

Document Version

Final published version

Citation (APA)

Kalntis, M., Kuipers, F. A., & Iosifidis, G. (2025). CHOMET: Conditional Handovers via Meta-Learning. In *Proceedings of the 2025 23rd International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)* IEEE. <https://doi.org/10.23919/WiOpt66569.2025.11123350>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

In case the licence states "Dutch Copyright Act (Article 25fa)", this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

**Green Open Access added to [TU Delft Institutional Repository](#)
as part of the Taverne amendment.**

More information about this copyright law amendment
can be found at <https://www.openaccess.nl>.

Otherwise as indicated in the copyright section:
the publisher is the copyright holder of this work and the
author uses the Dutch legislation to make this work public.

CHOMET: Conditional Handovers via Meta-Learning

Michail Kalntis, Fernando A. Kuipers, and George Iosifidis

Delft University of Technology, The Netherlands

Emails: {m.kalntis, f.a.kuipers, g.iosifidis}@tudelft.nl

Abstract—Handovers (HOs) are the cornerstone of modern cellular networks for enabling seamless connectivity to a vast and diverse number of mobile users. However, as mobile networks become more complex with more diverse users and smaller cells, traditional HOs face significant challenges, such as prolonged delays and increased failures. To mitigate these issues, 3GPP introduced conditional handovers (CHOs), a new type of HO that enables the preparation (i.e., resource allocation) of multiple cells for a single user to increase the chance of HO success and decrease the delays in the procedure. Despite its advantages, CHO introduces new challenges that must be addressed, including efficient resource allocation and managing signaling/communication overhead from frequent cell preparations and releases. This paper presents a novel framework aligned with the O-RAN paradigm that leverages meta-learning for CHO optimization, providing robust dynamic regret guarantees and demonstrating at least 180% superior performance than other 3GPP benchmarks in volatile signal conditions.

Index Terms—Handover, Conditional Handover, Online Learning, Meta-Learning, Network Optimization.

I. INTRODUCTION

A. Motivation

THE heterogeneity of modern mobile networks, characterized by multiple radio access technologies, dense cell deployments, and high frequencies, introduces significant challenges in mobility management that might lead to increased delays, service interruptions, and Quality of Service (QoS) degradation [1]. The handover (HO) mechanism is key to addressing these issues by enabling smooth transitions between cells as users move and stream, browse, text, or work [2]. Although *traditional* HO mechanisms have existed for decades, it has been shown that, even in new radio access technologies (5G and beyond), HOs still face many challenges. The reason lies in the HO procedure itself, as it relies on *reactive* decision-making: a HO is initiated only when the signal quality of the serving cell deteriorates significantly [3]. This approach often leads to increased HO delays and failures (HOFs), particularly in dense and high-frequency environments where signal conditions fluctuate rapidly [1], [4].

To mitigate these issues, 3GPP recently introduced Conditional Handovers (CHOs) [2]. The main idea lies in reserving resources (e.g., physical resource blocks and spectrum) to multiple *candidate* cells for a single user *proactively*, namely, when signal conditions are still favorable (in contrast to traditional HOs that happen when conditions are poor), and offloading the final HO decision to the user. The novel idea of *network-configured, user-decided HOs*, in contrast to the traditional *network-configured, user-assisted HOs*, has recently been shown to significantly reduce HO delays and HOFs, rendering them a promising solution in different environments, such as 5G beamformed [5] or non-terrestrial networks [6].

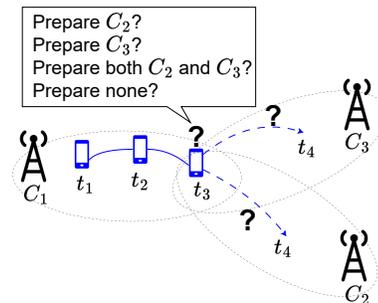


Fig. 1. CHO candidate cells for a moving user in consecutive times.

Nevertheless, CHOs introduce new challenges that must be carefully studied to ensure seamless and optimal mobility management in heterogeneous cellular networks. Specifically, the key difficulty lies in determining which cells to prepare while signal conditions vary, as preparing too many leads to resource waste and increased *signaling* due to extra communications, while too few might lead to traditional HO [4], [7]–[9].

Figure 1 illustrates these challenges: if at the intersection of the three cells (i.e., at t_3), we prepare only C_2 (C_3) but the user moves towards C_3 (C_2), then C_2 (C_3) resources will be wasted; at t_4 , extra signaling/communication will be needed to release these resources and a traditional HO will occur. Preparing both C_2 and C_3 , if they have enough capacity, is beneficial to the user, and soon after (i.e., at t_4) the trajectory becomes clear and a CHO is executed, the resources allocated to the unneeded cell can be released, creating, however, some signaling cost. Ideally, if only C_2 (C_3) is prepared at t_3 and the user’s movement towards C_2 (C_3) is confirmed at t_4 , the CHO can be executed efficiently, ensuring optimal resource allocation with minimal signaling overhead.

The goal of this paper is to design a *robust* algorithm that decides which cells to prepare in order to increase the rate of the users and reduce the traditional HO possibility (which comes with increased HOFs and delays), while at the same time minimizing the signaling needed for cell preparations and waste of resources. Robustness in this context refers to maintaining reliable performance without access to accurate information about the signal quality of the users at all times or their mobility patterns.

B. Methods & Contributions

After demonstrating the key limitations of traditional HOs and the promising capabilities of CHOs, we introduce a realistic system model that works in time slots, in accordance with online convex optimization (OCO) [10], and captures the impact of preparation decisions on both users (UEs) and cells. We consider a central network controller that focuses

on providing UEs with high rates and cells with saving their scarce resources, as well as reducing signaling overheads. Similar to other works [7], we provide tunable parameters that prioritize these different CHO criteria based on the preferences of a Mobile Network Operator (MNO). However, for the first time in CHO, (i) these parameters can change over time if preferences change, and can be different for each user and cell (e.g., different slices [11]), and (ii) assumptions for access to future signal measurements are dropped [12], rendering our modeling applicable to real-world scenarios.

Building on this modeling and problem formulation, we propose a learning algorithm, CHOMET, that updates its preparation decisions dynamically. To ensure CHOMET is adaptive to fast-changing and unpredictable signal conditions, multiple similar (but with different parameters) learners are employed, as well as a meta-learner that tracks and combines their performances, ensuring the best-performing ones are selected. To account for the signaling cost associated with frequent cell preparations and releases (i.e., decision changes across consecutive time slots), we leverage Smoothed Online Learning [13]. Following a rigorous theoretical analysis, we compare CHOMET through dynamic regret with an omniscient benchmark that has full information on future signal qualities, ensuring that their gap diminishes with time. Given that the operations of our algorithm are relatively lightweight, it can be implemented in the *near-real-time* of O-RAN [14], similarly to, e.g., [12], [15]. We conclude our study with rigorous experimental evaluations in both “easy” (stationary) and volatile scenarios of signal conditions, showing that CHOMET outperforms existing 3GPP-compliant mechanisms. The key contributions can be summarized as follows:

- We model, for the first time, CHO as a *smoothed online learning* problem, where the signaling/switching cost can even change over time.
- We propose a meta-learning algorithm that adjusts cell preparations dynamically, maximizing UEs’ rate and minimizing resource waste and signaling overheads. Our algorithm provides strong theoretical dynamic regret guarantees even under volatile signal conditions.
- We evaluate our algorithm in volatile and slow-changing signal conditions against 3GPP-compliant benchmarks, finding more than 180% performance improvement.

