Proactive Low Level Mobility in Cellular Networks

Anna Prado, Aaron Jakumar, Serkut Ayvasik, Fidan Mehmeti, Wolfgang Kellerer
Chair of Communication Networks, Technical University of Munich, Germany.
{anna.prado, aaron.jakumar, serkut.ayvasik, fidan.mehmeti, wolfgang.kellerer}@tum.de

Abstract*—Mobile users frequently face significant interruptions in transmission and reception during handovers from one Base Station (BS) to another, resulting in latencies that are incompatible with the stringent requirements of Ultra-Reliable Low Latency Communications (URLLC). To address this, 3GPP introduced a novel handover procedure, called Layer 1/Layer 2 Triggered Mobility (LTM), in Release 18. LTM uses lower level signaling to respond quicker to mobility events, bypassing the reconfiguration of higher layers while keeping modifications to the lower layers at a minimal level. This drastically reduces service interruptions during handovers, making them practically negligible. However, since LTM uses more frequent L1 measurements, it has a higher handover and ping-pong handover rates, as well as signaling overhead. In this work, we propose to incorporate future channel predictions in LTM to perform cell preparations and handover decisions with the goal of reducing signaling overhead and resource reservation. We focus on a controlled indoor scenario, where future user channel predictions are possible with a high accuracy. Our proactive algorithm reduces the cell preparation rate by 76% and the handover rate by 72%, without compromising the network sum throughput. Moreover, the resource reservation time at the target BS is reduced to nearly 0 ms.

***Index Terms*—Handover, mobility management, proactive, 5G.**

I. Introduction

5G technology has transformed multiple industries by enabling a wide range of innovative applications, including augmented and extended reality, smart factories, and telemedicine. Looking ahead, 6G is expected to build on this progress and introduce new applications, including holographic communication, immersive extended reality, and advanced healthcare solutions [1]. These applications will require reliable and uninterrupted connectivity. However, the mobility procedures available up to Release 17 fall short of meeting the requirements for Ultra-Reliable Low Latency Communications (URLLC). This shortcoming is due to either a high Handover Interruption Time (HIT) of 80 ms [2] with both the Baseline Handover (BHO) and Conditional Handover (CHO), or because the handover method that avoids HIT, known as Dual Active Protocol Stack (DAPS) [3], is only implemented for Frequency Range (FR) 1, and not for FR 2.

In 5G Advanced Release 18 [3], a novel approach called Layer 1/Layer 2 Triggered Mobility (LTM) was introduced to enable seamless handovers between Base Stations (BSs). LTM leverages Layer 1 (L1)/L2 measurements to trigger handovers, eliminating the need for reconfiguring higher layers such as the Radio Resource Control (RRC) and Packet Data Convergence Protocol (PDCP). Furthermore, LTM aims to minimize the number of reconfigurations at the lower layers, such as Medium Access Control (MAC) and Physical (PHY). The usefulness of using LTM can be best reflected in HIT values as low as 1 ms [4].

With LTM, the cell[1] preparation is performed similarly to the BHO and CHO, using L3 measurements. The user can then preemptively synchronize and acquire timing advance of the prepared cells, even before receiving a handover command. The user periodically measures the channel of all prepared BSs and sends Measurement Reports (MRs) containing L1 measurements to the serving BS, which uses them to decide when to execute a handover. Release 18 [3] specifies the signaling messages required for LTM, while the exact conditions for executing a handover are determined by the operator.

While LTM reduces HIT to a negligible level, it increases signaling traffic and the energy consumption as users receive and send more control messages and have to monitor more prepared BSs. Additionally, LTM can lead to higher handover failure rates [4]. Combined with an elevated rate of ping-pong handovers, this suggests that handovers may sometimes be executed too early or to a wrong BS. The optimal BS selection becomes especially challenging at higher frequencies, where there are significant channel fluctuations. This makes it difficult to know based on the current channel which BS will offer a Line of Sight (LoS) to the user in the next time slots.

