Instability in Distributed Mobility Management

Revisiting Configuration Management in 3G/4G Mobile Networks

Yuanjie Li†, Haotian Deng‡, Jiayao Li†, Chunyi Peng‡, Songwu Lu†
†University of California, Los Angeles, ⋆The Ohio State University
yuanjie.li@cs.ucla.edu, deng.264@buckeyemail.osu.edu, likayo@ucla.edu,
chunyi@cse.ohio-state.edu, slu@cs.ucla.edu

ABSTRACT

Mobility support is critical to offering seamless data service to mobile devices in 3G/4G cellular networks. To accommodate policy requests by users and carriers, micro-mobility management scheme among cells (*i.e.*, handoff) is designated to be configurable. Each cell and mobile device can configure or even customize its own handoff procedure. In this paper, we examine the handoff misconfiguration issues in 3G/4G networks. We show that they may incur handoff instability in the form of persistent loops, where the device oscillates between cells even without radio-link and location changes. Such instability is mainly triggered by uncoordinated parameter configurations and inconsistent decision logic in the handoff procedure. It can degrade user data performance, incur excessive signaling overhead, and violate network's expected handoff goals. We derive the instability conditions, and validate them on two major US mobile carrier networks. We further design a software tool for automatic loop detection, and run it over operational networks. We discuss possible fixes to such uncoordinated configurations among devices and cells.

1. INTRODUCTION

The 3G/4G cellular network is the only large-scale infrastructure that offers "anytime, anywhere" mobility support to smartphones and tablets in reality. The key lies in its micro-mobility management scheme, which *determines* the serving cell (also known as base station)[1] and migrates the mobile device from the current cell to the next neighboring one. This procedure is called handoff, the fine-grained mobility scheme in practice.

In this paper, we study how configurations on the management plane affects the handoff behaviors of mobile devices. At first glance, this micro-mobility support scheme seems to warrant no further research; it has been operated for many years and extensive studies have been documented. However, our fresh perspective from the management-plane configurations yields some interesting, yet surprising results.

[1]Each base station may manage multiple cells (antennas), each of which covers a geographical area. In this paper, we use cells and base stations interchangeably, for a slight abuse of notations.

SIGMETRICS '16, June 14-18, 2016, Antibes Juan-Les-Pins, France

DOI: http://dx.doi.org/10.1145/2896377.2901457

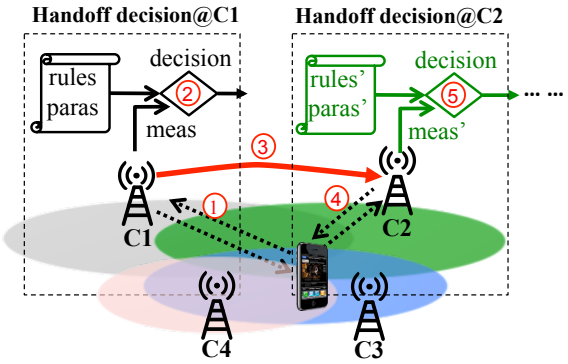


Figure 1: The distributed handoff is configurable.

In a nutshell, the handoff process is distributed in nature, as illustrated in Figure 1. There is no central point which collects all the information and makes decision. Instead, each decision is made locally at a cell or by the mobile device. The target cell is selected by the handoff decision made by the current serving cell or the mobile device. The operation has three components: the local decision logic (rules), tunable parameters and runtime measurements. The decision logic takes both parameters and measurements as inputs (①), and selects the next cell (②). Once the handoff is executed (③), it switches to a new serving cell and starts another handoff decision iteration (④ and ⑤). Note that, both the decision logic and parameters are configurable, in order to meet diverse requirements, such as selecting the best radio quality, letting operators specify their preferences, *etc.*. Moreover, coexistence of heterogeneous technologies (*e.g.*, 3G, 4G LTE, LTE-advanced, small cells) further results in diverse handoff configurations.

In this work, we conduct a systematic study on handoff configurations. We show that, uncoordinated configurations among cells can lead to handoff instability: the mobile device oscillates between a set of cells covering the same area, even when no radio-link or location change is detected. Load balancing is also not the main concern. Instead, uncoordinated parameter settings and loop-prone decision logic among devices and cells are the key drivers to instability. The concrete cases of such instability are covered in five distinctive categories to be reported in §5 and §6.

In each of the five categories, our effort starts from a formal analysis on the handoff (in)stability conditions. While we understand many parameters could be involved, our highlight is not on the qualitative or quantitative impact of each parameter by processing large traces from operational networks. Instead, our focus is on the fundamental understanding from the perspective of structural properties. We thus derive instability conditions that exhibit persis-

tent loops for various forms of handoff misconfigurations. To this end, our modeling setting strives to be as simple as possible, if not overly simplistic in some cases, while still capturing the essence and neglecting secondary details.

We then validate the existence of handoff loops, as well as their impacts, on two US mobile carrier networks. Such empirical assessments are conducted on each problematic category discovered in modeling and analysis. We further design a software tool for automatic loop detection, which applies domain-specific knowledge to check for configuration conflicts in real networks.

Our results show that, handoff instability is not uncommon in reality. We have found 21 instances in the tested scenarios. Instability may occur among homogeneous cells (within 3G or 4G), and among heterogeneous cells (3G/4G/Femtocell). The conflicts can occur not only within existing handoff policies, but also upon configuration updates by carriers. Both configurations and decision logic, from both the network and the user, could result in persistent loops. Certain policy conflicts are even rooted in the problematic network mechanism design (*e.g.*, radio connection control). We notice that persistent loops may be less common than those transient ones (*e.g.*, ping-pong effects), which oscillate between cells due to frequent movement and radio signal strength fluctuations [30, 31]. However, such persistent loops *could* be fully avoided while transient loops cannot due to the nature of the environment dynamics. To this end, we devise loop detection and resolution solutions on both the device and network side, and validate their effectiveness. To our knowledge, this is the first work to report that instability may exhibit due to uncoordinated configurations (among cells and devices) on the management plane of 3G/4G mobile networks.

The rest of the paper is organized as follows. §2 introduces background on handoff configurations and its formulation as a distributed decision process. §3 describes our methodology. §4 gives a brief overview of findings. Using both analysis and empirical study, §5 and §6 describe five categories of instability due to uncoordinated parameters and loop-prone logic, respectively. §7 reports our design of an automatic detection tool and §8 discusses the possible changes on both the carrier and the device to eliminate the loops. §9 compares with the related work and §10 concludes the paper.

2. HANDOFF MANAGEMENT

The 3G/4G network is the largest wireless infrastructure deployed in reality. Its architecture has two main components: radio access network (provisioned by base stations) and core network. The geographic area served by a base station (BS) is called a cell, denoting the coverage of radio access to devices in proximity. At a given location, a device is usually covered by multiple, possibly overlapping cells to ensure seamless service.

In a larger area with many cells, the 3G/4G network offers wide-area mobility support to roaming devices with a two-tier structure. The base stations provide the fine-grained roaming support by handoff, which enables the device to switch its serving cell/base station as it moves. At the coarse granularity, base stations in a geographic area are grouped and managed by a central controller (i.e., mobile switching center (MSC) in 3G, and mobility management entity (MME) in 4G). The controller permits roaming between areas, and further tracks user locations, manages voice/data service contexts and configures forwarding paths.

2.1 Configuration Management for Handoff

Micro-mobility management in current 3G/4G networks is *configurable*. Three components work in concert to make each handoff decision: the *decision logic*, the *tunable parameters* and the *runtime observations (i.e., measurements)*. The decision logic takes both parameters and observations as inputs, and selects the next cell. Tunable parameters specify what kinds of metrics are of interest to the device and the operator. Runtime observations collect latest measurements, thus capturing the dynamic network state.

There are two types of handoffs in 3G/4G networks: (1) *Idle-state handoff*: It is performed by the mobile device, when the device is at the idle state (without ongoing traffic) and has no active connection to the serving cell. This is to prepare the device for network access at any time. (2) *Active-state handoff*: It is initiated by the serving cell, when the device has established the radio connection to the network. We next elaborate on their differences in three components.

∘ **Decision Logics.** This is the algorithm to choose the target cell. The idle-state handoff logic is standardized in 3GPP specifications [5, 10]. Its exact form will be described in §2.2. In contrast, the active-state handoff logic is customizable by the networks.

∘ **Configurable parameters.** They are used by the decision logic. For idle-state handoff, two types of parameters are used: the cell preference, and the radio assessment thresholds. Table 1 summarizes the parameter notations, which are abstracted from actual configurations in operational networks. For active-state handoff, it can further customize its parameter set. Both idle and active state handoffs' parameters are configured by the network, and distributed to the mobile device through broadcast (for idle-state handoff [10]) or dedicated signaling channel (for active-state handoff [6]).

∘ **Runtime observations.** They are usually on the dynamic radio qualities measured by the device, and serve as inputs to the handoff decision logic. The device collects and reports such observations to the handoff decision logic for the decision-making. The idle-state handoff accepts cell radio quality assessments as input, while the active-state can use both the radio quality and customizable observations (*e.g.*, cell load). In practice, to tolerate signal strength oscillations, these observation metrics are typically pre-processed before handoff decision. For example, the received signal strengths used in handoff have been averaged to filter out noises and transients [5, 10]. To stay focused, we assume the observations remain unchanged during each handoff decision process. In reality, we find that this usually holds at the same location, since the handoff decision making is faster than signal strength fluctuations (§7.3).