II. BACKGROUND

A. Traditional & Conditional HOs

Traditional HOs occur when the signal from the serving cell falls below a threshold (A2 event) or when the signal from a neighboring cell becomes offset better than the signal from the serving cell (A3 event) [3], [16], [17]. This *reactive* approach, where HOs are initiated *after* the signal conditions are getting/remaining poor for a predefined time, leads to high HO delay and HOF rates, particularly due to the high small cell density and the rapid signal degradation of FR2 (i.e., frequency ranges > 24 GHz) [1], [18]. Specifically, and as can be seen in Fig. 2, it is common for HOFs to occur when a user attempts to send a measurement report (MR) under deteriorating signal conditions; or even if the MR is

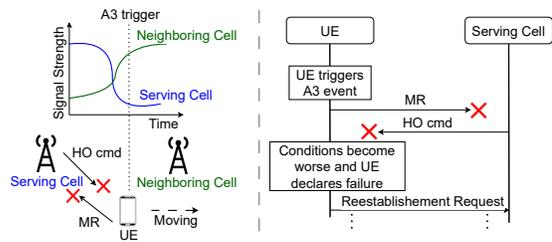


Fig. 2. A failed traditional HO (left) and its event timeline (right), where the MR or HO command may not reach the serving cell or UE, respectively.

successfully sent, the worsening signal conditions may prevent the user from receiving the subsequent HO command.

CHOs, designed as part of 3GPP Release 16 and enhanced in subsequent releases, address the limitations mentioned above by offloading part of the HO decision-making to the user, before the signal conditions deteriorate [2]. As can be seen in Fig. 3, a CHO decision is taken while the signal quality is still adequate (step 2), and the source cell can pre-configure multiple *candidate* target cells (steps 3, 4, 5) based on the MR sent from the UE (step 1). To conclude the *preparation phase* for the CHOs, the source cell provides monitoring conditions (step 6), such as hysteresis and offset parameters, and the *execution phase* starts: the user applies these conditions continuously (step 7) to evaluate the signal strength and quality –e.g., through signal-to-interference-plus-noise ratio (SINR)– of the source and candidate target cells.

If any of the predefined conditions are met (step 8), the UE directly executes the stored HO command (as if this command was just received) without needing to send a MR and await a reply from the source cell: procedures that could delay or/and fail due to the degradation of the signal. If more than a single cell meets the execution condition, the UE decides which cell to access. In our modeling (see Sec. III), we assume that this cell is the one with the highest SINR adjusted by a cell-specific threshold; however, it is straightforward to consider otherwise. In addition, the UE is not allowed to try the HO execution multiple times or to execute several HO attempts in parallel, towards different candidate cells [4]. CHO is finalized in the *completion phase*, where resources are released from the source cell, and a new path is established to the target cell (steps 9, 10, and after). We highlight that once the resources are reserved for a UE through the admission control (step 4), they are kept until a cancellation is sent (step 10b).

Conversely, in traditional HO, the HO decision (step 2) is made when the signal conditions, included in the MR (step 1), are deteriorating. Moreover, the HO decision in traditional HOs dictates the *one* target cell that the UE should connect to; and this transition (steps 6, 7) is executed *immediately* after the UE receives the command. In other words, the execution phase in traditional HOs occurs immediately after preparation, in contrast to CHOs where the gap between the two phases can be as large as 9–10 sec. [8].

B. CHO Key Trade-Offs & Selecting Criteria

Determining which cells to prepare is critical in CHOs for balancing resource efficiency and mobility robustness. Ideally,

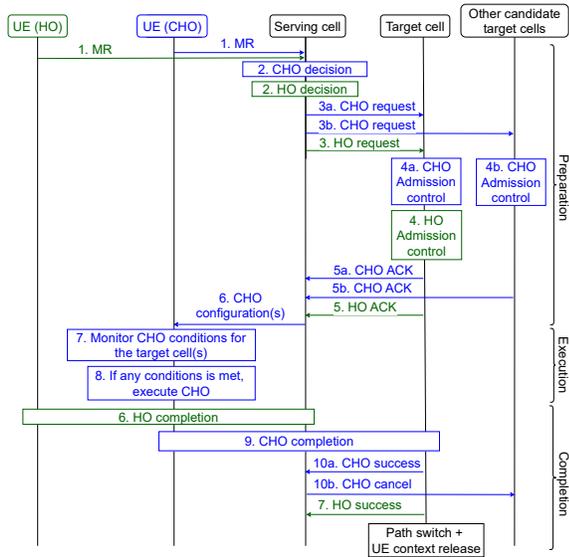


Fig. 3. Basic steps/procedures of UEs performing traditional (green) and conditional (blue) handover.

the MNO, whose goal is to save the scarce and valuable resources of its cells, would allow the preparation of the single cell to which the UE will connect in the next time slot and whose signal strength is high, thus, maximizing the UEs' rate. Since this cannot always be predicted with certainty because, e.g., the user is located at the edge of a cell and/or its trajectory is not known, multiple candidate cells are often prepared. These candidate cells are typically selected based on signal quality metrics (e.g., SINR), which have to exceed a predefined threshold for a specific time [4] (these thresholds are lower than those for traditional HO). This approach, while straightforward and easy to implement, can lead to inefficient resource usage, as resources of prepared cells cannot be used by another UE until they are released.

At the same time, a long list of prepared cells does not necessarily increase the likelihood of including the “correct” target cell (i.e., the cell that the user will connect in the next time slot), especially if all prepared cells exhibit low signal quality. Conversely, while preparing fewer cells may conserve resources, it increases the risk of traditional HOs (and thus, more HOFs) to occur, if the correct target cell is not included in the prepared list; preparing no cells leads to traditional HOs.

Another aspect to consider for optimizing CHOs is the signaling cost of preparing cells. For instance, in environments such as FR2, the small cell density and rapid signal fluctuations result already in more frequent HOs and HOFs, which contribute to significant signaling overhead. The additional burden of continuously preparing and releasing cells due to constant signal variations further exacerbates this overhead [8], [19]. For that reason, while initial 3GPP releases state that UEs should release CHO candidate cells after (any) successful HO completion to save resources [2] (see also Fig. 3), follow-up works suggest that doing so is not always optimal in terms of reducing HOFs and signaling overhead [18].

These trade-offs emphasize that the selection process for cell preparations should focus on finding the balance (and sometimes prioritizing) among identifying the “best” –in terms

of signal quality– cells, while keeping the number of prepared cells as small as possible to save resources and avoid frequent preparations and releases to reduce signaling overheads.