Therefore, minimizing the occurrence of handovers and carefully selecting the BS that needs to be prepared in cellular networks is of utmost importance. In an indoor scenario, the channel can be predicted in the next $100-500$ ms [5], [6]. The cell preparation takes less than 30 ms, so even for a smaller prediction window, there is enough time for the serving BS to prepare target BS, and for the user to synchronize and obtain the timing advance of the target BS. This can be accomplished in an indoor scenario, where we have a complete control of the environment, using proactive channel prediction as in [5]. Leveraging these channel predictions enables proactive mobility management, allowing for more informed, proactive, and efficient decisions in handover processes. Furthermore, reducing the cell preparation rate helps to save valuable

This work was supported in part by the Bavarian Ministry of Economic Affairs, Regional Development and Energy under the project "6G Future Lab Bavaria", and in part by the Federal Ministry of Education and Research of Germany (BMBF) under the project "6G-Life", with project identification number 16KISK002.

[1]Note that we use the words cell and BS interchangeably.

radio resources and the limited contention free random access resources [7].

Some important questions related to proactive mobility management in 5G and beyond networks that arise are:

- How can handover-related decisions be performed reliably to support URLLC while simultaneously reducing signaling, energy consumption, and resource reservation at the target BS?
- How does this approach perform compared to the conventional baseline algorithms?

To address these questions, this paper proposes a proactive approach that leverages predicted future channel information to make handover decisions. The results that we present here are important for the operators aiming to support URLLC in dynamic environments. Specifically, our main contributions are:

- We propose a proactive mobility management solution to enhance LTM that significantly reduces the handover execution and cell preparation rates, as well as the resource reservation time.
- We perform extensive realistic 3GPP-compliant simulations in a dynamic FR2 environment clearly showing the advantages our approach offers.

II. Related Work

Several studies have proposed prediction-based algorithms using machine learning to improve mobility management [8], [9], [10]. In [8], the authors analyze in depth mobility-related issues and their measurements show that frequent handovers in 5G diminish throughput and deplete user batteries, causing in the worst case a complete service outage. Moreover, they propose to predict the future signal strength and then to learn the BS's logic for handover execution. The authors consider the BHO and aim at improving the quality of experience of mobile users.

To reduce the cell preparation rate in CHO, the authors in [9] predict each user's future trajectory and prepare the BSs accordingly along the predicted path. The authors in [10] predict the future target BS based on the measured Reference Signal Received Power (RSRP) values. They take advantage of the fact that users experience blockages in certain geographical locations. By learning these blockage patterns from historical data, the algorithm can incorporate this knowledge into the cell preparation decision-making process in CHO. However, their approach only considers geometric blockages and may not account for sudden LoS blockages and shadowing, which cause unpredictable fluctuations in channel conditions. On the other hand, in this work, we focus on an indoor scenario, in which the multipath effect makes channel predictions and handover decisions more challenging.

The works most closely related to ours in the realms of enhancements for LTM and proactive handovers are [4] and [6]. The authors in [4] propose adding L2 filtering on the L1 measurements to reduce signal fluctuations, thereby decreasing the ping-pong handover rate. They also explore the benefits of dynamic switching, in which users keep prepared cell configurations after a handover instead of releasing them. While this approach lowers the cell preparation rate and subsequently reduces signaling, it results in increased resource reservation time, and the overall signaling overhead remains high. In [6], the authors conclude that the data rate degradation in mmWave cannot be predicted solely based on the signal strength. Hence, they propose to include camera images in the state of their Deep RL (DRL) algorithm with the goal of predicting long-term data rates. They consider the BHO and rely on the DRL agent to make handover decisions. Even though they execute handovers proactively before the signal degradation happens and increase the data rate, the HIT cannot be avoided with the BHO, making it unsuitable for URLLC. Moreover, the user is more likely to experience a handover failure with the BHO, especially in FR 2 [11]. Therefore, in this work, we propose enhancing LTM by incorporating channel predictions based from camera images to enable effective cell preparations and handover executions. If the DRL model fails, the system retains the ability to revert to the default LTM, which relies on measured data.

III. Baseline Models

In this section, we describe the two baseline models against which we are going to compare our approach in Section IV.

