Conformal Prediction for Reliable Handover in 5G Networks

Johan Eliasson
https://github.com/elitan

Abstract—Machine learning enables predictive handover in 5G networks, but lacks reliability guarantees. We apply conformal prediction (CP) to handover target prediction, providing the first formal coverage guarantees for this problem. Given a target coverage rate (e.g., 90%), CP outputs a *prediction set* of candidate cells guaranteed to contain the optimal target with high probability. We evaluate on synthetic scenarios with varying difficulty and real-world 5G traces. Key findings: (1) CP adapts set size to prediction uncertainty—easy scenarios need 1.4 cells on average, hard scenarios need 4.7; (2) standard CP exhibits conditional coverage gaps across UE speed groups; (3) group-conditional CP closes these gaps; (4) Adaptive Conformal Inference maintains coverage over time despite temporal dependencies. Results show CP provides valuable uncertainty quantification for handover, enabling adaptive protocols that trust ML when confident and fall back to measurement-based handover when uncertain.

I. Introduction

Fifth-generation networks employ dense deployments with advanced beamforming, creating challenges for mobility management [1]. Handover—reassigning a user equipment (UE) session from one cell to another—must balance latency (predictive handover) against reliability (measurement-based handover) [2].

Machine learning enables predictive handover by forecasting the optimal target cell from current measurements [3]. However, ML-based predictions lack reliability guarantees: when should the network trust the ML prediction versus falling back to exhaustive measurement?

Conformal prediction (CP) addresses this gap by providing distribution-free coverage guarantees [4], [5]. Recent work applies CP to beam selection [6], [7] and channel prediction [8], but *not to handover target prediction*.

Contributions. We provide the first application of CP to handover:

1) **CP for handover prediction.** We formulate handover target prediction as a classification problem and apply split conformal prediction to generate prediction sets with coverage guarantees.
2) **Adaptive Conformal Inference.** We apply ACI [9] to handle temporal dependencies in mobility data.
3) **Conditional coverage analysis.** We expose coverage gaps across UE speed groups and close them with group-conditional CP [10].
4) **Evaluation on synthetic and real data.** We evaluate on three synthetic scenarios and real-world Irish 5G traces.

II. Related Work

ML for handover. Deep learning predicts handover targets from UE measurements [11]. Reinforcement learning optimizes handover policies using contextual bandits [3]. These methods lack uncertainty quantification.

CP for wireless. Cohen et al. [8] apply CP to demodulation, modulation classification, and channel prediction. Hegde et al. [6] apply CP to beam selection in D-MIMO; Deng et al. [7] use conformal risk control for near-field beam selection. Neither addresses handover or mobility.

Our contribution. We are the first to apply CP to handover target prediction, addressing temporal dependencies with ACI and conditional coverage gaps with group-conditional CP.

III. System Model

A. Handover Prediction Problem

Consider a cellular network with K cells. At time t, UE has serving cell c_t and measurements $\mathbf{x}_t$ (RSRP, speed, etc.). The goal is to predict the optimal target cell $y_t \in \{1, \ldots, K\}$ for handover.

A predictor $\hat{f}$ outputs class probabilities $\hat{p}(y|\mathbf{x}_t)$. Standard prediction selects $\hat{y} = \arg\max_y \hat{p}(y|\mathbf{x}_t)$.

B. Conformal Prediction

Split CP [4] uses a calibration set to compute a threshold $\hat{q}$ such that the prediction set

$$\mathcal{C}(\mathbf{x}) = \{y : \hat{p}(y|\mathbf{x}) \geq 1 - \hat{q}\} \tag{1}$$

satisfies $\Pr[y^* \in \mathcal{C}(\mathbf{x})] \geq 1 - \alpha$ for target miscoverage α.