In summary, handoff configuration specifies key parameters and observation requirements for its decision logic. The parameters span both the user device and the network. Such parameters for the idle-state handoff have been standardized in [10], whereas those for active-state handoff are not standardized and carriers have freedom to customize them. In this work, we do not study abnormal factors, such as inaccurate observations, failure report, etc., but focus on common-case factors. To highlight our findings, we mainly focus on cell preferences and radio quality evaluation thresholds in our analysis and experiments.

2.2 Configurations in Idle-State Handoff

In this section, we use idle-state handoff as an example to present the configurable parameters, decision logic and runtime observations. We denote a handoff execution as $s \xrightarrow{\Omega_s(G_s,O_s)} t$, where s is the serving cell, and t is the target cell selected from candidate cells. Given the serving cell s, Ω_s, G_s and O_s denote the handoff decision logic, tunable parameters and runtime observations, respectively. If the serving cell does not exist (*e.g.*, upon device power-on), we have $s = \emptyset$ and the decision is made by the device.

Figure 2 shows the standardized decision logic for idle-state handoff [3,6]. The decision logic chooses the target cell by comparing the serving cell with each candidate. The runtime observation

Symbol	Description
γ_c	Received signal strength of cell c
$P_{s,c}$	Idle-state preference of cell c at cell s
Θ_s^{serv}	Idle-state threshold of γ_s when s is the serving cell
$\Theta_{s,c}^{low}$	Idle-state threshold of γ_c when s is serving and $P_{s,c} < P_{s,s}$
$\Theta_{s,c}^{eq}$	Idle-state threshold of γ_c when s is serving and $P_{s,s} = P_{s,c}$
$\Theta_{s,c}^{high}$	Idle-state threshold of γ_c when s is serving and $P_{s,c} > P_{s,s}$
$\Theta_{s,c}$	Active-state threshold of γ_c when s is serving with absolute comparison
$\Theta 1_s$	Active-state threshold of γ_s when s is serving with indirect relative comparison
$\Theta 2_{s,c}$	Active-state threshold of γ_c when s is serving with indirect relative comparison
$\Theta_{s,c}$	Active-state threshold of γ_c when s is serving with direct relative comparison

Table 1: Notations.

Idle-state handoff

Input: serving cell s, neighboring cell list C, radio measurements γ for each neighboring cell

Output: target cell t

Step1: initialize candidate cell list $L \leftarrow [\,]$

Step2: pairwise cell comparison

for each cell c **in** C:

if $P_{s,c} > Ps, s$ **and** $\gamma_c > \Theta_{s,c}^{high}$

L.append(c)

elif $P_{s,c} = Ps, s$ **and** $\gamma_c > \gamma_s + \Theta_{s,c}^{eq}$

L.append(c)

elif $P_{s,c} < Ps, s$ **and** $\gamma_s < \Theta_s^{serv}$ **and** $\gamma_c > \Theta_{s,c}^{low}$

L.append(c)

Step3: target cell decision

$$t = \begin{cases} s & \text{if } L \text{ is empty} \\ c & \text{if } c \text{ is the cell in } L \text{ with highest preference } P_{s,c} \end{cases}$$

Figure 2: Idle-state handoff decision logic.

is the received signal strength γ_c from each candidate cell c, measured by the user device. For each candidate cell c, the serving cell s defines two types of configurable parameters: the preference level ($P_{s,c}$) concerning a candidate cell c and a series of signal strength thresholds ($\Theta_s^{serv}, \Theta_{s,c}^{low}, \Theta_{s,c}^{eq}, \Theta_{s,c}^{high}$) that help Ω_s to make a decision. Note that both parameters are needed. Radio signal strength is directly related to wireless transmission performance, as well as the cell type (3G, 4G, macro-cells, or femtocells). The cell preference reflects the precedence of cell types from the perspective of the carrier or the user or both. It supplies a flexible mechanism for the device/network to adjust the priorities.

Specifically, each cell is evaluated with the pre-configured preference and the runtime received signal strength. A target cell is chosen when (1) it is more preferred than the serving cell, and its signal strength is higher than a threshold; or (2) it is equally preferred to the serving cell, and its signal strength is offset higher than the serving cell's, or (3) it is less preferred than the serving cell, but the serving cell's signal strength is lower than a threshold, while the target cell's signal strength is higher than another threshold. If more than one cell outperforms the serving cell, the one with highest preference could be chosen. If a tie exists, the signal strength is used to break the tie.

3. METHODOLOGY

We take a two-step approach to studying the instability of the configured handoff. We first use a modeling framework to derive (in)stability conditions. We then run validation experiments to detect loops, uncover the root causes, and quantify the impact in operational 3G/4G networks.

Modeling and Analysis. Our modeling generally follows a discrete-event style. Each handoff is abstracted as a transition from serving cell s to target cell t. The handoff execution for $s \rightarrow t$ is acted on the serving cell s and the mobile device. So after handoff to a new cell, the execution would change. Consecutive handoffs may occur even with the same observations (*e.g.*, no location/radio condition change). Stability is ensured, if for any invariant observation, a device initially associated with any cell s will always converge to the target t but not move to other cells, *i.e.*

$$\mathbf{s} \rightarrow c_1 \rightarrow c_2 \rightarrow \cdots \rightarrow c_k \rightarrow \mathbf{t}$$

If this property is violated, a *persistent loop* can happen between a set of cells even for some unchanged measurements (or measurements fluctuatiing within a small range, see §5–6). The user would experience data/voice performance degradation. The carrier cannot achieve the designated handoff goal, and may suffer from excessive signaling overhead. The handoff is assumed to always succeed without failures (*e.g.*, we ignore radio-link outage).

Empirical Validation. We next conduct experiments in two metropolitan areas from both west and east coasts over two top-tier U.S. carriers: US-I and US-II. The goal is to (in)validate the existence of each handoff loop and quantify its negative impact. The validation takes two steps. First, we develop a loop detection tool (§7) to check the carriers' handoff policies. The detection algorithm is based on our analytical (in)stability results. It reports the stability violations and its condition for runtime observations. Second, for each stability violation, we conduct validation experiments to test its existence, and quantify its negative impact if it exists.

We run both outdoor and indoor experiments. The outdoor experiments cover 63 different locations over 240 km^2 in the west coast and 260 km^2 in the east coast. Each location is selected by at least 2km apart, in order to obtain different cell coverage. We also collect information on indoor experiments at 50 spots in two 8-floor office buildings and one apartment. In the indoor settings, we mainly collect the radio quality observations at various spots, since most cells, as well as their configurations, are similar across locations. We also deploy four 3G Femtocells in office and at home for indoor tests. We use four Android phone models: Samsung Galaxy S4, S5 and Note 3, and LG Optimus G.

4. OVERVIEW OF FINDINGS

In this work, we classify the instability (IS) cases based on the *causes* of the configuration conflicts. Broadly speaking, there are two classes: *parameter misconfiguration*, and *loop-prone decision logic*. Figure 3 exemplifies more concrete subcategories, and Table 2 summarizes our main findings from operational networks.

uncoordinated parameter configuration (§5). In this category, the instability is observed when cells' tunable parameters are not well coordinated. Such misconfigurations can happen within idle-state handoff and active-state handoff (§5.1 and §5.2). It can also happen between idle and active-state handoffs (§5.3). For idle-state handoff, note that identical decision logic is used among all devices at each location [5, 10]. The uncoordinated parameter configuration is the *only* cause of the persistent loops (proved in §5.2).

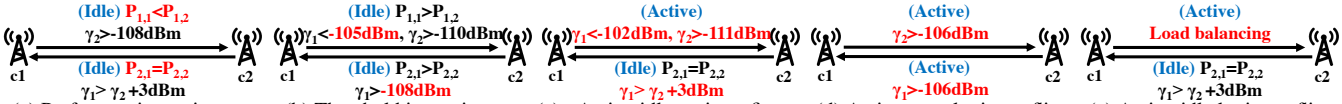
loop-prone decision logics (§6). The instability also occurs when different cells apply conflicting decision logics. No matter how well parameters are fine tuned, conflicts always exist between decision logics. The fundamental reason is that, the active-state logic is customizable at each cell and the current standards do not mandate the same decision algorithm. Specifically, the conflicts can happen between active-state logic engines (§6.1), and between active and idle-state decision logics (§6.2). However, no conflicts exist between idle-state engines since they follow identical, standardized algorithms.

5. INSTABILITY BY UNCOORDINATED PARAMETER CONFIGURATION

We have discovered three categories in this class of instability, all of which are caused by uncoordinated parameter configurations.

Category	Loop type	Carrier	Cause of configuration conflicts	Impact
Uncoordinated parameter configurations (§5)	4G-Femtocell-3G	US-I	Conflicting demands between offloading to private Femtocell and high-speed data service to user.	Data drop/delay; excessive signaling overhead; no offloading or high-speed service achieved
	4G-Femtocell-2G-3G	US-I	The mobile device's service-recovery handoff improperly configures high preference to 2G.	No voice service; data drop/delay; excessive signaling overhead; no offloading or high-speed service achieved
	4G-4G	US-I	Partial policy update with incremental 4G infrastructure upgrade.	Failure of user migration to new cells
	Femtocell-3G	US-I	Isolated Femtocell applies aggressive handoff.	Data/voice disruption
Loop-prone decision logic (§6)	4G-4G	US-I	Uncoordinated load balancing between cells.	Load balancing failure
	3G-3G	US-I,US-II	Radio-agnostic connection control by design.	Data delay/drop; offloading failure
	Femtocell-3G	US-I	Radio-agnostic connection control by design.	Data delay/drop; offloading failure

Table 2: Summary of persistent handoff loops.



