Can Online Learning Make Extreme Mobility More Reliable?

Yuanjie Li [1],Esha Datta [2], Jiaxin Ding [3], Ness Shroff [4], Xin Liu [2]
[1]Tsinghua University, [2]University of California, Davis, [3]Shanghai Jiao Tong University, [4]Ohio State University

ABSTRACT

The demand for seamless Internet access under extreme user mobility, such as on high-speed trains and vehicles, has become a norm rather than an exception. However, the 4G/5G mobile network is not always reliable to meet this demand, with non-negligible failures during the handover between base stations. A fundamental challenge is to balance the *exploration* of more measurements for satisfactory handover, and *exploitation* for timely handover before the fast-moving user leaves the serving base station's coverage.

We ask a simple question: *Can the mobile network automatically learn a reliable handover policy for its users in extreme mobility?* Our preliminary study yields a positive sign. We formulate the exploration-exploitation trade-off in extreme mobility as a composition of two online learning problems. Then we showcase how multi-armed bandits help solve both problems with a provable $O(\log J \log T)$ regret. Our theoretical analysis and tests over a real LTE dataset from Chinese high-speed rails confirm the potential of online learning for reliable extreme mobility.

1 INTRODUCTION

The wide adoption of high-speed rail has made extreme user mobility increasingly common. Today, high-speed trains can move up to 350 km/hr with passengers who require always-on Internet access. A common solution is to use cellular networks, such as 4G LTE and 5G. But while the existing 4G/5G network can successfully support billions of stationary or low-mobility users, it struggles to maintain the same reliable service for users with extreme mobility. It has been reported that [? ?], the high speed causes more failures: empirical studies of 4G LTE from high-speed rail shows the handover failure ratio can range from 5.5% to 12.6%, which is about 2× higher than low-mobility scenarios such as walking and driving.

There are various causes of the failures in extreme user mobility, such as weak coverage, dramatic wireless dynamics, signaling loss, sluggish channel feedback, to name a few [?]. To address them, a unique and fundamental challenge for the radio base station (cell) is to tackle the *exploration-exploitation trade-off.* To decide the next target cell a client should migrate to, the serving cell should ask the client to measure the candidate cells' radio quality (Figure 1). With extreme mobility, the serving cell has to balance the need to take more measurements of available cells (exploration) for satisfactory decision, and the demand to make a timely, successful handover (exploitation) before the fast-moving user leaves its coverage (thus losing network service). This is a difficult trade-off that depends on various dynamic factors, such as the user movement speed, base station's runtime operation modes, and the external environment change. A static or manually-crafted handover policy today fall short in adapting the trade-off to these dynamic factors.

To this end, this work takes the first step to explore if online machine learning can help automate the exploration-exploitation trade-off in extreme mobility. We analyze this trade-off in 4G/5G and formulate it as a composition of two distinct problems (§3). For each mobile user, the serving cell should first identify an optimal serving cell threshold value to trigger the handover and measurement procedure. Then, the serving cell should next determine when and in what sequence to take a measurement, and when to execute a handover. This problem decomposition is compatible with the readily-available mechanisms in 4G/5G, thus facilitating the implementations in practice.

We next explore the solution directions to both problems using online learning. While the ultimate solutions are still under investigation, our early attempts have yielded positive results. Specifically, we showcase a policy called `BaTT` (Bandit and Threshold Tuning) based on multi-armed bandits (§4). To determine *when* to start the measurements, we formulate it as a J-armed stochastic bandit problem over T rounds, and solve it with ϵ-Binary-Search-First with $O(\log J \log T)$ regret. Then, to optimize the handover reliability, `BaTT` decides *what sequence* of target cells to measure. This can be formulated as an opportunistic bandit with side observations. We introduce the opportunistic Thompson sampling algorithm to solve this problem with $O(\log T)$ regret. `BaTT` has exhibited viable benefits. Our experiments in §6 with a large-scale LTE dataset on the Chinese high-speed trains show `BaTT` reduces 29.1% handover failures compared to the state-of-the-art 4G/5G handover policies and lower regret than traditional UCB and Thompson sampling. We discuss the remaining open issues and the next steps in §8.

2 PRELIMINARIES

The 4G LTE and 5G are the largest wireless infrastructure that, together with wired Internet, enable ubiquitous Internet access and wide-area mobility management for users. 4G/5G deploys base stations ("cells") in different geographical areas.