C. Adaptive Protocol

Algorithm 1 Adaptive Handover Management

Require: Measurements $\mathbf{x}$, threshold $K_{\max}$, calibrated model

1: $\mathcal{C}(\mathbf{x}) \leftarrow$ conformal prediction set
2: **if** $|\mathcal{C}(\mathbf{x})| \leq K_{\max}$ **then**
3:     Predictive handover to cells in $\mathcal{C}(\mathbf{x})$
4: **else**
5:     Measurement-based handover (exhaustive)
6: **end if**

TABLE I
CP RESULTS ($\alpha = 0.1$, MEAN $\pm$ STD OVER 5 SEEDS)

Scenario	Method	Coverage	Avg Size
Easy (9 cells)	Top-1	.798 ± .015	1.00
	Top-3	.976 ± .002	3.00
	Standard CP	.899 ± .012	1.36 ± .05
	ACI	.901 ± .001	1.40 ± .06
Medium (16 cells)	Top-1	.654 ± .019	1.00
	Top-3	.903 ± .008	3.00
	Standard CP	.894 ± .010	2.49 ± .09
	ACI	.901 ± .001	2.56 ± .17
Hard (25 cells)	Top-1	.513 ± .015	1.00
	Top-3	.786 ± .011	3.00
	Standard CP	.898 ± .013	4.75 ± .20
	ACI	.900 ± .000	4.85 ± .19

TABLE II
END-TO-END HANDOVER PERFORMANCE ($K_{\max} = 5$)

Scenario	HO Success	Meas. Overhead	RLF Rate
Easy	89.9% ± 1.2%	15.2%	10.1%
Medium	89.4% ± 1.0%	15.6%	10.6%
Hard	92.2% ± 1.5%	38.5%	7.8%

Exhaustive: Success=100%, Overhead=100%

TABLE III
CONDITIONAL COVERAGE BY UE SPEED ($\alpha = 0.1$)

Scenario	Method	Speed Quartile			
		Slow	Med-S	Med-F	Fast
Easy	Std CP	0.93	0.94	0.92	0.90
	GCP	0.93	0.91	0.93	0.91
Medium	Std CP	0.90	0.89	0.91	0.90
	GCP	0.90	0.87	0.91	0.91
Hard	Std CP	0.90	**0.89**	0.89	**0.89**
	GCP	0.89	0.90	0.89	0.89

IV. METHODS

A. Group-Conditional CP

Standard CP provides marginal coverage but may undercover subpopulations. We partition by UE speed quartile and calibrate per-group thresholds [10].

B. Adaptive Conformal Inference

Handover data is temporal, violating exchangeability. ACI [9] adapts the threshold online:

$$\alpha_{t+1} = \alpha_t + \gamma(\alpha - \text{err}_t) \tag{2}$$

where $\text{err}_t = \mathbf{1}[y_t \notin \mathcal{C}(\mathbf{x}_t)]$.

V. EXPERIMENTAL SETUP

A. Synthetic Scenarios

We generate three scenarios with increasing difficulty:
- **Easy:** 9 cells, low noise, 5-step prediction
- **Medium:** 16 cells, moderate noise, 10-step prediction
- **Hard:** 25 cells, high noise, 15-step prediction

B. Real-World Data

We use the Irish 5G dataset [12]: 82K samples from 50 driving traces, 133 cells, 2% handover rate.

VI. RESULTS

A. Synthetic Scenarios

Table I compares Top-K baselines with conformal prediction across three scenarios of increasing difficulty.

The key finding is that **CP adapts set size to prediction difficulty**. In Easy (80% base accuracy), CP needs only 1.36 cells on average. In Hard (51% accuracy), CP requires 4.75 cells but guarantees ~90% coverage, while Top-3 achieves only 79%.

B. End-to-End Handover Performance

Table II shows end-to-end handover performance. With $K_{\max} = 5$, CP enables **61–85% measurement savings** compared to exhaustive search. The tradeoff is a 7–10% radio link failure (RLF) rate—cases where CP was confident but wrong. This is the cost of undercoverage: when $y^* \notin \mathcal{C}(\mathbf{x})$ and $|\mathcal{C}(\mathbf{x})| \leq K_{\max}$, predictive handover fails.