(a) Preference inconsistency (b) Threshold inconsistency (c) Active-idle misconfiguration (d) Active-state logic conflict (e) Active-idle logic conflict

Figure 3: Instability with uncoordinated parameter configuration (a,b,c) and loop-prone decision logic (d,e).

They are illustrated in Figures 3a-3c. We first derive the conditions for parameters and runtime observations that would (not) trigger instability. Our analysis aims to answer two questions: (1) Under which parameter configurations, there *exists* some runtime observations that can trigger instability? (2) Given improper parameter configurations, which runtime observation values will eventually trigger instability? Then we conduct empirical assessments to validate their instances in operational 3G/4G networks.

5.1 Inconsistent Preference Values

In this category, each cell locally configures its preference, but these preferences are not globally coordinated. Figure 3a illustrates a simple two-cell case. In this setting, c_1 configures c_2 to be more preferred to c_1 itself, but c_2 assigns equal preference to both cells. The persistent loop happens if the signal strength satisfies $\gamma_2 > \Theta_{2,1}^{high}(-108\text{dBm})$, and $\gamma_1 > \gamma_2 + \Theta_{2,1}^{high}(3\text{dBm})$. Note that, this loop can occur for *any* threshold settings (in the achievable range).

5.1.1 Deriving Instability Conditions

We first derive the instability conditions for this category. Recall that, in the idle-state handoff (Figure 2), when associated with cell s, the mobile device evaluates each candidate c with the preconfigured preference $P_{s,c}$, and its runtime signal strength γ_c. Each cell c is compared with serving cell s, and would be selected if (1) it is more preferred than the serving cell, and its signal strength is higher than a threshold ($(\gamma_c > \Theta_{s,c}^{high})$); (2) it is equally preferred, and its signal strength is offset higher than the serving cell's ($\gamma_c > \gamma_s + \Theta_{s,c}^{eqn}$), or (3) it is less preferred, but the serving cell's strength is weak ($\gamma_s < \Theta_s^{serv}$), *and* the target cell's signal strength is satisfying ($\gamma_c > \Theta_{s,c}^{low}$). If more than one cell satisfies above condition, the one with highest preference could be chosen.

The following result shows that, persistent loop can be caused by improper configurations of preference values. The good news is that, such persistent loops can be eliminated, when the derived preference conditions are avoided (the proof is in Appendix A):

Proposition 1. *Consider n cells $c_1, c_2, ..., c_n$ that use idle-state handoffs only. A loop $c_1 \rightarrow c_2 \rightarrow ... \rightarrow c_n \rightarrow c_1$ can always happen, if and only if their preference settings satisfy: (1) at least one cell c_i sets $P_{i,i} < P_{i,i+1}$, and (2) every cell c_j sets $P_{j,j} \leq P_{j,j+1}$, and $P_{n,n} \leq P_{n,1}$.* □

Two results follow from Proposition 1. First, some preference settings would *always* trigger persistent loops with some runtime observations. For stability, they should always be avoided. Second, with consistent preference configuration, the idle-state decision logic can always ensure stability for a device. This serves as the foundation for stability analysis on other forms of handoff logic. As we will see in §5.2, there exists *pairwise* coordination methods for loop freedom. Enumerating all possible loops is not needed.

5.1.2 Empirical Validation

In this category, we have been able to identify 17 instances that can cause loops in US-I, using the detection tool to be described in Section 7. These configuration conflicts can happen at the same areas (with different runtime observations), and are reported in all locations. Figure 4 summarizes these loops. The smallest loop involves 3 cells, while the largest one includes 7 cells. Among these loops, 16 out of 17 conflicts would occur with Femtocell deployed, while the remaining one can occur without Femtocell. Our outdoor tests first show that all 2G/3G/4G Macrocells have the problematic configurations, and 61 out of 63 locations (96.8%) have all Macrocells deployed. This implies that a potential loop would exist if a Femtocell were deployed at the spot. We further deploy a Femtocell in a campus building, and conduct indoor experiments in all viable locations. In that floor, 25% of the testing locations satisfy both configuration and signal strength conditions, thus triggering loops. For US-II, we do not observe loops in this category. Based on the causes of the preference conflicts, the loops found in US-I can be further classified in three categories:

- **L1: uncoordinated handoff goals.** In this category, 8 variants of loops are reported, all happening between 4G Macrocell, Femtocell and 3G Macrocells. These loops are caused by preference settings for conflicting goals. The 4G Macrocells intend to offload user to his/her private Femtocells, so it assigns Femtocell with highest preference (over 4G/3G Macrocells). The 3G Femtocell has equal preference to all 3G/4G cells. But 3G Macrocells prefers to move the user to high-speed 4G network, so it assigns 3G Femtocell lower preference to 4G. This violates Proposition 1.

We next quantify the negative impacts. We observe that the loop frequently occurs. Figure 5 plots a two-hour log of serving cells in the 40-hour test in one 4G-Femtocell-3G example. Our test further shows that, more than 90% loops happen every <200 seconds. With such high frequency, the carrier can neither offload the users to Femtocell, nor offer high-speed 4G service.

Such frequent handoff loops incur large signaling overhead between the phone and the network. Figure 6a shows that, compared with the 4G case, the current loop increases its signaling to the cell

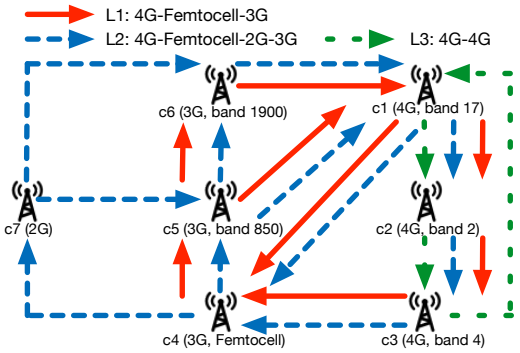


Figure 4: Idle-state persistent loops detected in US-I.

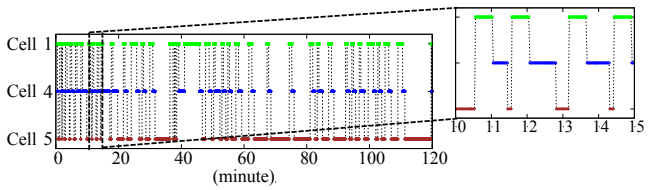


Figure 5: A two-hour log of associated cells at one static phone. The loop is observed despite varying loop cycles.

and to the core network by 7.6x (6329 : 827) and 23.5x (1226 : 58), respectively. These signaling messages are triggered by location update [4], and include messages related to radio resource allocation, data forwarding path reconfigurations and authentications.

These frequent loops can also degrade data performance. To evaluate it, we load a webpage (www.cnn.com) and a music file (about 5MB) every five minutes using Firefox, and record the loading time with/without loop at the same location in Figure 6b and Figure 6c. The loop slows down the webpage downloading by 11x than 4G (33x fold in the worst case). In the worst case, it takes 1.5 minute to download this webpage, whereas it would take 3 seconds using 4G. We observe similar performance slump in the music case with delay increase by 10x (50th) and 14.5x (worst case). The music file can be stably downloaded within 12 seconds using 4G, whereas it takes up to 180 seconds in our test. Such performance degrade is mainly caused by repetitive location update. The location update may trigger data path forwarding reconfiguration and re-authentication, during which the incoming/outgoing traffic would be delayed or dropped. Each 3G and Femtocell location update typically takes 3 to 6 seconds. In the worst case, the Femtocell location update can be further delayed to up to 30s.

Besides data performance, such idle-state persistent loops can also cause call drops. In the same experiment setting, we launch voice calls every five minutes using phone's dialer, and record the failures of voice call setup. We find that the call drop rate is 9.6%. The reason is that, when the call is initiated in presence of loops, the network cannot locate the user to the specific cell, thus unable to establish the voice call session.

◦ **L2: device-side preference misconfiguration.** Our loop detector further reports 8 variants of loops between 4G Macrocells, Femtocell, 2G and 3G Macrocells. Compared with previous category, when leaving the Femtocell, the mobile device handoffs to 2G first, then handoffs to 3G Macrocells. This happens when the Femtocell's signal strength is weak (< -115dBm) but still higher than 4G's high-preference handoff threshold (-116dBm in this scenario). It turns out that, this extra handoff is caused by improper preference configuration on the mobile device. With low signal strength, the device may temporarily lose association to Femtocell. Based on the 3GPP standard, the mobile device resumes the service by scanning all the cells, and associates to the first available one [10]. The order of the scanning is based on a preference list in the phone's SIM card. For some phones, the 2G is listed as highest preference, so the phone moves to 2G instead of 3G Macrocells. Once associated with 2G, the device would immediately handoff to 3G Macrocells, because the 2G cell assigns 3G cells higher preference. This way, the persistent loop continues.

For loops in this category, all negative impacts in 4G-Femtocell-3G loops also retain here. Besides, the device may further lose voice services in 2G cells. The reason is that, some 2G cells cannot support voice and data service *concurrently*. When the device is transmitting data, the voice service would be disabled.