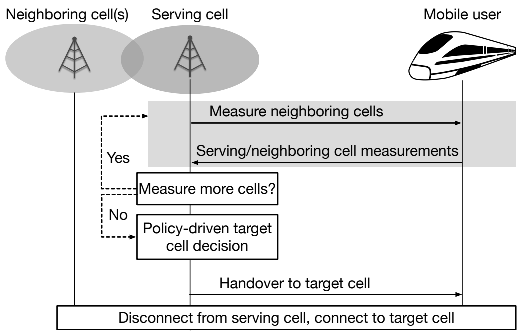


Figure 1: Mobility management in 4G/5G today.

Table 1: Handover failures in extreme mobility

User speed (km/h)	200	300	350
Total handover failures	5.5% (100%)	12.1% (100%)	12.6% (100%)
• **Due to serving cell**	4.9% (90.0%)	9.3% (77.1%)	11.0% (87.3%)
• **Due to target cell**	0.6% (10.0%)	2.8% (22.9%)	1.6% (12.7%)

When a user leaves one cell's radio coverage, it migrates to another cell (a *handover*) to retain its network service.

Figure 1 shows the standard 4G/5G handover procedure [**?** **?**]. When a mobile user connects to a serving cell, it receives a list of neighboring cells. The user can then measure these neighboring cells' signal strengths one by one. If any neighboring cell satisfies the standard triggering criteria (e.g., a neighboring cell's signal strength is offset better than the serving cell's [**? ?**]), the user will report this cell's and serving cell's signal strengths to the serving cell. The serving cell will then run its local policy to decide if more neighboring cells should be measured, whether handover should begin, and which target cell the user should hand over to. If the serving cell chooses to take new measurements, it will provide the user with a new neighboring cell list. If it chooses to handover, the serving cell will (in coordination with the target cell) send the handover command with the target cell's identifier to the user. The user will disconnect from the serving cell and connect to the target cell.

3 PROBLEM AND FORMULATION

3.1 Is 4G/5G Reliable in Extreme Mobility?

The current 4G/5G handover design is primarily meant for static and low-mobility scenarios. Recent studies [**? ?**] have shown that fast-moving users experience non-negligible handover failures, thus frequently losing Internet access. Table 1 shows the 4G LTE handover failure ratios of a smartphone on a Chinese high-speed train from Beijing to Shanghai based on the dataset from [**?**] (elaborated in §6). On average, 5.5%, 12.1% and 12.6% handovers fail at the train speed of 200km/h, 300km/h and 350km/h, respectively. The failure ratio becomes higher with faster train speed.

By analyzing the LTE signaling messages of these failures, we find that 77.1%–90.0% of these failures are caused by the *late handover*, i.e., by the user not receiving the handover command from the serving cell by the time it leaves the serving cell's radio coverage. The remaining handover failures occur when the user receives the handover command from the serving cell, but fails to connect to the new target cell. In this case, the selected target cell is unreliable.

3.2 New Challenge: Exploration-Exploitation Trade-off

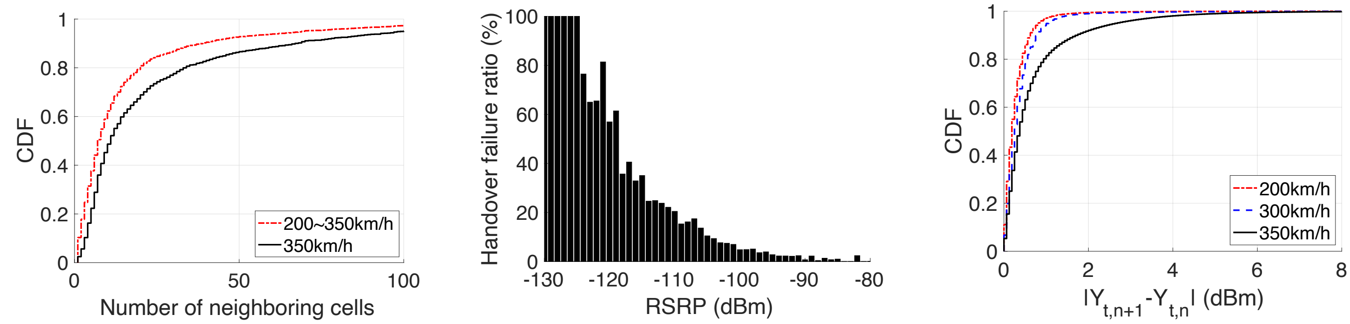
Frequent handover failures occur in extreme mobility because the serving cell faces a fundamental dilemma between *exploration* (more measurements for satisfactory target cell selection) and *exploitation* (fast measurements for timely handover). In 4G/5G, the serving cell relies on the user to measure and report the cells' signal strengths for the handover decision. To retain Internet access, the user must deliver these measurements *before* it leaves the serving cell's radio coverage. But finding a reliable target cell may require scanning and measuring *all* available cells, in principle. Per Figure 2a, on average, a mobile user on a Chinese high-speed train should measure 16 different neighboring cells before making a handover decision. Note that the user has to measure these cells *sequentially*. But, if the user is moving very fast, it may not be able to deliver all its measurements *and* trigger a handover before leaving its serving cell's radio coverage (resulting in a late handover failure). Reducing the number of cells to measure can mitigate late handovers. But this risks missing better target cells and therefore committing a handover to an unreliable target cell (thus failures).