A. L1/L2 Triggered Mobility (LTM)

The user measures the channel and sends L3 MRs to its serving BS periodically. Before reporting the measurements, the user applies Layer-3 filtering and averages RSRP values over 200 ms [3]. Based on these measurements, the serving BS selects candidate BSs that should be prepared for a potential handover. Then, the user starts sending L1 MRs with the prepared cells, which are used to trigger a handover. LTM uses an A3 event for handover preparation, execution, and cell release [12]. The A3 event is triggered when a neighboring BS becomes better than the serving BS by a preparation or execution margin. In LTM, the reconfiguration of upper layers (e.g., RRC or PDCP) is avoided, as well as the changes to lower layers (e.g., MAC and PHY) are kept to a minimum level. This results in a considerable reduction of HIT. Another enhancement in LTM is Random Access Channel (RACH)-less handover (if the timing advance of the target BS is available at the user side). As a result, HIT can be reduced to as little as 1 ms [4]. LTM can be applied to Centralized Unit (CU)-Distributed Unit (DU) split architecture, with CU handling the PDCP and RRC layers, and the DU managing the PHY, MAC, and Radio Link Control (RLC) layers of the protocol stack.

B. Enhanced LTM

The authors in [4] investigate LTM with two enhancements: L2 filtering and dynamic switching as discussed previously in Section II, denoted as Lower Layer Mobility (LLM) further in this work. Introducing the additional filtering delays handovers, which reduces the ping-pong handover rate without decreasing the reliability. Dynamic switching allows the user

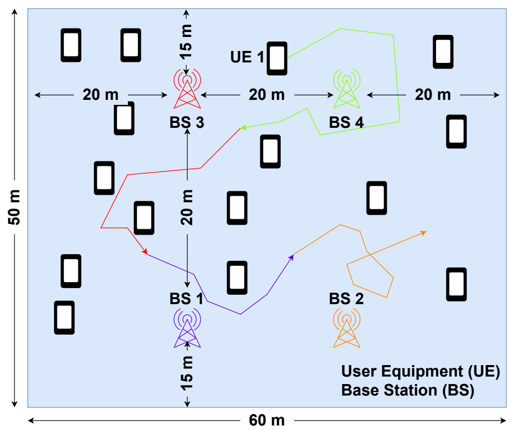


Fig. 1: Illustration of the system model.

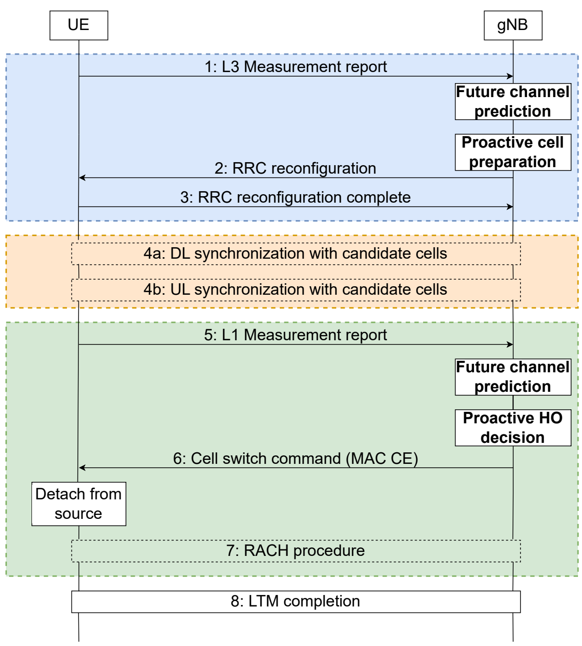


Fig. 2: The signaling in PLTM.

to keep the list of the prepared cells after a handover, instead of releasing them. This reduces signaling since the user is likely to execute a handover to one of the already prepared BSs.

IV. Proposed Proactive Handover Algorithm

In this section, we first present the system model and then explain the proposed Future LTM algorithm.

A. System Model

We consider an indoor factory floor covered with multiple BSs, where robot workers move around performing different tasks. The scenario is illustrated in Fig. 1.

In 5G, Physical Resource Block (PRB) is the unit of resource allocation [3]. One PRB is defined as 12 consecutive subcarriers in the frequency domain and one slot in the time domain. We consider FR 2, and therefore, set the numerology

TABLE I: Simulation Parameters [3], [13]

Parameter	Value
FR2 Carrier frequency	30 GHz
Channel measurement periodicity	10 ms
L3 filtering time constant	200 ms
Ping-pong window	1000 ms
HIT for LTM	1 ms
HIT for CHO	80 ms
Handover preparation time	28.5 ms
Cell preparation margin	2 dB
Handover execution margin	3 dB

to $\mu = 2$, using a subcarrier spacing of 60 kHz. The bandwidth per PRB is 720 kHz. The per-PRB data rate that a user receives from BS is calculated according to 3GPP's specification [14]. It is a function of the Signal to Interference plus Noise Ratio (SINR) of user and the bandwidth of the PRB of BS. The total user rate depends on the number of PRBs allocated to the user by the serving BS. After the handover decisions are made, the resources of a BS are split equally among connected users. Note that the number of PRBs might not be divisible by the number of users connected to a BS at a given time slot, when this is the case, some users might receive one PRB more.