◦ **L3: incremental 4G infrastructure upgrade.** The last variant is a 4G-only loop that appears until recently. We observe that US-I is upgrading its 4G infrastructure, and deploying cells under a new frequency band (c_2 in Figure 4). Before the upgrade, existing 4G cells (c_1 and c_3 in Figure 4) assign equal preferences to each other. US-I intends to migrate users to the new cells, which offers higher bandwidth. To achieve it, some old cells (c_1) assign higher preference to new cells. However, not all cells' preferences are updated timely: equal preference still exists on some cells (c_2). Such partial update cannot migrate user to the new cells: it violates Proposition 1, and incurs loops between cells. This loop has no direct impact on users, because all cells belong to the same location area. But this incurs larger 4G-Femtocell-3G and 4G-Femtocell-2G-3G loops, and indirectly amplifies their negative impacts.

5.2 Inconsistent Thresholds

In this category of instability, handoffs may oscillate among cells with uncoordinated thresholds. This may occur even when the preference values are globally consistent. It can be exemplified in Figure 3b. In the setting, both cells c_1 and c_2 agree that c_1 is preferred, but they apply different rules. Specifically, c_1 uses high-preference rule ($\gamma_2 > \Theta_{1,2}^{high}$ (-108dBm)), and c_2 uses equal-preference rule ($\gamma_1 > \gamma_2 + \Theta_{2,1}^{eq}$ (3dBm)). Therefore, the loop exists as long as the received signal strength meets the above condition. Similar misconfigurations can also occur in active-state handoff.

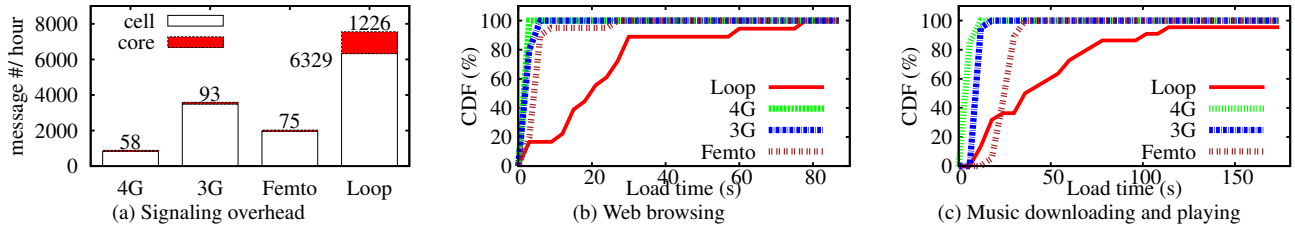
5.2.1 *Instability Condition*

We next derive the instability condition with respect to the radio threshold. We assume the preference settings are globally consistent, *i.e.* they all see the non-conflicting ordering on cell preference values. The following result shows the necessary and sufficient threshold configurations for any loop-free handoff (the proof is in Appendix B):

Proposition 2. *Consider n cells that use the idle-state handoffs only, and configure consistent preferences. The handoff stability is guaranteed iff. the radio thresholds are coordinated as follows: for every two cells c_i and c_j, (1)*$\min_{c_i \to c_k} \Theta_{i,k}^{high} \geq \Theta_j^{serv}$ *if* $P_i > P_j$[2]*; (2)* $\min_{c_j \to c_k} \Theta_{j,k}^{high} \geq \Theta_i^{serv}$ *if* $P_i < P_j$*; (3)* $\Theta_{i,j}^{eq} + \Theta_{j,i}^{eq} \geq 0$ *if* $P_i = P_j$. □

Compared with Proposition 1, Proposition 2 offers a *pairwise* configuration between any two cells. With consistent parameter configurations, two-cell loop avoidance also implies larger loop avoidance.

We further show that instability detection is still polynomial even with inconsistent preference (the proof is in Appendix C):



(a) Signaling overhead (b) Web browsing (c) Music downloading and playing

Figure 6: Impacts of loops of L1.

Proposition 3. *Given n cells at a location, the complexity of finding persistent loops is $O(mn)$, where m is the number of idle-state handoff rules from all cells.* □

5.2.2 Empirical Assessment

For this category of idle-state threshold misconfigurations, our experiments report no conflicts in US-I/US-II. The traces show that, both carriers impose stricter conditions over the idle-state thresholds than required (by Proposition 2). The real threshold settings are *fully decoupled* from the preferences: no matter how preferences are configured, the high-preference threshold Θ^{high} (US-I: [-114dBm, -110dBm], US-II: [-114dBm, -111dBm]) is *always* higher than the serving threshold Θ^{serv} (US-I/II: [-120dBm,-116dBm]) between any two cells. This signifies prudent engineering practice, contributing to good operations by both carriers in reality most of the time.

5.3 Active-Idle Misconfiguration

Instability may also be observed when the idle-state handoff is used in some cells but the active-state handoff is adopted in others. For instance, this could occur when the device exchanges highly bursty traffic, e.g., during Web browsing or instant messaging. The device thus stays active with traffic for a while, but then remains idle without traffic. This active/idle state switching is driven by the setup/release of radio connections, and is regulated by the Radio Resource Control (RRC) protocol [6]. For this scenario, uncoordinated configurations between idle and active state handoff may incur instability.

Figure 3c illustrates such an example of two-cell loop. In the setting, c_1's active-state handoff policy evaluates c_1 and c_2's signal strength with two thresholds. But it does not coordinate with c_2's idle-state handoff. So the persistent loops between them can happen when -111dBm$< \gamma_2 < \gamma_1 -$ 3dBm<-105dBm.

5.3.1 Stability Analysis

We next derive the stability conditions when active-state handoff is involved. Different from idle-state handoff, the active-state handoff logic is customizable. It can thus decide whether to access cells' radio qualities in decision. We first assume all cells' active-state handoffs evaluate radio conditions (the remaining cases will be handled by Proposition 6 of next section). Since there can be many ways to define the radio evaluation criteria, we cannot derive the (in)stability conditions in an arbitrary setting. Instead, we analyze a class of criteria, which are widely used in engineering practice [2, 6, 7] and research [24, 27]. To this end, we assume that the active-state handoff adopts the following radio criteria.

Assumption 1. *For any active-state handoff policy $c_i \rightarrow c_j$, it evaluates cell radio in the decision logic, and takes one of the following forms: (a)* ***absolute comparison****: $\gamma_j > \Theta_{i,j}$ (b)* ***indirect relative comparison****: $\gamma_i < \Theta 1_i, \gamma_j > \Theta 2_{i,j}$; (c)* ***direct relative comparison****: $\gamma_j > \gamma_j + \Theta_{i,j}$.*

In §5.3.2, we will validate that this assumption holds in real operational networks. The following configurations ensure stability for active-idle misconfiguration (the proof is in Appendix D):

Proposition 4. *Consider n cells $c_1, c_2, ..., c_n$ that satisfy Assumption 1. The stability is guaranteed, if all cells' active and idle-state handoffs' radio thresholds are coordinated as follows. For every two cells c_i and c_j, consider c_i's idle-state and c_j's active-state parameters: (1)$\min_{c_i \rightarrow c_k} \Theta^{high}_{i,k} \geq \Theta 1_j$ if $P_i > P_j$ and c_j uses absolute comparison; (2) $\min_{c_j \rightarrow c_k} \Theta_{j,k} \geq \Theta^{serv}_i$ if $P_i < P_j$ and c_j uses indirect relative comparison; (3) $\Theta^{eq}_{i,j} + \Theta_{j,i} \geq 0$ if $P_i = P_j$ and c_j uses direct relative comparison.* □

Given Assumption 1, any active-state handoff can be split into two parts: an equivalent "idle-state handoff" and decisions over other observations. Indeed, when only active-state handoffs are used, Proposition 4 is sufficient but not always necessary. Stability can be ensured through other means (*e.g.* coordinating other parameters). The merit of Proposition 4 stems from its support for idle-state handoffs: it ensures stability within active-state handoffs only, *and* between idle and active state handoffs.

5.3.2 Empirical Validation

Validation of Assumption 1. We first exam whether Assumption 1 holds in real mobile networks. Constrained by no access to the operator's internal handoff decision logics, we gauge it from the runtime cellular messages exchanged between the device and the network (the serving cell). This is viable because in 3G/4G active-state handoff decision is device-assisted: the mobile device reports the radio measurement results to the serving cell, and then the serving cell makes the decision. We find that, both US-I and US-II's active-state handoff commands are usually *directly* triggered by 3G/4G RRC radio measurement reports, whose triggering conditions are standardized in [3, 6] and satisfy Assumption 1[3]. So effectively, the serving base station's handoff logic evaluates radio quality that satisfies Assumption 1. To demonstrate this, we calculate (1) the probability that the active-state handoff occurs after the measurement report, and (2) the device-perceived time interval between the delivery of last measurement reports and receipt of the active-state handoff command. Higher occurrence probability and smaller time interval would imply closer relation between active-state handoff and RRC measurement reports.

[3]For example, in 4G RRC, the report criteria A4 and B1 are the absolute comparisons, criteria A5 and B2 are the indirect relative comparisons, and criteria A3 is the direct relative comparison.