III. SYSTEM MODEL & PROBLEM FORMULATION

A. Modeling Components

We consider a heterogeneous cellular network comprising a set \mathcal{J} of J cells serving a set \mathcal{I} of I UEs. We assume that a central network controller takes decisions for multiple UEs/cells in a time-slotted manner for a set \mathcal{T} of T slots [12], [20] with 10msec–1sec duration for each slot, as realized with xApps in *near-real-time* (10msec–1sec granularity) of O-RAN [14]. The key metric is the SINR for the signal delivered by cell $j \in \mathcal{J}$ to user $i \in \mathcal{I}$ in slot $t \in \mathcal{T}$:

$$s_{ij}(t) = \frac{q_j \phi_{ij}(t)}{W_j \sigma^2 + \sum_{k \in \mathcal{B}_j} q_k \phi_{ik}(t)}, \quad (1)$$

where q_j is the transmit power of cell j , \mathcal{B}_j the set of cells that operate in the same frequency as j , $\phi_{ij}(t)$ the channel gain (including pathloss, shadowing, and antenna gains), W_j is the bandwidth of cell j , and σ^2 the power spectral density. In line with previous works [20]–[23], $s_{ij}(t)$ is the average SINR in the slot t , since UEs report multiple SINRs during each slot. The maximum possible rate for UE i , if it uses cell j exclusively, can be expressed as $c_{ij}(t) = W_j \log_2(1 + s_{ij}(t))$. Moreover, we introduce the random variable $u_j(t) \in [0, 1]$ to show the part of the bandwidth W_j that is available.

We denote by $x_{ij}(t) \in \{0, 1\}$ the preparation decision: $x_{ij}(t) = 1$ means that cell j is prepared in slot t for user i , and $x_{ij}(t) = 0$ otherwise. Also, we define the vector $\mathbf{x}_t = (x_{ij}(t) \in \{0, 1\}, i \in \mathcal{I}, j \in \mathcal{J})$, with the decision set:

$$\mathcal{X} = \left\{ \mathbf{x} \in \{0, 1\}^{I \cdot J} \mid \sum_{j \in \mathcal{J}} x_{ij} \leq J, i \in \mathcal{I} \right\},$$

and its convex hull $\mathcal{X}^c = \text{co}(\mathcal{X})$ that relaxes the integrality, i.e., $\mathbf{x} \in [0, 1]^{I \cdot J}$. For $p_{ij}(t)$ being the probability that cell j will be the highest-rate cell for UE i during slot t , then this CHO will be realized with probability $p_{ij}(t)x_{ij}(t)$, where $\mathbf{p}_t = (p_{ij}(t) \in \{0, 1\}, i \in \mathcal{I}, j \in \mathcal{J})$ and $p_{ij}(t) = \mathbb{1}\{j = \arg \max_k (s_{ik}(t) - o_k(t))\}$, with $o_j(t)$ being the cell-specific offsets that can vary in each slot t , resembling the offsets used in, e.g., the A3 event [3], [9], [16] to avoid often HOs/CHOs. It is important to note that a user can be served from only one cell at a time, which is inferred from the definition of \mathbf{p} .

B. Problem Statement

As mentioned earlier, CHOs are *network-configured, user-decided HOs*. Thus, it is the network controller that decides the preparations of the cells based on its belief regarding the cell that will be the best for each UE, while minimizing resource waste and signaling overheads; and each UE is allowed to execute the CHO directly to one of these prepared cells. Otherwise, a traditional HO occurs, which incurs additional costs due to higher HO delays and HOF probabilities.

With these in mind, we introduce the *utility* function that the network controller wishes to maximize as follows:

$$g_t(\mathbf{x}) \triangleq \sum_{i=1}^I \sum_{j=1}^J \left(x_{ij}(t) p_{ij}(t) u_j(t) \log c_{ij}(t) - \beta_t x_{ij}(t) (1 - p_{ij}(t)) - \gamma_t p_{ij}(t) (1 - x_{ij}(t)) \right), \mathbf{x} \in \mathcal{X}.$$

The first term of g_t defines the rate of a user, which is non-zero only for its served cell and is discounted by u_j due to exogenous (i.e., independent of the preparation decisions) effects; e.g., other UEs that executed traditional HO. The logarithmic transformation balances the sum-rate across all users to achieve fairness [20]; however, we note that other mappings (e.g., linear) can be used to capture the specifics of different applications. The second term of g_t introduces a penalty if cells other than the highest-rate ones are prepared to reduce resource waste, and the third term represents an additional cost if the highest-rate cell is not prepared. The scalarization parameters β_t and γ_t are used to normalize units and prioritize one criterion over the other according to the preferences of each MNO that can even change over time or be different for each UE-cell pair due to different slices [11].

At the same time, the goal of the network controller is to minimize the signaling/switching overheads (or, similarly, maximize their negation), which are captured using $-\delta_t \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_{B_t}$, with $\mathbf{x}_t, \mathbf{x}_{t-1} \in \mathcal{X}$. This presents the *switching cost* induced by the signaling of preparing and releasing cells (see Fig. 3) scaled by $b_{ij}(t)$, as each cell j might have different costs for preparing and releasing cells for each user in slot t due to fluctuating traffic demands. More precisely, we define with $\mathbf{B}_t = \text{diag}(\mathbf{b}_n(t) > 0)$ a positive definite matrix which has on its diagonal the signaling weights $b_n(t) \in [0, 1]$, $n = i \cdot j$ when UE i prepares and releases cell j , and $\|\cdot\|_{B_t}$ is its induced norm that can change over time t , i.e., $\|\mathbf{x}\|_{B_t}^2 = \sum_n b_n(t) x_n^2$ and its dual $\|\mathbf{x}\|_{B_t^*}^2 = \sum_n x_n^2 / b_n(t)$ [24]. The role of the parameter δ_t is similar to β_t and γ_t . Thus, the overall problem that the network controller wishes to solve can be expressed as:

$$\begin{aligned} \mathbb{P} : \max_{\{\mathbf{x}_t\}_t} & \sum_{t=1}^T \left(g_t(\mathbf{x}_t) - \delta_t \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_{B_t} \right) \\ \text{s.t.} & \mathbf{x}_t \in \{0, 1\}^{I \cdot J}, \forall t \in \mathcal{T}, \end{aligned}$$

where $g_t(\mathbf{x}_t) - \delta_t \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_{B_t}$ is the *objective* function. Solving the problem at the beginning of the horizon T is challenging due to several factors. First, the controller at $t = 1$ has no knowledge of the future SINR values for each UE (it has no access or effect on how the UEs will move in the future). In fact, SINRs remain unknown even at the beginning of each slot, as knowing the current SINR and deciding afterwards the cell preparations lead to trivial solutions¹. In other words, the cell preparations should be decided based on the history (one or multiple slots) of the UE's signal, a problem that is further perplexed by the discreteness of the preparation

¹A similar way to approach the problem would be to decide the cell preparations in the previous slot, and then access the SINRs at the beginning of each slot, see [7].

variables [25]. What is more, the signaling overhead depends on the change in preparation state, as a cell that remains prepared does not incur additional signaling. This introduces a memory effect into the system, as past decisions influence current preparation decisions (due to signaling costs). Finally, we highlight that the B_t -norm implies that the switching costs change over time; however, this cost in time t does not change based on previous preparation decisions $\{\mathbf{x}_\tau\}$, $\tau = 1, \dots, t-1$.