B. Proactive LTM (PLTM)

One or multiple conditions need to be satisfied to start the handover preparation or execution. 3GPP does not define these conditions [3], leaving it to the operator to decide when to start the corresponding procedure. We propose to use the predicted future channel for cell preparation and handover execution.

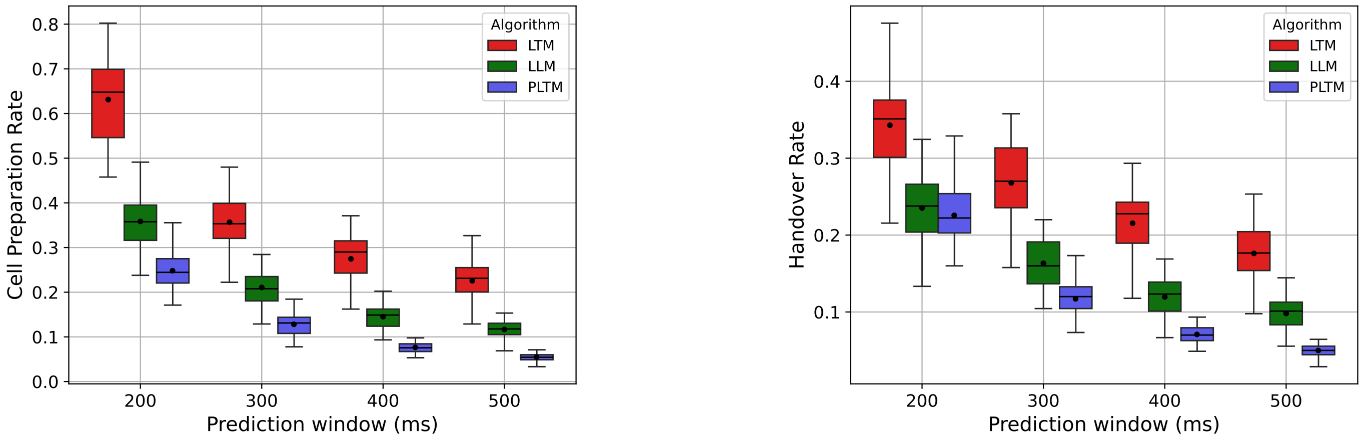
Fig. 2 presents the signaling of our proposed algorithm, coined Proactive LTM (PLTM). The user sends periodically L3 MRs to its serving BS, using which the BS checks if the preparation condition is satisfied. In PLTM, the usual A3 event for the cell preparation decision is used, but we use predicted future channel with the look ahead window length of w. The formula for the preparation event is

$$S_{b'}^{PredL3}(t) - S_b^{PredL3}(t) > O_{prep}, \text{ for } \tau \in [t; t+w], \quad (1)$$

where $S_b^{PredL3}(t)$ and $S_{b'}^{PredL3}(t)$ are predicted L3 SINR values of the serving and target BSs, O_{prep} is the preparation margin, and t is the current time. To summarize, the SINR of the target BS should be higher than that of the serving one by O_{prep} during the prediction window w to trigger a cell preparation. Then, if one or multiple BSs were prepared, the serving BS communicates the configurations of the prepared cells to the user.

Afterwards, the user starts reporting periodically L1 measurements of all prepared BSs to the serving BS (step 5 in Fig. 2), based on which the handover decision is made. We use the actual L1 channel measurement at the current time slot and the predicted channel to make handover decisions. The handover execution condition is similar to the one in (1), and is defined as