		US-I	US-II
#. Active-state handoff		11,050	10,178
Occurrence probability after RRC measurement report	Total	99.7%	97.8%
	Absolute comparison	1.0%	3.2%
	Indirect relative comparison	21.2%	34.2%
	Direct relative comparison	77.5%	60.4%
Interval between handoff command and last RRC measurement report	Min	40.6ms	20.1ms
	Med	79.9ms	60.0ms
	Max	141.4ms	266.5ms

Table 3: Probability and the elapsed time of active-state handoff which is triggered by RRC measurement report [3,6] (satisfying Assumption 1).

Table 3 shows both results. For both US-I and US-II, 99.7% and 97.8% active-state handoff happens after device sends RRC measurement reports, respectively. Among these handoff commands, 77.5% (60.4%) are initiated after RRC measurements triggered by direct relative comparison (*e.g.*, event A3 in 4G RRC). The time interval between the last measurement report (from device) and the handoff commands are short: the medium value is 79.9ms (60.0ms), while the maximum value is 141.4ms (266.5ms) in US-I (US-II). Both imply that, on reception of device's radio measurement reports (satisfying Assumption 1), the serving base station's decision logic immediately determines to trigger active-state handoff. This is coherent with the public literature [2, 6, 7, 24, 27].

Validation of active-idle misconfiguration. In this subcategory, our detection tool has found one instance in US-I (L4). We observe threshold incoordination between 3G Macrocell's idle-state handoff and Femtocell's active-state handoff for voice. The scenario is similar to that of Figure 3c. The device oscillates between 3G Macrocell (c_1) and Femtocell (c_2) when $\gamma_1 < -102dBm, \gamma_2 > -111dBm, \gamma_1 > \gamma_2 + 3dBm$. Interestingly, though no threshold misconfiguration is observed at the idle state, it occurs between idle and active-state handoffs.

This threshold incoordination is not shown without reasons. The Femtocell tends to move the active-state user to the Macrocell, even when the Macrocell's signal strength is weaker than the Femtocell's. This is because the Femtocell is deployed by users in an unplanned and isolated fashion. Its radio coverage is smaller than that of the Macrocell's. So the device has a higher chance to leave the Femtocell coverage. To avoid the potential voice call disruption, the Femtocell proactively switches the device to the Macrocell earlier than needed. Unfortunately, this configuration violates Proposition 4, and fails to achieve the expected goal. The device may handoff back to the Femtocell at the idle state under the same observations.

We test this loop at all viable indoor locations. At each spot, we launch a 24-hour test, and periodically load the webpage with Firefox every five minutes. We count the total number of connections, and how many experience the looped transition between two cells. The active-state handoff condition is satisfied with probability 9.4% (see Table 4). However, once satisfied, the loop always occurs in our test. We run the same webpage loading test to assess its impact on user traffic. As shown in Figure 7, this loop incurs extra delay about 40–90 seconds.

6. INSTABILITY BY LOOP-PRONE DECISION LOGIC

In this major class of instability, persistent loops may occur if cells apply conflicting handoff decision logics. Different from uncoordinated parameters, no matter how well parameters are tuned, conflicts always exist between decision engines. The fundamental root cause is that, the active-state handoff logic is customizable at each cell. We further discover two categories of instability in this class, one is between active-state logic (§6.1), and the other between active and idle-state engines (§6.2).

6.1 Active-Active Logic Conflicts

The category of instability may appear due to loop-prone decision logic, despite identical parameters. The root cause is that, active-state handoff may use decision logic different from the common one defined by its idle-state counterpart. Operators may customize the logic.

Figure 3d illustrates a two-cell-loop example. The active-state handoff decision logic at each cell adopts a simple rule: both agree to switch to the other if the signal strength at the neighboring cell is good enough (*e.g.* >-106dBm). However, if both cells satisfy the signal strength condition, the loop would occur. Note that regardless of the radio threshold to be set, the signal strength that satisfies the loop condition always exists.

6.1.1 Analytical Result

We now derive the stability conditions for the active-state decision logic. When all active-state decision engines assess radio quality and satisfy Assumption 1, loop-prone logic can be eliminated as follows (the proof is similar to Appendix D):

Proposition 5. *Consider n cells $c_1, c_2, ..., c_n$ that satisfy Assumption 1. The stability is guaranteed, if all cells' active and idle-state handoffs' radio thresholds are coordinated as follows. For every two cells c_i and c_j, consider c_i's idle-state and c_j's active-state parameters: (1)$\min_{c_i \to c_k} \Theta_{i,k} \geq \Theta 1_j$ if c_i uses indirect relative comparison, and c_j uses absolute comparison; (2) $\min_{c_j \to c_k} \Theta_{j,k} \geq \Theta 1_i$ if c_i uses absolute comparison, and c_j uses indirect relative comparison; (3) $\Theta_{i,j} + \Theta_{j,i} \geq 0$ if c_i and c_j use direct relative comparison.* □

Proposition 5 specifies a sufficient condition for loop-free, active handoffs. It is applicable to *any* handoff decision logic among cells, as long as it assesses radio quality. By coordinating the radio evaluation portion, no observations on the radio quality can lead to a loop. There are two benefits to this approach. First, it does not require coordination of other unknown parameters, thus being flexible and extensible. Second, it is also backward compatible with the idle-state handoff. This helps to prevent persistent loops between active-state handoffs, *and* between active and idle-state handoffs.

6.1.2 Empirical Validation

Our tool also detects one instance (L5) between two 4G cells in US-I at one location (Figure 3d). Both cells try to offload users to each other, when the other's signal strength is higher than certain threshold. However, such load-balancing decisions are not coordinated. A user thus oscillates between these two cells. Fortunately, this loop is not commonly observed. Among all 4G cells we collect, 67% of them use the same policy for the active-state handoff, but its neighboring cells are not observed to use the same rule except at one location. At this location, we conduct 6-hour ping tests and observe 8 loops (every 45 minutes on average) and the minimum one lasts only 43 seconds.

We further discover that, active-state loop-prone handoff engines are less observed due to prudent engineering practice. In US-I/US-II, most Macrocells' active-state decision logic is *more conservative* than their idle-state handoffs' counterpart. The radio quality measurements, which may trigger idle-state handoffs, cannot always trigger active-state handoffs in the same cell. The RRC measurement configuration for active handoffs uses stricter thresholds and report criteria than their idle-state handoff counterpart. The

reason is that, active-state handoff is usually activated with data traffic. It tends to be more conservative to ensure the seamless data/voice service. Since the idle-state handoffs' thresholds follow the conservative loop-free setting, it is not surprising to see fewer active-state-only handoff loops.

Another related finding is that, the active-state loop-prone handoff logics are more common in 4G. In both US-I and US-II's 3G Macrocells, we didn't observe loop-prone handoff logics in active state. Our experiments show that, both operators' 3G active-state handoff logics are even more conservative than 4G's active-state counterparts: only handoffs between cells under the same frequency (intra-frequency handoff) are used, whose triggering condition is based on the direct relative comparison between serving and neighboring cells' radio qualities (Assumption 1). The reason is that, different from 4G LTE, 3G UMTS supports *soft handover* between cells under the same frequency. Because of this, intra-frequency handoff offers more seamless data service in mobility, thus more preferred by network operators. This further limits the available candidate cells, thus less prone to loops than 4G.

6.2 Active-Idle Logic Conflicts

This category of instability occurs when some cells apply idle-state handoffs, while others use active-state handoffs. Instability is always triggered if the handoff does not assess radio quality. This can be illustrated by the example of Figure 3e. In this case, c_1's active-state decision never considers the signal strength used by the c_2's idle-state handoff. Instead, c_1 uses load-balancing, regardless of the radio quality. Consequently, the serving cell oscillates between c_1 and c_2 once the switch conditions are satisfied in both handoff iterations. It shows that incoordination between decision logic functions is responsible for this unnecessary loop.

6.2.1 Analytical Result

If the active-state handoff does not assess radio quality, we show that idle-active loops would always occur:

Proposition 6. *Two-cell persistent loop between c_i and c_j always exists, if c_i's active-state logic to c_j does not evaluate radio quality, while c_j applies idle-state decision logic to c_i.* □

The proof can be found in Appendix E. In practice, the idle-state decision logic is available at all devices. Proposition 6 implies that, if a cell c_i's active-state handoff logic does not assess radio quality, the only possibility to avoid loops is that all its neighboring cells do not allow handoffs to c_i. However, this cell would consequently become isolated from others.

6.2.2 Empirical Validation

We have found two instances in this category, one (L6) in US-I and US-II, and the other (L7) in US-I only. L6 happens between two 3G Macrocells, whereas L7 happens between a 3G Macrocell and a Femtocell. The handoff decision logic is shown in Figure 3e, with c_1 = 3G and c_2 = 3G/Femtocell. This setting violates Proposition 4.

Our study shows that, both instances are caused by a design defect in 3G Radio Resource Control (RRC) protocol. The 3G RRC defines an offloading mechanism during connection setup. When a device attempts to setup a radio connection, the cell can reject the device's request, and redirects the device to a nearby cell [3]. However, this redirection *cannot* take cells' radio quality into consideration. Without a connection, the device cannot report the observations to the cell. If the current cell's radio quality is better than neighbors', the offloading cannot succeed, because the device would shift back in idle state.

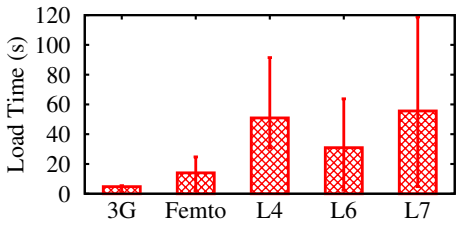


Figure 7: Impacts of loops of active-idle misconfiguration logic conflicts (web page loading).