3.3 Problem Formulation

As discussed in §3.2, a fast-moving user has a short but critical time window to conduct effective measurements for handover. It needs to use this period effectively by measuring the right sequence of target cells *before* leaving the serving cell's coverage. For reliable handover, we must answer two questions: 1) When does this critical moment start? 2) what is the right sequence of target cells to measure?

To answer both questions, we formulate the reliable handover problem as follows. Consider a serving cell with K neighbor cells. Given a set of mobile users $t = 1, \ldots, T$, our goal is to minimize the handover failure ratio over all T users.

When should the measurements for handover begin? As a user is moving away from the serving cell, its signal strength weakens. So roughly speaking, the user's critical time starts when the user-perceived signal strength of the serving cell is at a certain threshold. In 4G/5G, this threshold has been standardized (A2 in [**? ?**]) and configurable for each cell by operators. Manually tuning this threshold is a challenging task. On one hand, this threshold should be high enough so that 1) handover failure will not often occur due to weak serving cell; and 2) the user has sufficient time to



(a) Number of neighboring cells **(b) Failures w.r.t. signal strength (350km/h)** **(c) Dynamics of adjacent measurements**

Figure 2: Characteristics of LTE handovers over Chinese high-speed train.

measure neighboring cells to obtain a good target cell for handover. On the other hand, this threshold should also be low enough 1) to avoid the "ping-pong" loops where a user oscillates between two cells due to signal fluctuations (which incurs a lot of signaling overhead and more failures), and 2) to avoid a false start when a desirable target cell is still too far away to be measured appropriately.

To automatically learn this threshold, we formulate this "when" question as a *closest sufficient threshold identification problem*. Given a serving cell, let $\{Z_j\} = Z_1, Z_2, \ldots Z_j, \ldots, Z_{J-1}, Z_J$ be the sequence of J serving cell signal strength observed by a mobile user. Let $[J]$ denote the list $\{1, 2, \ldots, J\}$. Note we only consider *discrete* signal strengths which are standardized in 4G/5G [**? ?**]. Let the random variable $f(Z_j)$ represent handover failure due to the serving cell's signal strength Z_j. Note $f(Z_j) \in \{0, 1\}$, where 0 indicates handover failure and 1 indicates a success. The probability of a handover failure due to signal strength Z_j is $r_j = \mathbb{P}[f(Z_j) = 0]$. Let R be the given serving cell handover failure tolerance level. We assume that as the serving cell's signal strength Z_j increases, the handover failure probability r_j monotonically decreases. This is coherent with the physical laws and empirical results from the high-speed rail dataset in Fig. 2b. Our goal is to find a threshold $M \in \{Z_j\}$ that is the smallest Z_j such that $r_j \leq R$. That is, M is the lowest signal strength at which the probability of handover failure is no larger than R.

What sequence of target cells to measure? Given the threshold M, once a mobile user starts the measurement, the key issue is to decide the sequence of the target cells to measure and the time to stop measurement for handover.

Consider a mobile user t. Once the measurement procedure is triggered, the serving cell starts a sequence of measurements of neighboring cell's signal strengths, indexed by n. The decision of whether to take more measurements or to execute a handover is the central exploitation-exploration dilemma faced by the serving cell. Note that the total number of measurements may vary from user to user (depending on their movement speed). Further, once the handover is decided, the serving cell will terminate measurement.