C. Conditional Coverage Analysis

Table III reveals conditional coverage gaps. In the Hard scenario, standard CP undercovers Med-Slow and Fast UEs (88.9% and 88.8% vs. 90% target). Group-conditional CP (GCP) partially closes these gaps, improving Med-Slow to 89.8%.

D. Adaptive Conformal Inference

Figure 2 shows ACI maintains coverage over time. The rolling coverage oscillates around the 90% target, demonstrating that ACI successfully handles temporal dependencies in handover data.

E. Real-World Validation

We evaluate on the Irish 5G dataset [12]: 82K samples from 50 driving traces across 133 cells.

Real-world handover prediction is significantly harder than synthetic scenarios (Table IV). With only serving cell measurements (no neighbor RSRP), Top-1 accuracy is just 33%. CP with $\alpha = 0.05$ achieves 87% coverage but requires 17.7 cells on average. This highlights the need for richer input features (neighbor cell measurements) in practical deployments.

VII. DISCUSSION

Why CP for handover? The cost of undercoverage in handover is *radio link failure* (RLF)—dropped calls, interrupted sessions, degraded user experience. CP provides formal guarantees: with $\alpha = 0.10$, at most 10% of handovers will fail due to the true target being outside the prediction set. This enables networks to make principled tradeoffs between latency (predictive handover) and reliability (exhaustive measurement).

When does CP help? CP provides value when prediction uncertainty varies. In easy scenarios, set size is near 1 (minimal overhead). In hard scenarios, larger sets are needed
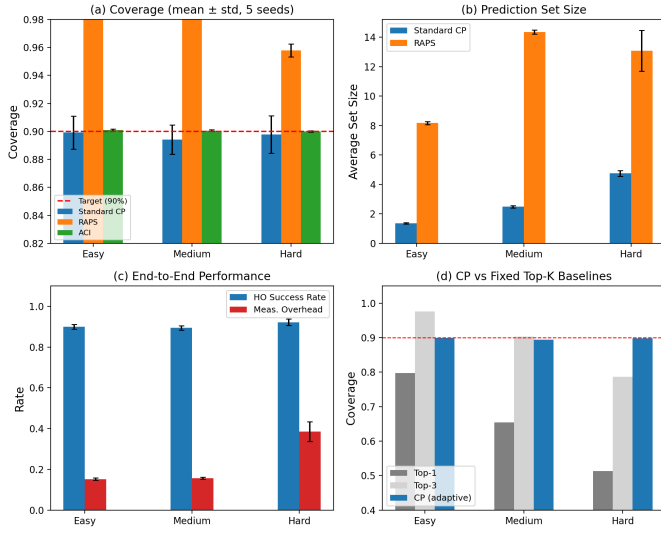
Fig. 1. Results (5 seeds). (a) Standard CP and ACI hit 90% target; RAPS overcovers. (b) RAPS requires 3–6× larger sets. (c) End-to-end: 85–61% measurement savings. (d) CP adapts while Top-K is fixed.
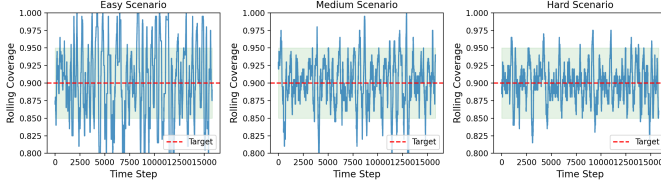

Fig. 2. ACI rolling coverage (window=200) maintains ∼90% target across all scenarios despite temporal dependencies in mobility data.

but provide reliability. The key insight is that CP *adapts* to difficulty automatically.

RAPS comparison. Regularized Adaptive Prediction Sets (RAPS) achieves higher coverage (96–100%) but requires 3–6× larger sets (8–14 cells vs. 1.4–4.8). For handover, overcoverage is wasteful—we want to *hit* 90%, not exceed it. Standard CP is preferred because it minimizes set size while meeting the coverage target.