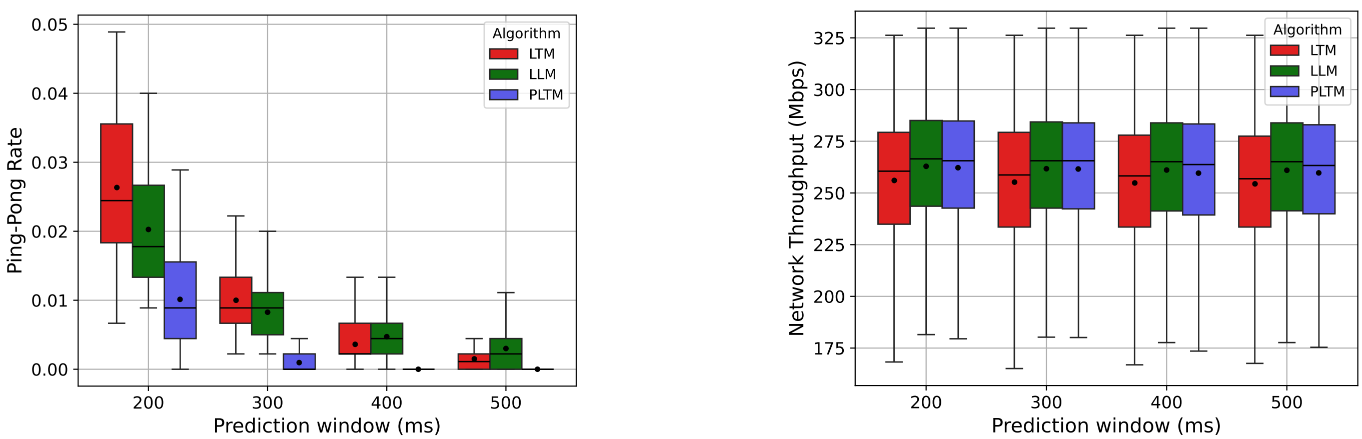
$$S_{b'}^{PredL1}(t) - S_b^{PredL1}(t) > O_{exec}, \text{ for } \tau \in [t; t+TTT_{exec}], \quad (2)$$



(a) Cell preparation rate (per user per second).

(b) Handover rate (per user per second).

Fig. 3: Performance evaluation with LTM, LLM and PLTM for different prediction windows.



(a) Ping-pong handover rate (per user per second).

(b) Network sum throughput.

Fig. 4: Performance evaluation with LTM, LLM and PLTM for different prediction windows.

where $S_b^{PredL1}(t)$ and $S_{b'}^{PredL1}(t)$ are predicted L1 SINR values of the serving and target BSs, whereas TTT_{exec} is Time-to-Trigger (TTT) for handover execution.

To increase the reliability, an extra condition could be added to enable cell preparation and handover execution using the actual MR if there is a neighbouring BS whose SINR is significantly better than the one of the serving BS. A preparation threshold could be set significantly higher than the preparation and execution margins used for the predicted channel.

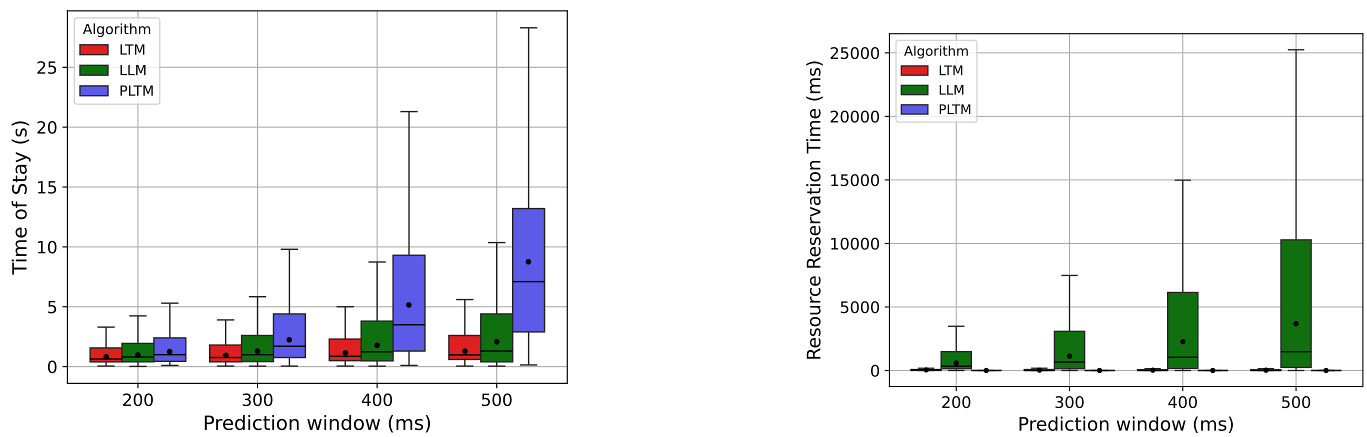
Prepared BSs need to be occasionally replaced or removed from the list, and the operator needs to define conditions for these events as well. The conditions are similar to the ones in (1)-(2). We also use the predicted channel for these events; specifically, a prepared BS is replaced if there is another BS with SINR by O_{repl} higher during TTT_{repl}. Similarly, if SINR of a prepared BS is by O_{rem} lower than that of the serving one at any time slot during TTT_{rem}, this BS is released and removed from the list of prepared BSs.