At the nth measurement, let $I_{t,n}$ be the index of the target cell to measure. In 4G/5G, the user can observe the serving cell signal strength $Y_{t,n}$ and the target cell signal strength $X_{I_{t,n}}$. Let X_{best} be the strongest target cell observed thus far. After n measurements, if the handover is decided, the user will migrate to the best target cell with signal strength X_{best}.

Recall that $f(Y_{t,n}) \in \{0, 1\}$ is the handover failure caused by serving cell signal strength $Y_{t,n}$. Similarly, we can define $g(X)$ as the handover failure caused by the target cell with signal strength X, where $g(\cdot)$ and $f(\cdot)$ may be distinct functions. The handover failure probability of user t is $\mathbb{E}[f(Y_{t,n})g(X_{best})]$ when the handover happens after n measurements and X_{best} is the best target cell. In general, $Y_{t,n}$ decreases with n as the mobile user is moving away from the serving cell. Therefore, the tradeoff is whether to make more measurements, which improves X_{best}, but at the risk of decreasing $Y_{t,n}$. The typical practice today is to measure target cells following a fixed sequence and trigger actual handover when X_{best} is greater than or equal to $Y_{t,n}$ plus an offset quantity determined by the network provider [**? ?**]. The objective of the "what sequence" question is to decide the best order of target cell measurement and when to stop measurement for handover.

4 A SHOWCASE POLICY: BaTT

We showcase the potentials of using online learning to solve both problems above for reliable extreme mobility. Our example design, BaTT, explores the multi-armed bandit algorithms. We choose bandit algorithms as our first candidate solution for three reasons: (1) They are lightweight and thus responsive for fast-moving users; (2) They can provably guarantee reliability (i.e., minimization of the regret of handovers in §5); (3) They are highly adaptive to environmental dynamics, changes of the network configurations, and user movement variations. We are also exploring other online learning algorithms that can achieve these goals and outperform BaTT.

4.1 When: ϵ-Binary-Search-First

Recall that our objective is to find the handover threshold M, i.e., the lowest signal strength at which the probability of serving cell handover failure is no larger than R. Clearly, exploring each value of the J signal strengths is expensive. Instead, we should leverage the monotonicity property between the signal strength and handover failure rate.

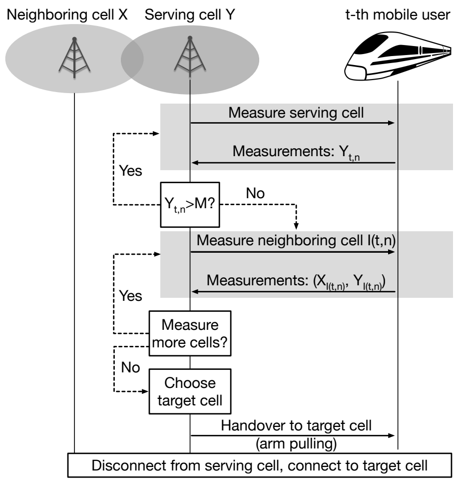


Figure 3: Workflow of BaTT in extreme mobility.

Algorithm 1 ϵ-Binary-Search-First

Input: J, T, R, $0 \le \epsilon \le 1$

Explore: Binary-Arm-Search(J, $\lfloor \frac{\epsilon T}{\log J} \rfloor$,1,$J$,R)
Select arm j such that $\hat{r}_j \ge R$ and $j \in \arg\min_{i \in [J]} |\hat{r}_i - R|$
for remaining rounds $n \le T$ **do**
  Play arm j
end for

Binary-Arm-Search(J, P, R, Start, End)

if End $\ge$ Start **then**
  Play arm $j = \lceil Start + \frac{End - Start}{2} \rceil$ for a total of P times. Denote the empirical mean reward $\hat{r}_j$.
  if $\hat{r}_j \ge R$ **then**
    Return Binary-Arm-Search(J, P, Start, $j-1$, R)
  else
    Return Binary-Arm-Search(J, P, $j+1$, End, R)
  end if
end if