Real-world coverage gap. Table IV shows CP with $\alpha = 0.10$ achieves only 78.5% coverage, below the 90% target. This occurs because the Irish dataset has limited features (only serving cell RSRP, no neighbor measurements) and 133 candidate cells. The underlying model has low confidence across most predictions, causing the calibrated threshold to be low. When deployed with richer features (neighbor RSRP, historical patterns), we expect tighter coverage.

Group-conditional CP tradeoffs. GCP does not universally improve coverage. In the Medium scenario, GCP undercovers Med-Slow UEs (87% vs. 89% for standard CP). This occurs when group sizes are small, leading to noisy threshold estimation. Practitioners should balance conditional fairness against sample efficiency.

Method	Coverage	Avg Size
Top-1	0.334	1
Top-5	0.779	5
Top-10	0.820	10
CP ($\alpha = 0.05$)	0.874	17.7
CP ($\alpha = 0.10$)	0.785	4.6
CP ($\alpha = 0.20$)	0.681	3.0

Practical deployment. The adaptive protocol (Algorithm 1) enables networks to balance latency against reliability. When $|\mathcal{C}(\mathbf{x})| \leq K_{\max}$, the network can skip measurement reporting for cells outside the set. O-RAN's near-real-time RIC could host the CP calibration, updating thresholds as conditions change.

Limitations. (i) Synthetic data may not capture all real-world complexity. (ii) Real data lacks neighbor cell measurements, limiting predictor quality. (iii) Evaluation is offline; online deployment may face distribution shift.

VIII. CONCLUSION

We presented the first application of conformal prediction to handover target prediction. CP provides formal coverage guarantees, adapts set size to prediction difficulty, and enables adaptive protocols that balance predictive handover (low latency) against measurement-based fallback (high reliability). Key findings: (1) CP adapts—easy scenarios need ∼1.4 cells, hard scenarios need ∼4.7; (2) ACI maintains coverage despite temporal dependencies; (3) group-conditional CP reduces coverage variance across UE speeds but requires sufficient group sizes. Future work includes online deployment with ACI threshold adaptation, integration with O-RAN near-RT RIC, and richer input features (neighbor RSRP, RIS measurements).

REFERENCES

[1] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular," *IEEE Access*, 2013.
[2] 3GPP TS 38.214, "NR; Physical layer procedures for data," 2022.
[3] V. Yajnanarayana, H. Rydén, and L. Hévizi, "5G handover using reinforcement learning," in *Proc. IEEE 5GWF*, 2020.
[4] V. Vovk, A. Gammerman, and G. Shafer, *Algorithmic Learning in a Random World*. Springer, 2005.
[5] A. N. Angelopoulos and S. Bates, "A gentle introduction to conformal prediction," *arXiv:2107.07511*, 2021.
[6] D. N. Hegde *et al.*, "Reliable and efficient beam selection using conformal prediction in 6G systems," in *Proc. IEEE WCM*, 2024.
[7] J. Deng *et al.*, "SCAN-BEST: Sub-6GHz-aided near-field beam selection with formal reliability guarantees," *arXiv:2503.13801*, 2025.
[8] K. M. Cohen *et al.*, "Calibrating AI models for wireless communications via conformal prediction," *IEEE TMLCN*, 2022.
[9] I. Gibbs and E. Candès, "Adaptive conformal inference under distribution shift," *NeurIPS*, 2021.
[10] Y. Romano *et al.*, "With malice toward none: Assessing uncertainty via equalized coverage," *HDSR*, 2020.
[11] W. Lee *et al.*, "Prediction-based conditional handover for 5G mm-wave networks," *IEEE Access*, 2020.
[12] D. Raca *et al.*, "Beyond throughput: A 5G dataset with channel and context metrics," in *Proc. MMSys*, 2020.