Molecular SISO communication testbed, see also [4].

testbed as depicted in Fig. 1. We reflect its mobility by creating a channel impulse response (CIR) modeling the receiver's time-varying positions.

We train an RL model, which is deployed at the receiver, to set the proper threshold level and detect the incoming frame for synchronization. Our results clearly show a high probability for correct frame detection and low symbol-time offset.

II. SYSTEM MODEL

In this work,[1] we target an MC scenario with an emitter moving according to Brownian motion and a receiver located at a fixed position. We make use of the macroscale single-input single-output (SISO) MC testbed as introduced in our previous work [4]. The transmitter consists of an electronically controlled sprayer for the mechanical release of molecules and a reservoir of diluted Ethanol. We employ on-off keying (OOK) modulation for data transmission. Transmission of a `1` activates the sprayer to release molecules (a total number of 3.9×10^{21} molecules) and a `0` accounts for no release of molecules. The molecules' propagation from the sprayer to the sensor is subject to additional drift. The receiver contains an electrochemical sensor measuring the concentration level of molecules with a sampling frequency of $5\,\mathrm{Hz}$.

We simulate one-dimensional mobility in our system by moving the emitter according to Brownian motion. The movement changes the emitter's distance to the receiver and, thus, the CIR of the MC channel over time. To evaluate the spatio-temporal CIR, we collect testbed measurements for fixed distances of $0.5\,\mathrm{m}$, $0.75\,\mathrm{m}$, $1.0\,\mathrm{m}$, $1.25\,\mathrm{m}$ and $1.5\,\mathrm{m}$ and then interpolate the measured time series to approximate the CIR
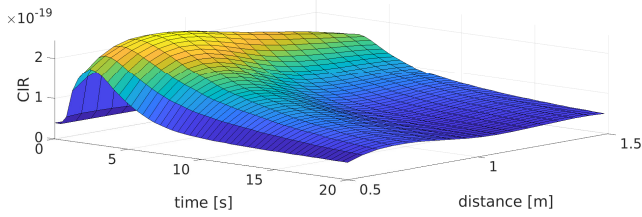
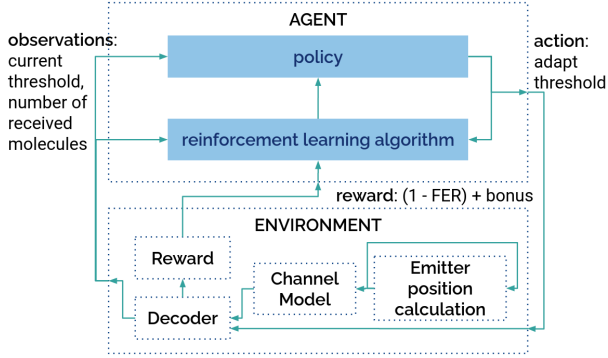Fig. 2: Average CIR of the testbed created by interpolating the recorded measurements.

Fig. 3: Training loop used for the training of threshold setting with a mobile emitter in the presented testbed.
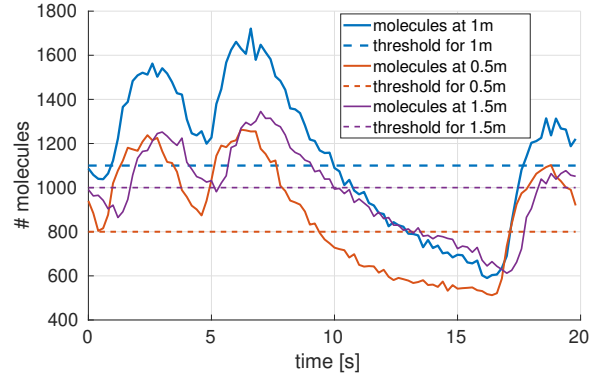

Fig. 4: Total number of received molecules in a synchronization frame and the threshold set by the PPO agent during its transmission for the distances of $1\,\text{m}$, $0.5\,\text{m}$ and $1.5\,\text{m}$.

for any distance between the recorded steps. Fig. 2 shows the resulting surface of the CIR.

Each time the emitter changes its position, we evaluate the distance to the receiver and sample the spatio-temporal CIR surface to collect the expected number of molecules over time on the receiver side. Using the sampled CIR, a transmission is performed, and the transmitter moves to the next position after all molecules were emitted.