Therefore, our goal is to design an algorithm that is oblivious to the time-varying and unknown parameters and maximizes the users' rate, while keeping the signaling costs and the amount of wasted resources to a minimum.

IV. ALGORITHM DESIGN

We approach \mathbb{P} as a *smoothed online learning* problem and address it through *meta-learning* based on the *experts* framework [26], [27]. More precisely, and due to the problem's unknown variability where optimal preparation decisions can change drastically between slots (due to the variability of the signal conditions), we deploy K similar learning agents, known as *experts*, but each with a different learning step $\theta = (\theta_k, k \in \mathcal{K})$ (to ensure at least one will perform well), and a meta-learner whose goal is to find the best-performing expert(s) on-the-fly, while signal conditions might vary.

More precisely, at the beginning of each slot t , with $t = 1, \dots, T$, each expert k shares its suggestion $\mathbf{x}_t^k \in \mathcal{X}$ regarding which cells should be prepared for each UE; we assume $\mathbf{x}_0^k = 0$, $k \in \mathcal{K}$. The meta-learner, which has assigned weights (i.e., "trust") $\mathbf{w}_t = (w_t^k, k \in \mathcal{K})$, with $\mathbf{w}_t^\top \mathbf{1}_K = 1$ to each expert, synthesizes them as follows:

$$\mathbf{x}_t^m = \sum_{k \in \mathcal{K}} w_t^k \mathbf{x}_t^k. \quad (2)$$

To ensure that the decision \mathbf{x}_t^m of the meta-learner is implementable (i.e., binary decision variable), we use a quantization function $Q_{\mathcal{X}}$ implementing any unbiased sampling technique, outputting $\mathbf{x}_t \in \mathcal{X}$; here, we opt for Madow sampling [28]. We underline that, for $\mathbf{x} \in \mathcal{X}^c$, g_t is concave (as a linear function), and subtracting the B_t -norm preserves the property.

Once cell preparations are finalized and implemented, the controller observes the actual signal qualities from this slot t , $s_{ij}(t)$, $\forall i \in \mathcal{I}$ and $j \in \mathcal{J}$ (we fill with zero the cells that are unreachable for each UE), and calculates the gradient of only the utility function for the "synthesized" decision (and not the decision of each expert \mathbf{x}_t^k), $\nabla g_t(\mathbf{x}_t)$, which is sent to all experts. Then, the meta-learner uses the surrogate (i.e., partially linearized) loss:

$$\ell_t(\mathbf{x}_t^k) = \langle \nabla g_t(\mathbf{x}_t), \mathbf{x}_t^k - \mathbf{x}_t \rangle - \delta_t \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_{B_t} \quad (3)$$

to update the weights:

$$w_{t+1}^k = \frac{w_t^k e^{\eta \ell_t(\mathbf{x}_t^k)}}{\sum_{k \in \mathcal{K}} w_t^k e^{\eta \ell_t(\mathbf{x}_t^k)}}. \quad (4)$$

Lastly, experts update their choices via online gradient ascent:

$$\mathbf{x}_{t+1}^k = \Pi_{\mathcal{X}} \left(\mathbf{x}_t^k + \theta_k \nabla g_t(\mathbf{x}_t) \right), \quad (5)$$

Algorithm 1: CHO via Meta-Learning (CHOMET)

1 **Required:** Step η for meta-learner and $\{\theta_k\}_{k \in \mathcal{K}}$ for experts
2 **Initialize:** Sort $\theta_1 \leq \theta_2 \leq \dots \leq \theta_K$ and set
 $w_1^k = \frac{1+(1/K)}{k(k+1)}, \forall k \in \mathcal{K}$
3 **for** $t = 1, 2, \dots, T$ **do**
4 Each expert $k \in \mathcal{K}$ shares its decision \mathbf{x}_t^k
5 The controller synthesizes \mathbf{x}_t^m using (2)
6 The controller implements $\mathbf{x}_t = Q_{\mathcal{X}}(\mathbf{x}_t^m)$
7 The controller sends $\nabla g_t(\mathbf{x}_t)$ to experts
8 The controller updates its weights using (4)
9 Each expert updates its decision using (5)
end

where η is the meta-learning step. A summarization of the steps can be seen in Algorithm 1 (CHOMET).

The performance of the algorithm will be assessed using the *Expected Dynamic Regret*, defined as in [13]:

$$\mathbb{E}[\mathcal{R}_T] \triangleq \sum_{t=1}^T g_t(\mathbf{x}_t^*) - \sum_{t=1}^T \mathbb{E} \left[g_t(\mathbf{x}_t) - \delta_t \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_{B_t} \right], \quad (6)$$

where $\{\mathbf{x}_t\}_t$ is the algorithm decisions, $\{\mathbf{x}_t^*\}_t$ is the benchmark that is the solution of \mathbb{P} , and the expectation captures the randomization due to the quantization $Q_{\mathcal{X}}$. Our goal is to design an algorithm that ensures this gap diminishes with time, $\lim_{T \rightarrow \infty} \mathbb{E}[\mathcal{R}_T]/T = 0$ for any possible benchmark sequence $\{\mathbf{x}_t^*\}_t$. In other words, we compare our algorithm with the best oracle that has full information on SINR for all users and cells at the beginning and for the entire horizon T . This serves as a highly competitive benchmark, exceeding the static (single best solution for all time slots), or dynamic one (best solution for each time slot independently).

To proceed, we observe that from the problem formulation (see Section III), the gradients of all functions are bounded by G , i.e., $\|\nabla g_t(\mathbf{x})\|_2 \leq G$, and the diameter of the domain is also bounded by D , i.e., $\|\mathbf{x} - \mathbf{x}'\|_2 \leq D$, where $D = \sqrt{IJ}$ and $G = \sqrt{I(J-1)\beta_{\max}^2 + I(\log c_{\max} + \gamma_{\max})^2}$, respectively, with $\beta_{\max} = \max_{t \in \mathcal{T}} \{\beta_t\}$, $\gamma_{\max} = \max_{t \in \mathcal{T}} \{\gamma_t\}$, $c_{\max} = \max_{t \in \mathcal{T}} \{c_t\}$. Therefore, defining $b_{\max} = \max_{n \leq I \cdot J} \{b_n(t)\}$ for $t \in \mathcal{T}$, it holds that $\|\mathbf{x} - \mathbf{x}'\|_{B_t} \leq D\sqrt{b_{\max}} \triangleq D_B$, $\|\mathbf{x} - \mathbf{x}'\|_{B_t^*} \leq D/\sqrt{b_{\max}} \triangleq D_{B^*}$, and $\|\nabla g_t(\mathbf{x})\|_{B_t} \leq G\sqrt{b_{\max}} \triangleq G_B$.