V. Performance Evaluation

First, we describe the simulation setup and present the results with proactive LTM. Then, we show how another handover algorithm, CHO, can benefit from proactive decision-making.

A. Simulation Setup

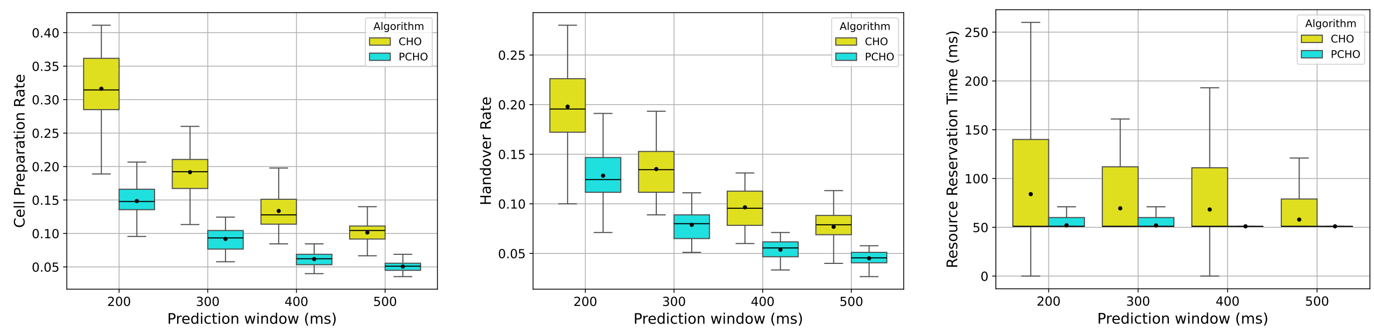
We evaluate the algorithms with a simulated channel from 3GPP 5G Release 18 [13]. We consider an indoor scenario with 4 BSs and 20 users over 1000 s. We model the path loss and shadowing for LoS and no LoS as in [13]. The correlation distance for shadowing is set to 10 m. Random Waypoint mobility model is used to generate the mobility traces [15] for mobile robots moving around the factory floor with the speed up to 3 m/s. The frequency reuse factor is 4, so there is no interference. For a fair comparison, we set the values for both TTT_{prep} and predicted window length w to be the same; specifically to $\{200, 300, 400, 500\}$ ms. Larger TTT_{prep} values reduce the cell preparation rate, and, as a result, the handover rate. The same happens for larger w. We set $TTT_{exec} = 0$ ms to enable fast reactions to channel drops.



(a) The time the user stays connected to the same BS.

(b) Resource reservation time at prepared BSs.

Fig. 5: Performance evaluation with LTM, LLM and PLTM for different prediction windows.



(a) Cell preparation rate.

(b) Handover rate.

(c) Resource reservation time.

Fig. 6: Performance evaluation with CHO and PCHO for different prediction windows.

The other simulation parameters are provided in Table I. To evaluate the proactive algorithms, we assume perfect channel knowledge in the next 500 ms.

B. Performance of Layer 1/Layer 2 Triggered Mobility (LTM)

We first performed simulations in which the predicted channel was used only for cell preparation. However, since the cell preparation is usually anyway performed earlier than the handover execution (by setting a smaller margin), there was no improvement comparing to using the actual channel. Therefore, we decided to utilize future channel knowledge for handover execution as well.