We quantity both loops' impact under similar experiment settings to those for IS1-C3. We find that L6 and L7 are not commonly observed even when the loop condition is satisfied. L6 and L7 occur at the probability of 2.15% and 0.49% in our observation at one location (Table 4). The serving-cell congestion probability largely decides the loop occurrence. As shown in Figure 7, both loops incur delay about 20–53 seconds (median), up to two minutes. We observe that some phones do not always follow the cell's handoff command. Instead, they seek to reconnect to the serving cell. This is why the user device may still not suffer from it even when the loop condition is met.

7. AUTOMATIC LOOP DETECTION

With above analytical findings, we next design and implement a software tool, which enables automatic detection on handoff instability. Given the parameters and decision logics from cells at a location, our tool reports the conflicting parameters/logics that may incur instability, uncovers their root causes, and identifies the possible runtime observations that will trigger the loop occurrence.

7.1 Design

At first glance, loop detection looks fairly simple. A straightforward straw man solution works as follows. Given handoff configurations from all cells, the tool first enumerates all the looped handoff transitions $c_1 \rightarrow c_2 \rightarrow ... \rightarrow c_n \rightarrow c_1$. For each of them, a persistent loop would be reported, if there exist some runtime observations that can satisfy handoff configurations concurrently. To uncover root causes, we search whether there exists a set of configurations for each handoff that can eliminate the loop. If yes, then the loop is incurred by uncoordinated configurations (§5); otherwise, it is due to loop-prone logic (§6).

However, this approach is deemed impractical. Enumerating loops implies state explosions on the number of cells and configurations. Uncovering root causes further incurs exponential search over the configuration space. To address both issues, we apply the domain-specific (in)stability conditions to reduce the complexity. We differentiate loops with/without active-state handoffs, and apply distinctive techniques for loop detection and cause inference.

∘ *Idle-state handoff only.* We adopt a two-step procedure to detect all idle-state handoff loops. First, given all cells at each location, we check if these cells' preference settings are consistent. If so, enumeration can be replaced with pairwise threshold check according to Proposition 2. Otherwise, enumeration with pruning is applied based on Proposition 1. In each round, we seek to find all loops with a cell c_i involved. Starting from c_i, we run a variant DFS algorithm over the graphs of handoff transitions. For each cell c_j to be visited, a variable γ_j is maintained, which denotes the potential radio measurement violating stability. We initialize γ_i, and derive γ_j recursively. When c_j is visited from c_k, (1) if $P_{k,j} > P_{k,k}$, then

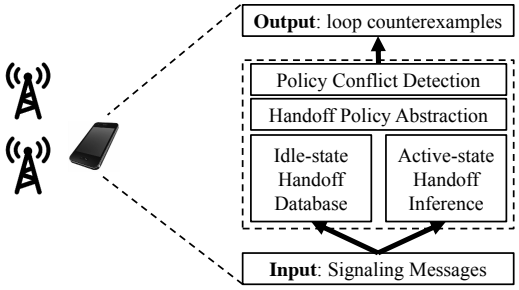


Figure 8: In-device for loop detection.

$\gamma_j \leftarrow \Theta^{high}_{k,j}$; (2) if $P_{k,j} = P_{k,k}$, then $\gamma_j \leftarrow \gamma_k + \Theta^{eq}_{k,j}$; (3) if $P_{k,j} < P_{k,k}$, c_j is pruned if $\gamma_k > \Theta^{serv}_k$; otherwise, $\gamma_j \leftarrow \Theta^{low}_{k,j}$. If c_i is visited again, a loop is reported if γ_i can satisfy the last handoff transition. Once all loops with c_i are found, we remove c_i from the cell list, and detect loops for the remaining cells. The root cause analysis is straightforward, since the loops in this category can only be caused by misconfigurations.

∘ *Active and idle-state handoff.* With active-state handoffs, we apply Propositions 4, 5 and 6 to eliminate unnecessary enumerations. For each cell c, we first check if its active-state handoff execution evaluates radio quality. If not, idle-active handoff loops are reported between c and its neighbors, with the cause "loop-prone logic" (Proposition 6). We further remove c, and replace each handoff transition from c with a new one from c's parents, with the configuration as that for $parent(c) \rightarrow c$. Next, if all cells' handoff executions apply pairwise radio quality evaluation, we check if they satisfy Propositions 4 and 5. If yes, no loops would be reported. This helps to avoid checking other parameters. Otherwise, enumeration would be invoked in the worst case.

7.2 In-phone Implementation

Since we have no access to the cellular network infrastructure, we implement the above design in the mobile device. Figure 8 illustrates our implementation. We first enable in-device cellular configuration collection with *MobileInsight* [1], a system app to collect signaling messages. With these inputs, our tool actively synthesizes the abstract model and applies the loop detection algorithm.

∘ *Handoff configuration collection.* Our implementation proactively switches the phone to every cell at each location (via *e.g.* secret code **#2263#* in Samsung S5), and collects both idle and active-state handoff configurations. The idle-state handoff configurations can be readily derived from the standardized logic and parameters from the *System Information Block (SIB)* messages. The active-state handoff ones are not visible to the device. To infer them, we observe that the serving cells' measurement configurations are available, which are designed for active-state handoffs. These report criteria are standardized in [3, 6] and satisfy Assumption 1. For each cell, our implementation monitors these configurations and handoff commands in *RRCReconfiguration* message, together with the measurements from the device. If a handoff command without measurement is observed, a configuration without radio evaluation is inferred. Otherwise, all measurement reports before the command are treated as potential configurations, and checked in the loop detection.

∘ *Parameter abstraction.* The actual configurations should be mapped to the abstract parameters before detection. Following [3, 6], we perform three parameter conversions before the detection: (1) *threshold set selection*, in which we select the corresponding thresholds based on the usage scenarios (*e.g.* RSRP or RSRQ as the radio metric, normal or VoLTE active-state handoff); (2) *threshold combination*, in which we combine different offsets/hysteresis from real parameters to the thresholds in the model. For example, in the equal-preference, idle-state handoff, the threshold $\Theta^{eq}_{s,c}$ in the model is the sum of the serving cell's hysteresis, the frequency-specific offset, and the cell-specific offset; (3) *threshold scaling*, in which we scale the thresholds based on the speed-scaling factors from SIBs.

∘ *Loop detection.* We implement the detector in Python, which accepts the abstract model as input, and reports the counterexamples. Each report includes the cells in the loop, the conflicting configurations, and the measurements that can trigger the loop. They are used to set up the validation experiments in Sections 5 and 6.

Accuracy analysis. For idle-state handoff only, our tool guarantees that it can find all loops without false positives/negatives. This is because that the phone has full knowledge of the idle-state handoffs, including its decision logics (standardized) and parameters (from the signaling messages). For active-state and active-idle loops, false positive exists. This is because our inference of the decision logics based on Assumption 1 may be incomplete. So the over-detection may exist: active-state and active-idle loops reported by our tool may not always happen in reality. But false negative is still prevented from active-state and active-idle loops: if a loop exists, it would always be detected by our tool.

Limitations. Our current design and implementation have three downsides. First, the loops reported in our design may not always be observed in practice, because the observations triggering the loops may not always appear. To test the loop existence, validation experiments should be conducted. However, they still offer hints for validations, and should be fixed because external measurements cannot always be controlled to avoid loops. Second, the active-state handoff configurations are not accessible to device-side implementation. This causes false positives, and prevents us from uncovering more insights on the active-state handoff. Third, without access to the carrier's handoff configurations, our tool has to be run area-by-area to detect the loops.

7.3 Experiments on Operational Networks

We run the designed tool to validate persistent loops from real carrier configurations, and quantify their negative impacts. Figure 9 summarizes the outdoor and indoor test settings. The cell distribution at different outdoor locations confirms that today's deployment is quite dense and hybrid. At most locations, there are about 8–16 cells. On average, there are about 11 cells in US-I and 10 cells in US-II. The number of unique cells, excluding those observed at multiple locations, are 275 (4G: 120, 3G: 97, 2G: 58) in US-I and 222 (4G: 92, 3G: 66, 2G: 64) in US-II. It confirms that 4G cells have smaller coverage and denser deployment whereas the 2G coverage is much larger. The indoor setting has similar cell density as the outdoor one. Figure 9c plots the median radio signal strength measured at 50 indoor spots in US-I networks. For 2G/3G/4G comparisons, we use normalized percentages obtained from OpenSignal[4], a popular network monitor app, where 0% indicates no coverage and 100% indicates strongest signal strength[5]. It implies that despite higher speed, 4G suffers worse coverage than 3G and 2G in indoor scenarios. The results in US-II are similar and thus omitted here.

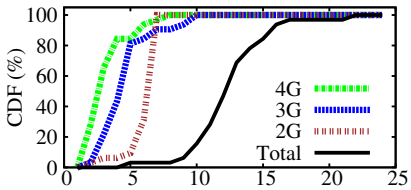
We also validate that, the assumption of invariant runtime ob-

[4]http://opensignal.com/android/

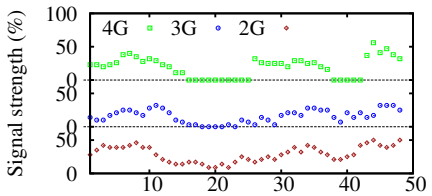
[5]4G uses different signal strength metrics from 3G/2G. The minimal strength observed in 3G/2G is -113dBm whereas it is around -125dBm in 4G.