III. RL-BASED SYNCHRONIZER

For our RL-based synchronizer, we implemented the training loop shown in Fig. 3. In each loop, a new emitter position is calculated based on its last position and the transmission of one frame is performed. The distance between the new emitter position and the static receiver is provided to the channel model to look up the CIR. Based on this time series, the number of received molecules per symbol is calculated and forwarded to the decoder. The decoder uses the threshold set by the RL agent to decode the number of molecules into a bit sequence.

A reward judging the correctness of the decoded sequence is passed on to the agent. At the same time, the number of received molecules and the currently used threshold are forwarded to it as the observation for the current environment state. Using the reward and the observations, the RL agent decides how to adapt the threshold for the next run of the training loop and forwards the change of the threshold via its action to the environment. By repeatedly interacting with the environment in this fashion, it learns to set a fitting threshold for the correct decoding of the synchronization frame.

We implemented the described training loop as an RL environment in Simulink and Matlab. A proximal policy optimization (PPO) agent with recurrent neural networks (RNN) learns the setting of the threshold with a mobile emitter. Based

on the reward, the agent evaluates the current threshold and decides on the necessary changes based on the observation.

In our experiments, we pass two observations per loop to the agent. The current threshold and the number of molecules received during the transmission are used to describe the state of the system. To interact with the environment, the PPO agent produces actions which raise or lower the current threshold. The resulting threshold is then passed as an observation to the agent in the next simulation loop. The reward judging the performance of the current threshold is calculated by comparing the decoded and the source synchronization frame. Using the success of the threshold calculated by dividing the number of correctly decoded bits by the frame length and an additional bonus for a completely correctly decoded frame, the reward is calculated and forwarded to the agent together with the mentioned observations. Based on those values, the agent decides the threshold for the MC link for the next loop.

In our setup, we use the synchronization frame `[11001]` (as follows from [7]) with a bit time of $4\,\text{s}$ aiming to reduce the impact of inter-symbol interference (ISI) and a sampling rate equal to the frame rate. The start distance of the emitter in our setup is set to $1\,\text{m}$. To make it possible for the agent to learn a fitting threshold, the movement of the emitter has to be appropriately slow so that the agent is able to react to the changed position. Based on our experiments, we were able to determine that a diffusion coefficient of $D_{\text{Tx}} = 8.4 \times 10^{-5}\,\text{m}^2/\text{s}$ or lower will give the agent enough time to react to the changed channel behavior.

The resulting RL agent was trained for 450 episodes in which a total number of 9×10^4 transmissions of the 5-bit synchronization frame was performed at one-dimensional changing distances between the emitter and the receiver. Fig. 4 shows the number of received molecules for transmitted synchronization frames and the appropriate threshold set by the RL agent for three example distances of $1\,\text{m}$, $0.5\,\text{m}$ and $1.5\,\text{m}$. The number of molecules reaching the receiver side of the system varies significantly, and the RL-based receiver adapts the threshold successfully.

IV. EVALUATION

In the following, we analyze the performance of the RL method for threshold settings. The simulation parameters
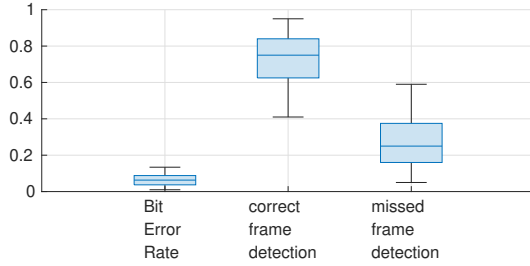
Fig. 5: Bit error rate (BER) for all transmitted bits and probability of correct or missed detection of synchronization frame achieved by the RL agent in 10^4 transmissions.

Fig. 6: Histogram of the misalignment of 700 transmissions for the RL-based method and the filter-based ML scheme.

used for the evaluation are set according to Section III. We comparatively evaluate the RL agent's performance with the filter-based ML synchronizer as an ideal receiver (see [7, Eqs. (7) and (8)]).

A. Probability of Missed and Correct Detection of the Synchronization Frame

As shown in Fig. 5, well over 90% of the received bits in the frames are decoded correctly. The probability of correctly decoding the sent synchronization frame, on the contrary is 72.16%. The implemented RL scheme is still subject to improvements as upon evaluating the synchronization capabilities of the filter-based ML, we find that with this method all synchronization frames are detected correctly. This indicates that the system is performing in a high signal to noise ratio (SNR) environment, which may account for the lower frame error rate (FER). While the performance of the agent must still be improved for it to be usable in real-world systems, the potential in using RL for threshold adaption in dynamically changing systems is apparent.