To this end, we propose the ϵ-Binary-Search-First based on multi-armed bandit algorithms. We provide an exploration parameter $0 \le \epsilon \le 1$. Each arm $j \in [J]$ is associated with a random variable $f(Z_j)$ where $\mathbb{E}[f(Z_1)] \le \mathbb{E}[f(Z_2)] \le \cdots \le \mathbb{E}[f(Z_J)]$. The goal is to identify the optimal threshold:

$$M = \arg\min_{Z_j} \left\{ \mathbb{E}[f(Z_j)] \ge R \right\} \tag{1}$$

Alg. 1 shows how ϵ-Binary-Search-First works. It has two phases: exploration and exploitation. During the exploration phase, Alg. 1 pulls the arms in a binary search manner as shown in the Binary-Arm-Search subroutine. Exploration lasts no more than ϵT rounds, where ϵ is optimized as $\epsilon = \frac{\log J}{T} - \frac{\log J}{2T\delta^2} \log\left(\frac{\log J}{6\delta^2 TJ}\right)$. During exploitation, Alg. 1 identifies the estimated best arm among those searched and pulls it for the remainder of the game. We choose an exploration-first policy here for easy illustration. Of course, UCB or Thompson Sampling can also be used with the same asymptotic bound.

4.2 What Sequence: Opportunistic-TS

Consider a mobile user t with a given threshold $\hat{M}$ decided by the ϵ-Binary-Search-First algorithm. Once the handover measurement is triggered by $Y_{t,0} < \hat{M}$, our goal is to determine the optimal sequence of target cells to measure.

Algorithm 2 Opportunistic Thompson Sampling (TS)

Input: $t, K, \hat{M}$, current TS posterior

1: $n = 0, X_{best} = 0, Y_{t,n} = \infty, B = \emptyset$
2: **if** $X_{best} < \hat{M}$ **then**
3:   **if** $Y_{t,n} > X_{best}$ **then**
4:     Measure target cell $I_{t,n}$ using TS, where $I_{t,n} \notin B$.
5:     Receive $(X_{I_{t,n}}, Y_{t,n})$
6:     Update
7:   **else**
8:     Handover to X_{best}
9:   **end if**
10: **else if** $Y_{t,n} \ge \hat{M} + c$ **then** % "free" observation
11:   Measure target cell $I_{t,n}$ using TS where $I_{t,n} \notin B$.
12:   Receive $(X_{I_{t,n}}, Y_{t,n})$
13:   Update
14: **else**
15:   Handover to X_{best}
16: **end if**
17: def **Update**:
18: **if** $X_{I_{t,n}} > X_{best}$ **then**
19:   $X_{best} \leftarrow X_{I_{t,n}}$
20: **end if**
21: $n \leftarrow n+1$, $B = B \cup I_{t,n}$. Update TS posterior distribution of arm $I_{t,n}$

We propose an opportunistic Thompson sampling algorithm motivated by a unique property observed in the real traces. In particular, we observe that the change in signal strength over consecutive measurement is bounded (as empirically validated in Fig.2c with the high-speed rails dataset). This allows us to make the following *regularity assumption* regarding the serving cell signal strength $Y_{t,n}$. We assume that there exists some positive constant c such that

$$|Y_{t,n} - Y_{t,n+1}| < c.$$

This ensures that the serving cell signal strength does not change "too quickly" between consecutive measurements. Under this assumption, we can have "free" measurements when the best target cell so far is good enough ($X_{best} \ge \hat{M}$) and the serving cell is still strong enough ($Y_{t,n} \ge M + c$). So the next measurement is risk-free. Therefore, we can first find the best target cell and then use the "free" observations when available to satisfy the need for exploration at no cost.

The algorithm is outlined in Alg. 2. It takes as input the number of available neighboring cells K, the index of the mobile user t, and the handover threshold $\hat{M}$ from ϵ-Binary-Search-First. If the best target cell X_{best} is not satisfactory (that is, $X_{best} < \hat{M}$) then the algorithm compares the serving cell to X_{best} (Line 2). If $Y_{t,n} > X_{best}$, then the algorithm continues to measure the best unmeasured target selected using TS (Thompson Sampling). If $Y_{t,n} < X_{best}$, the mobile user t handovers to X_{best} (line 7). Otherwise, as in line 9, if X_{best} is satisfactory and $Y_{t,n} \ge \hat{M} + c$, then the algorithm can make "free" measurements. In this case, the algorithm explores by selecting an unmeasured target selected using TS. We note that any bandit algorithms can be used in selecting target cells (to measure), such as UCB, greedy, and round-robin. In our experiments, we observe that TS achieves the best and most robust performance and thus adopt it here.