Fig. 3 shows that incorporating future channel knowledge reduces mobility-related signaling, especially for larger prediction windows. Specifically, cell preparation signaling is reduced by 53% and 76% when compared to state-of-the-art LLM and the baseline LTM (see Fig. 3a). With PLTM, users experience 34%, 56%, 67%, 72% and 4%, 28%, 41%, 49% fewer handovers compared to LTM and LLM for $w \in \{200, 300, 400, 500\}$ ms (see Fig. 3b). As anticipated, the ability to predict further into the future corresponds to a lower handover rate.

Fig. 4a shows that the ping-handover rate reduces by $50-99\%$ with PLTM, and for larger prediction windows of 400 and 500 ms, there are no ping-pong handovers at all. Interestingly, PLTM does not reduce the network sum throughput at the cost of reducing the handover-related events. PLTM manages to increase the sum throughput by 3% compared to the corresponding LTM algorithm (see Fig. 4b). As a result of the lower handover rate, the time of stay, which is the time the user stays connected to the same BS increases, as depicted in Fig. 5a.

Another advantage of the PLTM algorithm is that it reserves wireless resources at prepared cells for a shorter duration compared to LTM and LLM (see Fig. 5b). One of the enhancements of LLM is that it does not release the prepared cells after a successful handover, hence, it has the highest resource reservation time among the evaluated algorithms. Since HIT is 1 ms, the average time that the resources are reserved with PLTM is only a few milliseconds, as the cell is typically prepared only when a handover is about to occur. This allows the operator to conserve resources by reducing the amount required for random access and radio resources, especially in dynamic environments with a large number of connected users.

3GPP's specification [3] only states that user admission should be performed, without defining how and how many resources should be allocated to users who might execute a handover. For simplicity, we did not implement the actual radio resource reservation at prepared BSs. As a result, the actual throughput with LTM and LLM might be lower than the one in Fig. 4b because some resources may need to be reserved for incoming users. Moreover, certain users might not be admitted to a BS because it reserved resources for other users who may connect to that BS much later or potentially not at all.

Our proactive approach reduces signaling and resource reservation time while slightly increasing the overall throughput. Typically, there is a trade-off between sum throughput and handover rate [16], [17] because handover decisions are delayed to ensure the proper base station is selected. However, this trade-off is mitigated with the proactive strategy.

By leveraging channel prediction, URLLC can maintain high performance, even in dynamic conditions. Proactive decisions based on future channel states allow for seamless handovers, optimal resource allocation and dynamic adaptation to changing wireless channel, which are crucial for the success of URLLC in real-world scenarios such as autonomous driving, industrial Internet of Things (IoT), and smart healthcare.

C. Performance of Conditional Handover (CHO)

Another handover procedure, which was proposed in 5G to improve mobility robustness, is called CHO [3]. CHO decouples BS preparation and handover execution phases, and its performance was evaluated in [18], [19]. The main difference to LTM is that CHO uses L3 MRs, thus, there is a HIT of 80 ms. Furthermore, handover decisions are performed by the user to allow itself to execute a handover even when the user cannot communicate to the serving BS. CHO is not suitable for URLLC services, but it does not require such a strict CU and DU architecture as LTM.

Fig. 6 presents the results of the baseline CHO and Proactive CHO, denoted as PCHO. As anticipated, all LTM-based algorithms have a much higher signaling overhead compared to CHO and PCHO, which can be seen when comparing Fig. 3 and Fig. 6a-b. PCHO also benefits from the future channel knowledge, following the same trend as PLTM. Specifically, as can be seen from Fig. 6a and Fig. 6b, the cell preparation and handover rates are reduced by up to 54% and 44%, respectively, while the sum throughput increases by a few percent. Moreover, the resource reservation drops to zero. This is illustrated in Fig. 6c.

VI. Conclusion

In this paper, we examined the benefits of proactive mobility management that employs predicted future SINR values in a controlled indoor environment, such as a factory floor with mobile robots performing tasks. One of our findings is that utilizing channel prediction solely for cell preparation offers no significant advantages because in LTM and CHO cell preparation is already decoupled from handover execution and is performed in advance. Our proactive approach reduces signaling by 76% and almost completely avoids resource reservation while simultaneously increasing the overall throughput. Normally, there is a trade-off between network throughput and handover rate; to achieve a low handover rate, decisions are delayed, causing users to remain connected to a base station that may not have the highest SINR. However, this trade-off is minimized with our proactive solution. In the future, we plan to validate the performance of the proposed proactive algorithms on real hardware using OpenAirInterface.