Lemma 1 (Performance Analysis / Optimality Guarantee). *Using the parameters:*

- $K = \lceil \log_2 \sqrt{1+2T} \rceil + 1$,
- $\theta_k = 2^{k-1} \sqrt{\frac{D_B^2}{T(G^2+2G_B)}}$, $k = 1, \dots, K$,
- $\eta = 1/\sqrt{T\nu}$, with $\nu \triangleq (D_B+1/8)(GD+2D_B)^2$,
- $P_T = \sum_{t=1}^T \|\mathbf{x}_t^* - \mathbf{x}_{t-1}^*\|_{B_t}$ (path length),

the discrete decisions $\{\mathbf{x}_t\}_t$, where $\mathbf{x}_t \in \mathcal{X}$ ensure:

$$\mathbb{E}[\mathcal{R}_T] \leq \sqrt{T} \left(\sqrt{\nu} (1 + \ln(1/w_1^k)) + \sqrt{(G^2 + 2G_B)(D_B^2 + 2D_{B^*}P_T)} \right) + T(G + \sqrt{b_{\max}})\sqrt{IJ}/2, \quad (7)$$

Proof. The proof follows by tailoring the main result of [13], [29], [30], which, however, solve the UE-cell association problem. Specifically, after setting w.l.o.g. $\beta_t = \gamma_t = \delta_t = 1$, eq. (6) can be rewritten as:

$$\mathbb{E}[\mathcal{R}_T] = \mathbb{E} \left[\sum_{t=1}^T \left(g_t(\mathbf{x}_t^*) - (g_t(\mathbf{x}_t) - \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_{B_t}) + (g_t(\mathbf{x}_t^m) - \|\mathbf{x}_t^m - \mathbf{x}_{t-1}^m\|_{B_t}) - (g_t(\mathbf{x}_t^m) - \|\mathbf{x}_t^m - \mathbf{x}_{t-1}^m\|_{B_t}) \right) \right] = \sum_{t=1}^T \left(g_t(\mathbf{x}_t^*) - (g_t(\mathbf{x}_t^m) - \|\mathbf{x}_t^m - \mathbf{x}_{t-1}^m\|_{B_t}) \right) + \mathbb{E} \left[\sum_{t=1}^T \left((g_t(\mathbf{x}_t^m) - \|\mathbf{x}_t^m - \mathbf{x}_{t-1}^m\|_{B_t}) - (g_t(\mathbf{x}_t) - \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_{B_t}) \right) \right]. \quad (8)$$

Therefore, to characterize the performance of CHOMET, we begin by bounding the expected dynamic regret of the relaxed (continuous) preparation decisions $\{\mathbf{x}_t^m\}_t$, $\mathbf{x}_t^m \in \mathcal{X}^c$, and afterwards, the error of the (implementable) discrete decisions, which correspond to the first and second terms of eq. (8), respectively. As can be seen from the first term of the result in eq. (7), sublinear dynamic regret can be achieved for the relaxed preparation decisions. It follows by bounding the regret of each expert w.r.t. the benchmark and then the regret of the meta-learner w.r.t. any expert. For the extra cost/error introduced due to the discretization of the preparation decision through the quantization routine $Q_{\mathcal{X}}$, we bound the second term of eq. (8) as follows:

$$\mathbb{E} \left[\sum_{t=1}^T \left(g_t(\mathbf{x}_t^m) - \|\mathbf{x}_t^m - \mathbf{x}_{t-1}^m\|_{B_t} - g_t(\mathbf{x}_t) + \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_{B_t} \right) \right] \stackrel{(*)}{\leq} \left(G + \sqrt{b_{\max}} \right) \sum_{t=1}^T \sqrt{\mathbb{E} \left[\sum_{i=1}^I \sum_{j=1}^J \left(x_{ij}(t) - \mathbb{E}[x_{ij}(t)] \right)^2 \right]} \stackrel{(**)}{\leq} T(G + \sqrt{b_{\max}})\sqrt{IJ}/2,$$

leading to eq. (7). Specifically, for (*), we use Jensen's inequality, the linearity of expectation, and the Lipschitz continuity. In particular, for the Lipschitz continuity, considering two points $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{X}^c$, using the triangle inequality and the Lipschitz constant G of g_t and of the Euclidean norm, and defining $\mathbf{x}_{t-1} = C$, we have:

$$\left| (g_t(\mathbf{x}_1) - \|\mathbf{x}_1 - C\|_{B_t}) - (g_t(\mathbf{x}_2) - \|\mathbf{x}_2 - C\|_{B_t}) \right| \leq |g_t(\mathbf{x}_1) - g_t(\mathbf{x}_2)| + \left| \|\mathbf{x}_1 - C\|_{B_t} - \|\mathbf{x}_2 - C\|_{B_t} \right| \leq (G + \sqrt{b_{\max}})\|\mathbf{x}_1 - \mathbf{x}_2\|.$$

Lastly, we notice that (**) calculates the variance of the binary \mathbf{x}_t . Given that $\mathbf{x}_t \in \{0, 1\}^{I \cdot J}$ and the quantization routine $Q_{\mathcal{X}}$ implements Madow sampling, the maximum variance each component can obtain is 1/4 (the variance per element is $q(1-q)$ due to the Bernoulli trial with probability q for each x_{ij} , and maximizes for $p = 0.5$). Hence, we can upper bound

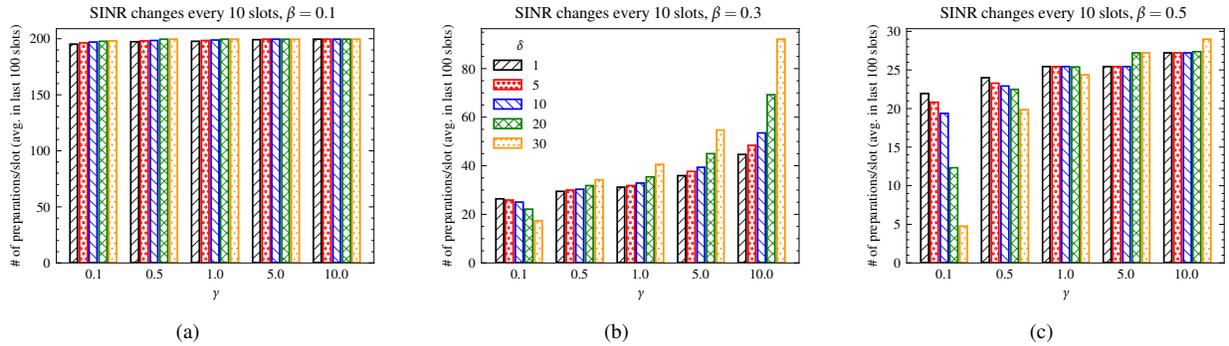


Fig. 4. Effect of parameters $\beta_t \equiv \beta, \gamma_t \equiv \gamma$, and $\delta_t \equiv \delta$, for $T = 5k$ slots in the volatile scenario (SINR changes ever 10 slots) when running CHOMET: (a) $\beta = 0.1$, (b) $\beta = 0.3$, and (c) $\beta = 0.5$. Y-axis symbolizes the average number (last 100 slots) of preparations per slot.

the expression in the square root with the constant $IJ/4$. \square

With Lemma 1, we bound the expected regret of the implementable cell preparations, in contrast to previous works [20], [21]. Even though the relaxed/continuous preparations achieve sublinear dynamic regret, the discretization introduces an unavoidable non-diminishing error. As has been shown in the UE-cell association context [29], this error is small in practical scenarios, and the algorithm converges towards the optimal solutions (see also Sec. V).