	Avg. cell#/spot		Unique cell#	
	US-I	**US-II**	**US-I**	**US-II**
#4G	2.6	2.1	120	92
#3G	3.4	2.4	97	66
#2G	5.4	5.6	58	64
#All	11.4	10.1	275	222

(a) Statistics of outdoor cell deployment



(b) Outdoor cell density in US-I



(c) Indoor radio signal strength at 50 spots in US-I

Figure 9: Summary of outdoor and indoor deployment.

	#Scenario instances	Occurrence of Misconfigurations or Loop-prone logic	Loop occurrence (parameter+logic +observation)
L1: 4G-Femto-3G	8	96.8%	25.0%
L2: 4G-Femto-2G-3G	8	96.8%	0.49%
L3: 4G-4G	1	2.2%	2.2%
L4: 3G-Femto	1	96.8%	9.4%
L5: 4G-4G	1	1.6%	1.6%
L6: 3G-3G	1	63.4%	2.15%
L7: 3G-Femto	1	96.8%	0.49%

Table 4: Loop occurrence probability in US-I.

servations in persistent loop is reasonable in practice: in response to runtime radio measurement, both US-I and US-II's handoff decision-makings takes no more than 141.4ms and 266.5ms (§3), respectively. For comparison, for all the tested indoor spots, 95% of cells' signal strength change at same spot takes more than 229.5ms (5.10s) in US-I (US-II), which is much slower than handoff decision making. So during the handoff decision, it is safe to make the invariant runtime observation assumption.

The detailed findings on each category have been described in Sections 5 and 6. With our tool, we have found 21 instances of potential misconfigurations and/or loop-prone logics, which are further classified into 7 categories. For each category, we further run indoor experiments to validate its existence, and estimate its occurrence probability. For each indoor spot, we run a 24-hour test and record the looped handoffs between cells. Table 4 lists the occurrence probability of problematic configurations (left column), and the occurrence probability of loops (right column) observed at one specific location. Other locations have similar results. It shows that, instabilities occur in 2G, 3G and 4G networks, with varying occurrence probabilities. From these instances, we show that, loops with both the uncoordinated configurations and loop-prone decision logic indeed exist. Although carriers have applied at least two prudent rules to mitigate loops, configuration conflicts still exist for various reasons, such as diverse handoff goals, the incremental and/or unplanned cell deployment, the device misconfiguration, and the design defects for the connection control mechanism. Loops incur negative impacts upon both the user and the network. We notice a big distinction between both columns, which reflect the gaps of the root causes and the actual impact. The reason is that, the occurrence of actual loops (right column) is also affected by another runtime observations, which may not always be satisfied. It has two implications. First, misconfigurations or loop-prone logics that may trigger persistent loops are not rare in reality. Most settings are problematic once femtocells are deployed. It indicates that the operator's network infrastructure is not fully upgraded to handle small cells which can be deployed by users. Second, although the misconfigurations occur with high probability, the satisfying signal strength that triggers loops do not always occur. For example, only L1 (25%) is relatively common and other loops like L2, L3, L5, L6 and L7 are rarely observed (below 2%). This is attributed to good practice and satisfactory coverage in radio planning and cell deployment.

8. DISCUSSIONS

We now discuss how to fix the configuration conflicts and their resulting loops. Given that persistent loops hurt both the user performance and the network's operation, we envision that both carriers and users have incentives to remove loops. We next propose solutions to both sides.

8.1 Network-Side Coordination

The carrier should coordinate cells' local configurations to avoid handoff instability. There are two issues to be addressed: (1) How to resolve loops from the existing handoff configurations, and (2) How to avoid new persistent loops from configuration updates?

Fixing existing loops. The carrier can take two steps. First, the operator should check if configuration conflicts exist at each location. This can be done with our loop detection tool (§7).

In the second step, conflicting configurations in loop may be coordinated for stability. It should be noted that, there can be more than one way to fix each loop. Consider the actual loop with the Femtocell involved in Figure 4 (§5.1.2). At least two fixes are available: *(i)* on 4G Macrocell, assign lower preference to Femtocell, or *(ii)* on Femtocell, assign lower preference to 3G Macrocell. The carrier may pick either one based on its demand. For example, applying (i) can provide users with high-speed data service in 4G, while choosing (ii) enables traffic offloading from 4G to 3G.

However, not all schemes are bullet-proof. New loops may appear when fixing old ones. For example, the following scheme can also fix the above loop: *(iii)* on Femtocell, assign higher preference to 3G Macrocell. However, this causes a new loop $Femtocell \rightarrow 3GMacrocell \rightarrow Femtocell$, because it violates Proposition 1. To address this, one could detect the loop again after the fix, and resolve new loops. But this requires enumeration of all configurations. More importantly, there is no guarantee that all loops will be finally fixed.

We propose a general guideline to determine a safe loop fix. Assume a fix requires to modify the configuration for $c_i \rightarrow c_j$. Our guideline imposes a monotonic condition over this fix:

> **Guideline 1. (Safe configuration update)** *For any runtime measurements that cannot trigger $c_i \rightarrow c_j$ before the configuration update, it should not trigger $c_i \rightarrow c_j$ after the update.*

If Guideline 1 is followed, it is straightforward to prove that, for any loop that exists *after* this configuration update, it must have already existed *before* the update. New loops cannot appear thereafter. No extra actions are needed after the old loop fix. So if all fixes obey this rule, all loops will be ultimately fixed.

Handling policy update. The handoff configurations or decision logics can change over time for various reasons, including incremental/unplanned cell deployment, tuning some cells' handoff goals, etc. The carrier may expect to retain stability after the update. We discuss how to achieve this in a single configuration update $c_i \rightarrow c_j$, assuming that stability is guaranteed before the change. Stability with multiple updates can be ensured with each step satisfying the following criteria.

Consider cell c_i adds a new handoff rule to c_j. Two approaches exist to prevent new loops. The first one is to detect loops after the addition, which however implies enumeration of loops (§7). The second is to apply constraints over the new update, without coordinating with others. In §5-6, we have seen two prudent rules from real carriers: (1) for a new idle-state handoff, set threshold Θ^{high} higher than threshold Θ^{serv}, regardless of the preference setting; (2) for a new active-state handoff, set it as at least conservative as the corresponding idle-state handoff. Both rules help to avoid new loops. Propositions 1– 6 also impose such conditions on new updates.

Next consider a configuration deletion: a cell c_i deletes its local handoff toward c_j. The stability is still retained, because no new handoff transitions are introduced.

The last scenario is to modify an existing handoff configuration for $c_i \rightarrow c_j$. This update is equivalent to a two-step procedure: delete the old configuration, and then add the new one. The first step does not cause new loops, while the second step can be safeguarded by above rules for new configurations. Guideline 1 is also applicable here: If the new policy is as conservative as the old one, the overall mobility are still stable after the update.

Runtime mitigation? In addition to configuration coordinations, we are aware that loops can be mitigated at runtime, similar to the transition loops due to radio dynamics. The mobility history can be used to stop the loops by blocking visited cells [9]. Signaling reduction techniques can reduce the impacts [4]. Despite helpful, they do not offer fundamental fixes to loops. Since the configuration-based loops are fixable, the carrier should not rely on runtime mitigations only. Instead, they should fix the configuration conflicts.

Implementation suggestions. To fix existing loops, a centralized controller is needed for configuration coordination. For loops between cells from the same area, loop resolution can be implemented at the location area controller. For other loops, higher-layer controller may be needed. For configuration updates, they can be handled in either centralized or distributed fashions. If loop detection is used, a central controller is still needed. If extra constraints are applied to configuration addition/modification, the resolution can be implemented at each cell. Every cell discovers all neighboring cells' handoff configurations under the self-organizing network (SoN) framework [7, 8], and apply these constraints over its own updates.

8.2 Device-Side Loop Prevention

When loops are not fixed by the carrier, the mobile device has the incentive to prevent itself from suffering from loops. Although the device cannot coordinate configurations among cells, it can configure itself to eliminate loops. At each location, the device can actively collect all cells' configurations, and run the loop detection algorithm (§7). Upon detecting a loop, the device can either statically block some cells, or stop the handoff procedure at runtime. Noted that, this device-side approach cannot replace the network-side configuration coordination. Without the network-side fix, users without this scheme still suffer from persistent loops.

9. RELATED WORK

In recent years, there have been extensive research efforts on 3G/4G mobile networks. They include radio analysis [11, 21], TCP over cellular network [18, 29], cross-layer optimization [13, 19], and software-defined cellular networks [17, 20]. On mobility support, extensive optimizations have been proposed for diverse goals, including reducing radio link failure and transient oscillations [14, 24], traffic offloading [12, 16], and supporting heterogeneous network [15, 23]. Instead of optimizing one specific handoff goal, our work focuses on the conflicts between handoff policies (with possibly different goals). Our preliminary work [22] discloses the existence of persistent loops in idle-state handoffs and this work greatly extends it. We conduct a systematic study with both analysis and empirical validation and covers both idle-state handoffs and active-state handoffs.

Misconfigurations and management-plane conflicts have been examined in other Internet systems such as BGP [25], DNS [26], and data center networks [28]. Our work is inspired by such early efforts, but focuses on the management plane for mobility management in 3G/4G mobile networks. We show that policy-driven configurations may lead to instability during handoffs.