5 REGRET ANALYSIS

We analyze the regrets of reliable handovers with BaTT.

5.1 ϵ-Binary-Search-First

To analyze Alg. 1, we first define its regret. Let $N_T(j)$ be the number of times a threshold setting j is pulled under a given policy Γ. We define the regret over T rounds as

$$R_\Gamma(T) = T - \mathbb{E}[N_T(a^*)] = \sum_{a \neq a^*} \mathbb{E}[N_T(a)].$$

We can now bound the regret accumulated by the ϵ-Binary-Search-First algorithm. Define $\Delta = r_M - R$, $D = \min_j |r_M - r_j|$, and $d = \min_j |r_j - R|$, and $\delta = \min(\Delta, D/2)$, where r_M is the probability associated with the signal strength M.

Theorem 5.1. *Alg. 1 achieves regret bounded by*

$$R(T) \leq \log J \left(\frac{\log 6\delta^2 T J}{2\delta^2} - \frac{\log \log J}{2\delta^2} + \frac{1}{2\delta^2} + 1 \right)$$

when $d < \sqrt{\frac{\log(T \log J)}{2P}}$, where d is the minimum absolute distance between a searched arm and R, and $\delta = \min(\Delta, D/2)$.

The proof is available in [?].

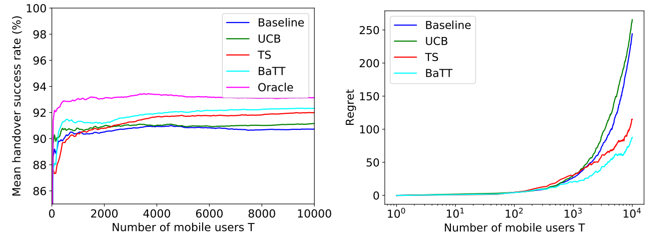
5.2 Opportunistic-TS

We note that the exact regret of the general "what sequence" is difficult to evaluate, even if we use a classic bandit algorithms such as TS and UCB. The reason is that $Y_{t,n}$ is an unknown and non-stationary process over n. However, a simpler case is where the mobile user is only allowed to measure one target cell and then handover to it. In this case, Opportunistic-TS reduces to the classic Thompson Sampling and thus yields $O(\log T)$ regret. We note the ability to select among multiple target cells in general should yield better performance than classic TS in practice. This is supported by our empirical evaluation using real traces. We are in the progress of extending this analysis to the general settings. Specifically, we hope to achieve $O(1)$ regret as in [?] because the probability of having "free" observations is non-negligible.

6 EMPIRICAL VALIDATIONS

We assess the potentials of online learning for reliable extreme mobility. The results here are based on BaTT, thus representing a *lower bound* of handover failure reductions.

Dataset: We use a large-scale 4G LTE dataset on Chinese high-speed trains from [?]. This dataset was collected on the high-speed rails between Beijing and Shanghai over 135,719 km of trips. In these tests, a smartphone using China Mobile or China Telecom 4G LTE runs continuous iperf data transfer on the high-speed train at 200–350 km/h. Meanwhile, the smartphone runs MobileInsight [?] to collect the 4G LTE signaling messages from the hardware modem. These messages include 38,646 runtime measurement configurations of neighboring cell lists and thresholds, 81,575 measurement reports of serving/neighboring cell's signal strengths, and 23,779 handover commands as exemplified in Figure 2.



(a) Reliability with M=-120 dBm (b) Regret with M=-120 dBm

Figure 4: BaTT and other algorithms with c=4.

Experimental setup: We first evaluate BaTT with a given threshold M. We then compare its regret and handover success rate with the following algorithms using the same threshold: (1) **Oracle**: We assume that the average handover failure rates of the target cells are known. Therefore, we measure the target cells in an increasing order of the failure rate. (2) **Baseline**: This is the state-of-the-art 4G/5G handover algorithm [? ? ?]. It compares the serving cell and target cell's signal strengths and selects the first neighboring cell with $X_{best} > Y_{t,n}$ as the target cell. This policy does not specify the ordering of cells to measure and relies on the device-specific cell scanning implementations instead. So, we assume the user's device measures the target cells randomly. (3) **UCB**: We assume that the serving cell maintains UCB estimates for the target cells and instructs a mobile user to measure target cells based these estimates. (4) **TS**: Similar to UCB, except using Thompson sampling. We note that the evaluation results are driven by real traces and not limited by the assumptions made for the analysis.