V. PERFORMANCE EVALUATION

We assess CHOMET under different synthetic scenarios to verify its robustness and showcase its learning convergence. For that, we compare our algorithm against the baseline 3GPP-compliant HO/CHO algorithm using the A3 event, which is the algorithm currently being used by MNOs and antenna vendors [2], [3]. We use the tuple notation ($\#$ Best BS, TTT) to refer to these comparators, where the first argument refers to the number of top- N best cells (i.e., highest SINR) that are prepared for each user in each slot, while the second is the time-to-trigger (TTT, i.e., number of consecutive slots a cell must have been in the top- N before is prepared), resembling the A3 event. For example, algorithm (3, 8) prepares the 3 highest-SINR cells for each user only if these cells remain the highest ones for at least 8 slots. Moreover, we compare CHOMET with an optimal Oracle that solves the optimization problem in every step using CVXPY [31]. We note that the comparison with the oracle that has complete knowledge of the future is computationally intensive for mixed-integer programs [21], [32]; however, our dynamic regret guarantees hold, as can be verified from Sec. IV.

In line with prior works [12], [29], we select two synthetic scenarios as follows, for $T = 5k$ slots: (i) *stationary*: the SINR $s_{ij}(t), i \in \mathcal{I}, j \in \mathcal{J}$ remains almost constant across all slots $t \in \mathcal{T}$, changing only once every 600 slots, and (ii) *volatile*: $s_{ij}(t)$ fluctuates every 10 slots within the range of [10, 30]dB [33], encompassing poor to excellent values. In both scenarios, we randomly select the bandwidths $W_j \in \{5, 10, 15, 20\}$ MHz [34], while B_t takes random values within [0, 1], as the offsets $o_j(t), \forall j \in \mathcal{J}$. We select $I = 20$ UEs and $J = 10$ cells to facilitate the calculation of the average regret, as determining the best oracle is computationally intensive. However, we underline that the best oracle is not needed to run CHOMET.

First, Fig. 4 shows the effect of the parameters β_t, γ_t , and δ_t on CHOMET, which, for simplicity, are assumed to be constant for the entire duration of the experiments, namely $\beta_t \equiv \beta, \gamma_t \equiv \gamma, \delta_t \equiv \delta$, and $u_t = 1, \forall t \in \mathcal{T}$, where $T = 5k$ and SINR changes every 10 slots (volatile scenario). It is important to note that due to the different values that each component of the objective function takes, we focus on the comparison of the values within the same parameter (i.e, they act as scalarization values too). For example, $\beta = 0.5 > \gamma = 5$ does not imply that more importance is given to γ ; however, choosing $\beta = 0.5$ instead of $\beta = 0.1$ does. We choose $\beta = \{0.1, 0.3, 0.5\}$, $\gamma = \{0.1, 0.5, 1, 5, 10\}$ and $\delta = \{1, 5, 10, 20, 30\}$. The y-axis of Fig. 4 shows the average number of preparations per slot, for the last 100 slots; notice that, due to considering 20 UEs and 10 cells, this number cannot exceed 200.

For a specific γ , the effect of β and δ on the number of prepared cells depends on the interplay of these parameters and should be carefully examined. If γ , the penalty for not preparing the best cell, is small (e.g., $\gamma = 0.1$) and β is high (e.g., $\beta = 0.5$ so fewer cells can be prepared), then as δ increases, the switching cost becomes more significant, making it less beneficial to prepare different cells to determine the best one. Consequently, the number of prepared cells *decreases*; e.g., setting $\beta = 0.5$ and $\gamma = 0.5$, we observe 24 and 19 preparations on average for $\delta = 1$ vs $\delta = 30$, respectively. Conversely, if γ is large, the penalty for not preparing the best cell is substantial, and as δ grows and switching becomes more costly, the algorithm decides to keep more cells prepared for more slots. As a result, the number of prepared cells *increases* even up to $\times 2$ times, as we can see for $\beta = 0.3, \gamma = 10$, and $\delta = 1$ vs $\delta = 30$.

For a fixed δ , the number of prepared cells *increases* with γ , as higher γ imposes a greater penalty for not preparing the highest-SINR cell. To mitigate this penalty, the algorithm prepares more cells to ensure that the best is found. This trend is consistent for all three β values considered, although a higher β limits the number of cells other than the highest-SINR that can be prepared, leading to differences in absolute values. For instance, our algorithm makes approximately 200 preparations on average per slot in Fig. 4(a), meaning that all cells for all users are prepared in each slot, while we observe at most 29 preparations per slot (i.e., 1–2 cells per UE) in Fig. 4(c); for that reason, lower values than $\beta = 0.1$ or higher than $\beta = 0.5$ are not considered in the evaluation. We underline

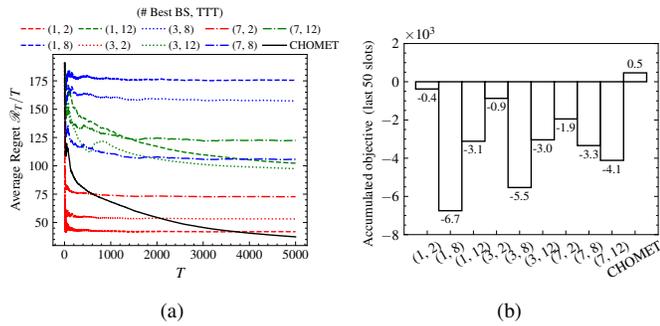


Fig. 5. Volatile scenario (SINR changes ever 10 slots) for $\beta_t = 0.5$, $\gamma_t = 10$, and $\delta_t = 5, \forall t \in \mathcal{T}$ with $T = 5k$ slots: (a) average dynamic regret and (b) total objective values for the last 50 slots, of CHOMET and benchmarks.

that the interplay of these parameters also changes significantly with the volatility/stationarity of the conditions (i.e., SINR).

In the volatile case, Fig. 5(a) shows the average dynamic regret of our proposed algorithm and the 3GPP-compliant competitors, with the former surpassing all (q, r) , for $q = \{1, 3, 7\}$ and $r = \{8, 12\}$ (blue and green lines) by up to 375% in slot $t = 5k$. Even though at first glance competitors $(1, 2)$, $(3, 2)$ and $(7, 2)$ seem to have comparable performance with CHOMET, we underline that their average dynamic regret stays almost constant for all slots (“stuck” in sub-optimal decisions). This claim can be verified from Fig. 5(b), where the accumulated objective in the last 50 slots of the best comparator, namely $(1, 2)$, is 180% less than CHOMET. Lastly, in a stationary (almost static) case, such as this of Fig. 6 where SINR changes very slowly (5 times in the total $T = 3k$), comparators $(1, 2)$, $(1, 8)$ and $(1, 12)$ behave similarly to CHOMET in terms of average regret and total objective values. This is reasonable, as preparing only the single highest-SINR cell (which is the one that the user is allocated) is the best policy to follow when conditions stay the same. Therefore, we verify with different scenarios that our algorithm, CHOMET, is adaptive to both volatile and stationary/static cases and provably approaches the behavior of an omniscient benchmark.