References

[1] M. Z. Chowdhury, M. Shahjalal, S. Ahmed, and Y. M. Jang, “6g wireless communication systems: Applications, requirements, technologies, challenges, and research directions,” *IEEE Open Journal of the Communications Society*, vol. 1, pp. 957–975, 2020.

[2] 3GPP, “Moderator summary for multi-beam enhancement: EVM,” 3rd Generation Partnership Project (3GPP), 3GPP TSG RAN WG1 102-e e-Meeting, Decision R1-2007151, 8 2020.

[3] ——, “NR; NR and NG-RAN Overall description; Stage-2,” 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.300, 7 2024, version 18.2.0. [Online]. Available: http://www.3gpp.org/DynaReport/38300.htm

[4] A. Gündogan, A. Badalıoğlu, P. Spapis, and A. Awada, “On the modelling and performance analysis of lower layer mobility in 5G-advanced,” in *Proc. of IEEE WCNC*, 2023.

[5] S. Ayvasik, F. Mehmeti, E. Babaians, and W. Kellerer, “Peach: Proactive and environment-aware channel state information prediction with depth images,” *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 7, no. 1, 2023.

[6] Y. Koda, K. Nakashima, K. Yamamoto, T. Nishio, and M. Morikura, “Handover management for mmWave networks with proactive performance prediction using camera images and deep reinforcement learning,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 2, 2019.

[7] U. Karabulut, A. Awada, I. Viering, A. N. Barreto, and G. P. Fettweis, “RACH optimization with decision tree based supervised learning for conditional handover in 5G beamformed systems,” *arXiv preprint arXiv:1910.11890*, 2019.

[8] A. Hassan, A. Narayanan, A. Zhang, W. Ye, R. Zhu, S. Jin, J. Carpenter, Z. M. Mao, F. Qian, and Z.-L. Zhang, “Vivisecting mobility management in 5G cellular networks,” in *Proc. of ACM SIGCOMM*, 2022.

[9] A. Prado, H. Vijayaraghavan, and W. Kellerer, “ECHO: Enhanced conditional handover boosted by trajectory prediction,” in *Proc. of IEEE GLOBECOM*, 2021.

[10] C. Lee, H. Cho, S. Song, and J.-M. Chung, “Prediction-based conditional handover for 5G mm-Wave networks: A deep-learning approach,” *IEEE Vehicular Technology Magazine*, vol. 15, no. 1, 2020.

[11] H. Martikainen, I. Viering, A. Lobinger, and T. Jokela, “On the basics of conditional handover for 5G mobility,” in *Proc. of IEEE PIMRC*, 2018.

[12] 3GPP, “NR; Radio Resource Control (RRC); Protocol specification,” 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.331, 7 2024, version 18.2.0. [Online]. Available: http://www.3gpp.org/DynaReport/38300.htm

[13] ——, “Study on channel model for frequencies from 0.5 to 100 GHz,” 3rd Generation Partnership Project (3GPP), Technical Report (TR) 38.901, 4 2024, version 18.0.0. [Online]. Available: http://www.3gpp.org/DynaReport/38901.htm

[14] ——, “User Equipment (UE) radio access capabilities,” 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.306, 9 2024, version 18.3.0. [Online]. Available: http://www.3gpp.org/DynaReport/38306.htm

[15] X. Lin, R. K. Ganti, P. J. Fleming, and J. G. Andrews, “Towards understanding the fundamentals of mobility in cellular networks,” *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, 2013.

[16] A. Prado, F. Stoeckeler, F. Mehmeti, K. Patrick, and W. Kellerer, “Enabling proportionally-fair mobility management with reinforcement learning in 5G networks,” *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 6, 2023.

[17] A. Prado, Z. Ding, F. Mehmeti, and W. Kellerer, “Mobility management for computation-intensive tasks in cellular networks with SD-RAN,” in *CNSM*, 2024.

[18] A. Prado, F. Mehmeti, and W. Kellerer, “Cost-efficient mobility management in 5G,” in *Proc. of IEEE WoWMoM*, 2023.

[19] S. B. Iqbal, A. Awada, U. Karabulut, I. Viering, P. Schulz, and G. P. Fettweis, “On the Modeling and Analysis of Fast Conditional Handover for 5G-Advanced,” *arXiv preprint arXiv:2204.06909*, 2022.