In our results, we see ISI as a major challenge for correctly setting the threshold for a whole synchronization frame. We found, that especially the decoding of the 0 in the third position of the used synchronization frame 11001 presents a problem for the threshold setting as it is subject to high ISI from the first two 1s in the synchronization frame. This problem is amplified in the case of a bigger distance between the emitter and receiver. As the CIR curve flattens with increasing distance, the ISI increases and leads to even higher numbers of molecules received during the first 0 of a synchronization frame.

B. Misalignment

For the misalignment of the synchronization frame, we measure the time offset when the first bit of the sent frame is detected correctly. A histogram showing the misalignment using the proposed RL-based approach compared to the filter-based ML scheme can be seen in Fig. 6. Compared to the results for the filter-based ML scheme, our approach achieves a lower misalignment.

The histogram shows that a majority of the transmissions have a misalignment of $0.25\,$s and the start of the frame is detected almost immediately after its transmission starts. A second spike for the misalignment is located at $1.25\,$s where we observe a steep incline in the number of received molecules at
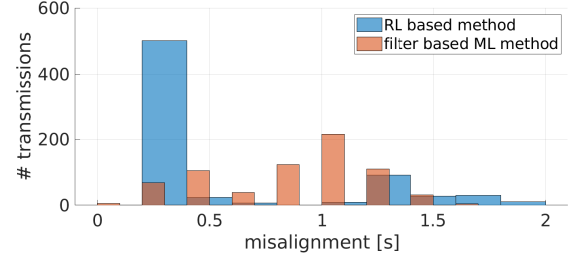
this point. Due to the increased number of molecules reaching the receiver at this point in the transmission, a threshold set to a higher value can be crossed at this point.

V. CONCLUSION

This research manifests the applicability of RL in overcoming the impact of mobility in MC scenarios. After training, our approach exhibits a high detection rate and a low misalignment for the decoded frames. In our future work, we intend to improve upon the agent's state observation system and add more robust system models to make RL a possible option in real-life deployments.

REFERENCES

[1] M. Bartunik, G. Fischer, and J. Kirchner, "The Development of a Biocompatible Testbed for Molecular Communication with Magnetic Nanoparticles," *IEEE Transactions on Molecular, Biological and Multi-Scale Communications*, pp. 1–1, 2023.
[2] L. Brand, M. Scherer, S. Lotter, S. Lotter, A. Burkovski, H. Sticht, K. Castiglione, and R. Schober, "Switchable Signaling Molecules for Media Modulation: Fundamentals, Applications, and Research Directions," arXiv, cs.NI 2302.10356, Feb. 2023. [Online]. Available: https://arxiv.org/pdf/2302.10356.pdf.
[3] V. Jamali, A. Ahmadzadeh, W. Wicke, A. Noel, and R. Schober, "Channel Modeling for Diffusive Molecular Communication - A Tutorial Review," *Proceedings of the IEEE*, vol. 107, no. 7, pp. 1256–1301, Jul. 2019.
[4] P. Hofmann, J. Torres Gómez, F. Dressler, and F. H. P. Fitzek, "Testbed-based Receiver Optimization for SISO Molecular Communication Channels," in *IEEE BalkanCom 2022*, Sarajevo, Bosnia and Herzegovina: IEEE, Aug. 2022, pp. 120–125.
[5] N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y.-C. Liang, and D. I. Kim, "Applications of Deep Reinforcement Learning in Communications and Networking: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3133–3174, 2019.
[6] L. Y. Debus, P. Hofmann, J. Torres Gómez, F. H. P. Fitzek, and F. Dressler, "Reinforcement Learning-based Receiver for Molecular Communication with Mobility," in *IEEE GLOBECOM 2023*, Kuala Lumpur, Malaysia: IEEE, Dec. 2023.
[7] V. Jamali, A. Ahmadzadeh, and R. Schober, "Symbol Synchronization for Diffusion-Based Molecular Communications," *IEEE Transactions on NanoBioscience*, vol. 16, no. 8, pp. 873–887, Dec. 2017.