VI. RELATED WORK

CHO is a novel solution included in 3GPP Release 16 that focuses on mobility robustness by reducing the HO delay and the increased number of HOFs that traditional HOs are facing [2], [35], [36]; kindly refer to [1], [29] for a comprehensive literature review of traditional HOs. Recently, CHO has been shown to be successful in several scenarios, e.g., in non-terrestrial networks [6], [37], [38], for integrated access and backhaul, and in 5G NR-unlicensed [4], as well as in beamforming and contention-free random access [5], [39]. In [9], by tweaking the CHO thresholds (also studied in [40]) for preparing cells and executing HO towards a prepared cell, the authors discover that CHOs offer mobility robustness (more than 60% reduced failures); however, improper parametrization could lead to increased signaling. Similarly, [4] discusses the basics of this new type of HO and reports that UEs with velocities greater than 30km/h benefit the most from CHO, exhibiting 3–4 times lower failures than UEs with no CHO enabled (i.e., traditional HOs). However, these works

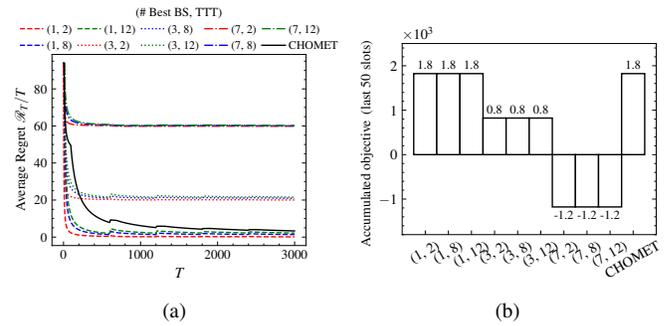


Fig. 6. Stationary scenario (SINR changes every 600 slots) for $\beta_t = 0.5$, $\gamma_t = 10$, and $\delta_t = 5, \forall t \in \mathcal{T}$ with $T = 3k$ slots: (a) average dynamic regret and (b) total objective values for the last 50 slots, of CHOMET and benchmarks.

offer either solely descriptive and simulation-based insights or propose solutions without performance guarantees.

In [41], [42], historical HO data (e.g., the likelihood of each cell becoming the next serving cell for a given UE) are used, and the cells to prepare for each user are determined based on past HO frequencies. To enhance CHOs, trajectory prediction is used in [43], where the controller decides which cells should be prepared/released based on the predicted distance between the UEs and the cells. Other prediction-based CHO schemes take as input the signal quality of each user at different slots and rely on Machine/Deep Learning (ML/DL) models to either predict the probabilities of the user’s connected (i.e., associated) cell in the next time slot [44], or the future signal measurements [45]. However, these approaches rely on static historical data, which do not account for changes in user mobility patterns over hours/days [1], or require frequent re-configuration to adapt to such changes, limiting their flexibility in dynamic environments.

Moreover, the authors of [8], [19] study the benefits of fast (i.e., consecutive) CHO proposed in [18] to facilitate the use of previously prepared cells and avoid the unnecessary signaling to prepare the same cells, after (any) successful HO. In our work, we introduce tweakable parameters based on the preferences of the MNOs that determine the signaling cost weights, allowing flexibility in cell preparation decisions. Higher signaling weights penalize heavily the frequent preparations, capturing the idea of [8], [18], [19], while lower weights allow them to occur often (i.e., even after any successful HO) [2]. Similarly to our work, the authors of [7] form an optimization problem to decide which cells to prepare for each user in order to minimize CHO costs. Our study differs from this work, as (i) we learn on the fly the parameters and do not require offline training, and (ii) our perturbation model is an adversarial one, where the various random parameters can even be selected by an attacker; still, all results hold.

VII. CONCLUSIONS

This study focuses on optimizing Conditional Handovers (CHOs), a novel type of handovers (HOs) introduced by 3GPP to mitigate the issues of traditional HOs, using Smooth Online Learning (SOL). We develop a dynamic meta-learning algorithm that maximizes the users’ rate, while at the same time minimizing the signaling needed for cell preparations and

the waste of resources, using tunable parameters for flexible and customizable prioritization of these criteria. Our proposed algorithm, CHOMET, aligned with the O-RAN paradigm, demonstrates robust theoretical performance guarantees even in challenging environments, as exemplified in its evaluation against the currently deployed 3GPP-compliant algorithms. Future works will focus on reducing the discretization error and validating the algorithm in real-world scenarios.

ACKNOWLEDGMENTS

We sincerely thank Naram Mhaisen and Dr. Andra Lutu for the insightful discussions. This work has been supported by the National Growth Fund through the Dutch 6G flagship project “Future Network Services”, and the European Commission through Grant No. 101139270 (ORIGAMI) and 101192462 (FLECON-6G).