6.1 Preliminary Results

We consider a serving cell with $K = 10$ neighboring cells with the empirical signal strength distribution drawn from the dataset. Then we draw each cell's expected reward by mapping their signal strength distribution to the handover success rate based on $f(\cdot)$ and $g(\cdot)$ from real traces in Figure 2b. This results in the reward (handover success rate) vector [0.76, 0.88, 0.90, 0.91, 0.92, 0.93, 0.94, 0.95, 0.97]. We then replay all sequences of serving cell's measurements before each handover command in the dataset. For each serving cell's measurement, we run each algorithm to decide the next neighboring cell to measure. We generate this neighboring cell's measurements based on its empirical distribution of signal strengths. With these measurements, each algorithm decides if the measurement should continue and selects the target cell if handover should start.

Comparison with state-of-the-art As shown in Figure 4a, compared to the baseline in 4G/5G today, BaTT improves mean handover success rate from 89.7% to 92.7%. That is, it prevents 29.1% handover failures in 4G/5G today. Further, note that BaTT approximates the Oracle, which is the optimal performance we can expect in reliable mobility today. Compared to the baseline, BaTT optimizes the ordering of the cells to measure when the serving cell's quality is decreasing, thus mitigating late handover failures.

Comparison with other bandit algorithms Figure 4a and 4b show BaTT outperforms UCB and TS. This is because BaTT adaptively balances the exploration and exploitation based on the runtime serving cell quality, while UCB and Thompson sampling do not. BaTT can accelerate the exploration when the serving cell quality is good and mitigate late handover failures when serving cell quality is not. This is crucial, since late handovers due to the serving cells dominate the handover failures in reality, as shown in Table 1.

7 RELATED WORK

Mobility management has been actively studied in recent years. Various deficiencies have been identified, such as suboptimal radio coverage [?], network misconfiguration [?], handover policy conflicts [?], and late/blind handovers [?], to name a few. Our work studies a different aspect of handover reliability in extreme mobility. In the context of extreme mobility, [? ?] report the non-negligible handover failures in reality, and [?] mitigates them by refining wireless communication to relax the exploration-exploitation trade-off *implicitly*. In contrast, our work *explicitly* tackles the exploration-exploitation trade-off using online learning policies. There are some efforts to refine the performance of handover policy with machine learning techniques like fuzzy logic [?], neural networks [?], and SVM [?]. Our work differs from them because we focus on reliability.

8 DISCUSSION AND CONCLUSION

In this work, we strive for reliable 4G/5G handover in extreme mobility using online learning techniques. We formulate and decompose the exploration-exploitation dilemma in extreme mobility into two online learning problems. To demonstrate the promise of online learning for reliable handovers, we showcase a multi-armed bandit-based strategy to search for the optimal threshold of signal strength to address this dilemma *and* opportunistically balance the exploration and exploitation of target cells based on the runtime serving cell's signal strength. Our analysis shows $O(\log J \log T)$ overall regret of handover failures. Experiments with large-scale operational LTE datasets from the Chinese high-speed trains demonstrate the viability of handover failure reduction.

This work is still in its early stages and demonstrates promising potentials. As a first attempt, many open issues remain. First, the current problem formulation mainly takes a single base station's perspective, i.e., minimize the handover failures among all clients served by one base station. Alternatively, a client-side problem formulation is also possible, i.e., minimize the handover failures of a single client during its fast movement across multiple base stations. We plan to explore if more insights and efficient solutions from this perspective. Second, BaTT currently relies on certain assumptions, such as the time-invariant distribution of each cell's signal strengths and the monotonicity of handover failures w.r.t. signal strengths. Relaxing these assumptions would broaden BaTT's applicability to operational mobile networks in reality. Third, the current BaTT can further explore the domain-specific knowledge of the 4G/5G handovers for further potentials for failure reduction. It could also benefit from recent advances of multi-armed bandit algorithms with $O(1)$ regret (e.g., opportunistic bandits [?]). Fourth, beyond multi-armed bandits, other online learning algorithms may also be beneficial for reliable handovers. Last but not least, our study so far relies on existing mechanisms in 4G/5G today. We will explore if other online learning techniques could be more beneficial with new mobility management scheme designs in 5G and beyond.

REFERENCES