REFERENCES

- [1] M. Kalntis, J. Suárez-Varela, J. O. Iglesias, A. K. Bhattacharjee, G. Iosifidis, F. A. Kuipers, and A. Lutu, “Through the Telco Lens: A Countrywide Empirical Study of Cellular Handovers,” in *Proc. of IMC*, 2024.
- [2] 3GPP, “5G; NR; NR and NG-RAN Overall description; Stage-2,” Technical Specification (TS) 38.300, 2024.
- [3] —, “5G; NR; Radio Resource Control (RRC); Protocol specification,” Technical Specification (TS) 38.331, 2024.
- [4] J. Stanczak, U. Karabulut, and A. Awada, “Conditional Handover in 5G - Principles, Future Use Cases and FR2 Performance,” in *Proc. of IWCMC*, 2022.
- [5] S. Bin Iqbal, U. Karabulut, A. Awada, A. Noll Barreto, P. Schulz, and G. P. Fettweis, “Mobility Performance Analysis of RACH Optimization Based on Decision Tree Supervised Learning for Conditional Handover in 5G Beamformed Networks,” in *Proc. of EW*, 2023.
- [6] E. Juan, M. Lauridsen, J. Wigard, and P. Mogensen, “Performance Evaluation of the 5G NR Conditional Handover in LEO-based Non-Terrestrial Networks,” in *Proc. of WCNC*, 2022.
- [7] A. Prado, F. Mehmeti, and W. Kellerer, “Cost-Efficient Mobility Management in 5G,” in *Proc. of WoWMoM*, 2023.
- [8] S. Bin 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,” in *Proc. of PIMRC*, 2022.
- [9] H. Martikainen, I. Viering, A. Lobinger, and T. Jokela, “On the Basics of Conditional Handover for 5G Mobility,” in *Proc. of PIMRC*, 2018.
- [10] E. Hazan, “Introduction to Online Convex Optimization,” *Foundations and Trends in Optimization*, vol. 2, no. 3–4, 2016.
- [11] P. Rost, C. Mannweiler, D. S. Michalopoulos, C. Sartori, V. Sciancalepore, N. Sastry, O. Holland, S. Tayade, B. Han, D. Bega, D. Aziz, and H. Bakker, “Network Slicing to Enable Scalability and Flexibility in 5G Mobile Networks,” *IEEE Communications Magazine*, vol. 55, no. 5, pp. 72–79, 2017.
- [12] M. Kalntis, G. Iosifidis, and F. A. Kuipers, “Adaptive Resource Allocation for Virtualized Base Stations in O-RAN With Online Learning,” *IEEE Trans. on Communications*, vol. 73, no. 3, pp. 1787–1800, 2025.
- [13] L. Zhang, W. Jiang, S. Lu, and T. Yang, “Revisiting Smoothed Online Learning,” in *Proc. of NeurIPS*, 2021.
- [14] A. Garcia-Saavedra and X. Costa-Pérez, “O-RAN: Disrupting the Virtualized RAN Ecosystem,” *IEEE Communications Standards Magazine*, vol. 5, no. 4, pp. 96–103, 2021.
- [15] M. Kalntis, and G. Iosifidis, “Energy-Aware Scheduling of Virtualized Base Stations in O-RAN with Online Learning,” in *Proc. of IEEE GLOBECOM*, 2022.
- [16] 3GPP, “LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification,” Technical Specification (TS) 36.331, 2024.
- [17] —, “LTE; General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access,” Technical Specification (TS) 23.401, 2024.
- [18] —, “Consecutive Conditional HO,” R2-1909862, 2019.
- [19] S. B. Iqbal, S. Nadaf, A. Awada, U. Karabulut, P. Schulz, and G. P. Fettweis, “On the Analysis and Optimization of Fast Conditional Handover With Hand Blockage for Mobility,” *IEEE Access*, vol. 11, pp. 30 040–30 056, 2023.
- [20] M. Gupta, S. Chinchali, P. Varkey, and J. G. Andrews, “Forecaster-Aided User Association and Load Balancing in Multi-Band Mobile Networks,” *IEEE Trans. on Wireless Communications*, vol. 23, no. 5, 2024.
- [21] A. Prado, F. Stöckeler, F. Mehmeti, P. Krämer, and W. Kellerer, “Enabling Proportionally-Fair Mobility Management With Reinforcement Learning in 5G Networks,” *IEEE JSAC*, vol. 41, no. 6, 2023.
- [22] M. Gupta, R. M. Dreifuerst, A. Yazdan, P.-H. Huang, S. Kasturia, and J. G. Andrews, “Load Balancing and Handover Optimization in Multi-band Networks using Deep Reinforcement Learning,” in *Proc. of IEEE GLOBECOM*, 2021.
- [23] J. Choi, W.-H. Lee, Y.-H. Kim, J.-H. Lee, and S.-C. Kim, “Throughput Estimation Based Distributed Base Station Selection in Heterogeneous Networks,” *IEEE Trans. on Wireless Comm.*, vol. 14, no. 11, 2015.
- [24] A. Beck, “First-Order Methods in Optimization,” *MOS-SIAM Series on Optimization*, 2017.
- [25] A. Lesage-Landry, J. A. Taylor, and D. S. Callaway, “Online Convex Optimization with Binary Constraints,” *IEEE Trans. on Automatic Control*, vol. 66, no. 12, 2021.
- [26] N. Littlestone and M.K. Warmuth, “The Weighted Majority Algorithm,” *Proc. of IEEE FOCS*, 1989.
- [27] S. Arora, E. Hazan, and S. Kale, “The Multiplicative Weights Update Method: a Meta-algorithm and Applications,” *Theory of Computing*, vol. 8, no. 6, 2012.
- [28] W. G. Madow, “On the Theory of Systematic Sampling,” *The Annals of Mathematical Statistics*, vol. 20, no. 3, 1949.
- [29] M. Kalntis, A. Lutu, J. O. Iglesias, F. A. Kuipers, and G. Iosifidis, “Smooth Handovers via Smoothed Online Learning,” in *Proc. of IEEE INFOCOM*, 2025.
- [30] L. Zhang, S. Lu, and Z.-Hua Zhou, “Adaptive Online Learning in Dynamic Environments,” in *Proc. of NeurIPS*, 2018.
- [31] S. Diamond and S. Boyd, “CVXPY: A Python-embedded modeling language for convex optimization,” *Journal of Machine Learning Research*, vol. 17, no. 83, pp. 1–5, 2016.
- [32] M. A. Bragin and E. L. Tucker, “Surrogate “Level-Based” Lagrangian Relaxation for mixed-integer linear programming,” *Scientific Reports*, vol. 12, no. 1, p. 22417, Dec 2022.
- [33] 3GPP, “Requirements for support of radio resource management,” Technical Specification (TS) 36.133, 2024.
- [34] —, “User Equipment (UE) radio transmission and reception,” Technical Specification (TS) 36.101, 2024.
- [35] S. Alraih, R. Nordin, A. Abu-Samah, I. Shayea, and N. F. Abdullah, “A Survey on Handover Optimization in Beyond 5G Mobile Networks: Challenges and Solutions,” *IEEE Access*, vol. 11, pp. 59 317–59 345, 2023.
- [36] A. Haghrah, M. P. Abdollahi, H. Azarhava, and J. M. Niya, “A survey on the handover management in 5G-NR cellular networks: aspects, approaches and challenges,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2023, no. 1, p. 52, Jun 2023.
- [37] M. I. Saglam, “Conditional Handover for Non-Terrestrial Networks,” in *Proc. of WINCOM*, 2023.
- [38] L. Yang, X. Yang, and Z. Bu, “A Conditional Handover Strategy Based on Trajectory Prediction for High-Speed Terminals in LEO Satellite Networks,” in *ICC Workshops*, 2024.
- [39] J. Stańczak, U. Karabulut, and A. Awada, “Conditional Handover Modelling for Increased Contention Free Resource Use in 5G-Advanced,” in *Proc. of PIMRC*, 2023.
- [40] S. C. Sundararaju, S. Ramamoorthy, D. P. Basavaraj, and V. Phanindhar, “Advanced Conditional Handover in 5G and Beyond Using Q-Learning,” in *Proc. of WCNC*, 2024.
- [41] C. Fiandrino, D. J. Martínez-Villanueva, and J. Widmer, “A study on 5G performance and fast conditional handover for public transit systems,” *Computer Communications*, vol. 209, pp. 499–512, 2023.
- [42] J. Stanczak, U. Karabulut, A. Awada, and P. Spapis, “Signalling-efficient CFRA Resource Updating for Conditional Handover in 5G-Advanced,” in *Proc. of CSCN*, 2023.
- [43] A. Prado, H. Vijayaraghavan, and W. Kellerer, “ECHO: Enhanced Conditional Handover boosted by Trajectory Prediction,” in *Proc. of GLOBECOM*, 2021.
- [44] 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, pp. 54–62, 2020.
- [45] H.-S. Park, H. Kim, C. Lee, and H. Lee, “Mobility Management Paradigm Shift: From Reactive to Proactive Handover Using AI/ML,” *IEEE Network*, vol. 38, no. 2, pp. 18–25, 2024.