10. CONCLUSION

Mobility support is a key utility function offered by 3G/4G cellular networks. As more mobile users are accessing the Internet from their smartphones through the 3G/4G infrastructure, mobility management is likely to become more critical. Like all operational networks, current mobile carriers allow for flexible configurations on their micro-mobility support scheme to address policy concerns. This management-plane aspect on mobility has been largely overlooked by past research efforts. In this work, we seek to conduct a first study toward this general direction.

Our results can be best interpreted from both viewpoints. On one hand, our effort yields some interesting and new findings. The discovered persistent loops, as well as their triggering conditions, have been partially validated in operational networks. Though the incurred damage, in terms of signaling overhead and performance degradation, is not appalling to some users, such problematic issues should be addressed as we seek to build a more dependable, high-performance, mobile network infrastructure. The presented analysis, as well as modeling, despite a little simplistic, also produces some interesting results not reported in the literature. On the other hand, this work is still at the early stage, and the obtained results are likely to be refined over time. Several important issues (e.g., other configuration parameters, and more forms of customized decision logic) have been overlooked so far. Moreover, other structural properties (e.g., whether the handoff process converges to the anticipated cell, and the convergence speed) warrant further efforts. In the broader context, moving beyond current focus on both data and control planes, management plane in 3G/4G mobile networks (hopefully also the upcoming 5G) is still a wide-open research area and deserves more attention. We hope our effort may stimulate more people in the community to work on this important direction.

Acknowledgments: We thank the anonymous reviewers for their constructive comments. We greatly appreciate our shepherd, Dr. Bozidar Radunovic for his valuable feedback. We also thank Jie Zhao for participating in trace collection in this work. This work is supported in part by NSF awards (CNS-1526456, CNS-1526985, CNS-1423576 and CNS-1421440).

11. REFERENCES

[1] Mobileinsight project. http://metro.cs.ucla.edu/mobile_insight.
[2] "ZTE UMTS Handover Description". *http://www.slideshare.net/quyetnguyenhong/zte-umtshandoverdescription*.
[3] 3GPP. TS25.331: Radio Resource Control (RRC), 2006.
[4] 3GPP. TS24.008: Mobile Radio Interface Layer 3, 2012.
[5] 3GPP. TS25.304: User Equipment (UE) Procedures in Idle Mode and Procedures for Cell Reselection in Connected Mode, 2012.
[6] 3GPP. TS36.331: E-UTRA; Radio Resource Control (RRC), 2012.

[7] 3GPP. TS32.522: Self-Organizing Networks (SON) Policy Network Resource Model (NRM) Integration Reference Point (IRP), 2013.
[8] 3GPP. TS32.511: Telecommunication Management; Automatic Neighbor Relation management; Concepts and Requirements, 2014.
[9] 3GPP. TS36.413: S1 Application Protocol (S1AP), 2014.
[10] 3GPP. TS36.304: E-UTRA; User Equipment Procedures in Idle Mode, 2015.
[11] P. K. Athivarapu, R. Bhagwan, S. Guha, V. Navda, and et.al. Radiojockey: mining program execution to optimize cellular radio usage. In *ACM MobiCom*, Aug. 2012.
[12] A. Balasubramanian, R. Mahajan, and A. Venkataramani. Augmenting mobile 3g using wifi. In *ACM MobiSys*, June 2010.
[13] N. Balasubramanian, A. Balasubramanian, and A. Venkataramani. Energy consumption in mobile phones: A measurement study and implications for network applications. In *IMC*, 2009.
[14] C. Brunner, A. Garavaglia, M. Mittal, M. Narang, and J. V. Bautista. Inter-system Handover Parameter Optimization. In *VTC Fall*, 2006.
[15] M. Z. Chowdhury, W. Ryu, E. Rhee, and Y. M. Jang. Handover between Macrocell and Femtocell for UMTS Based Networks. In *IEEE ICACT*, 2009.
[16] W. Dong, S. Rallapalli, R. Jana, L. Qiu, K. Ramakrishnan, L. Razoumov, Y. Zhang, and T. W. Cho. ideal: Incentivized dynamic cellular offloading via auctions. *TON*, 22(4):1271–1284, 2014.
[17] A. Gudipati, D. Perry, L. E. Li, and S. Katti. Softran: Software defined radio access network. In *Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking*, pages 25–30. ACM, 2013.
[18] J. Huang, F. Qian, Y. Guo, Y. Zhou, Q. Xu, Z. M. Mao, S. Sen, and O. Spatscheck. An in-depth study of LTE: Effect of network protocol and application behavior on performance. In *SIGCOMM*, 2013.
[19] U. Javed, D. Han, R. Caceres, J. Pang, S. Seshan, and A. Varshavsky. Predicting handoffs in 3g networks. In *MobiHeld*, 2011.
[20] X. Jin, L. E. Li, L. Vanbever, and J. Rexford. Softcell: scalable and flexible cellular core network architecture. In *CoNEXT*, 2013.
[21] S. Kumar, E. Hamed, D. Katabi, and L. Erran Li. LTE Radio Analytics Made Easy and Accessible. In *ACM SIGCOMM*, 2014.
[22] Y. Li, J. Xu, C. Peng, and S. Lu. A First Look at Unstable Mobility Management in Cellular Networks. In *HotMobile*, Feb 2016.
[23] M. Liu, Z. Li, X. Guo, and E. Dutkiewicz. Performance Analysis and Optimization of Handoff Algorithms in Heterogeneous Wireless Networks. *IEEE Transactions on Mobile Computing*, 7(7):846–857, July 2008.
[24] A. Lobinger, S. Stefanski, T. Jansen, and I. Balan. Coordinating Handover Parameter Optimization and Load Balancing in LTE Self-Optimizing Networks. In *VTC Spring*. IEEE, 2011.
[25] R. Mahajan, D. Wetherall, and T. Anderson. Understanding BGP Misconfiguration. In *ACM SIGCOMM CCR*, 2002.
[26] V. Pappas, Z. Xu, S. Lu, D. Massey, A. Terzis, and L. Zhang. Impact of Configuration Errors on DNS Robustness. In *SIGCOMM*, 2004.
[27] G. P. Pollini. Trends in Handover Design. *IEEE Communications Magazine*, 34(3):82–90, 1996.
[28] P. Sun, R. Mahajan, J. Rexford, L. Yuan, M. Zhang, and A. Arefin. A Network-State Management Service. In *ACM SIGCOMM*, 2014.
[29] F. P. Tso, J. Teng, W. Jia, and D. Xuan. Mobility: A Double-Edged Sword for HSPA Networks: A Large-Scale Test on Hong Kong Mobile HSPA Networks. *IEEE Transactions on Parallel and Distributed Systems*, 23(10):1895–1907, 2012.
[30] Wikipedia. Handover. *http://en.wikipedia.org/wiki/Handover*.
[31] D. Wong and T. J. Lim. Soft handoffs in CDMA Mobile Systems. *IEEE Personal Communications*, 4(6):6–17, 1997.

APPENDIX

A. PROOF FOR PROPOSITION 1

The proof is based on the fact that, the above preference setting is non-decreasing. We can always find a runtime observation that incurs loop as follows: (a) $\gamma_{i+1} > \Theta^{high}_{i,i+1}$, and (b) $\gamma_{j+1} > \Theta^{high}_{j,j+1}$ if $P_{j,j} < P_{j,j+1}$, or $\gamma_{j+1} > \gamma_j + \Theta^{eq}_{j,j+1}$ if $P_{j,j} = P_{j,j+1}$. On the other hand, if above preference setting does not hold, two cases may arise: (1) all cells are of equal preference, then Proposition 2 below will guarantee the loop freedom; (2) at least one cell c_k has $P_{k,k} > P_{k,k+1}$, then the loop can be avoided by setting Θ^{serv}_k equal to the minimum achievable signal strength. □

B. PROOF FOR PROPOSITION 2

The proof starts from the two-cell case. By enumerating all potential looped transitions, we verify that two-cell loops can be eliminated if and only if the above condition holds. This way, the necessity of Proposition 2 is readily guaranteed. The sufficiency of the condition is further proven via contradiction. Assume the above threshold conditions are satisfied, but a larger loop $c_1 \rightarrow c_2 ... \rightarrow c_n \rightarrow c_1$ still exists. There are two possibilities: (a) all cells are of equal preference. By listing their handoff decisions (inequalities) and summing them up, it can be readily shown that (3) is not satisfied; (b) some cells are of different preference. Then it can be shown that there always exists one cell applying high-preference decision, another cell applying low-preference decision, and any cells in between applying equal-preference decision. By listing their handoff decisions (Figure 2) and summing them up, it can be readily shown that (1)–(3) are not satisfied. □

C. PROOF FOR PROPOSITION 3

We prove it by constructing a corresponding algorithm. Given an arbitrary cell c_i, we find all loops with c_i involved. Starting from c_i, we run a variant DFS algorithm by following the idle-state decision rules for transition. In the process, for each cell c_j to be visited, a variable $\gamma_{i,j}$ is maintained, which denotes the potential signal condition violating stability. When c_j is visited by c_k in DFS, (1) if $P_{k,j} > P_{k,k}$, then $\gamma_{i,j} \leftarrow \Theta^{high}_{k,j}$; (2) if $P_{k,j} = P_{k,k}$, then $\gamma_{i,j} \leftarrow \gamma_{i,k} + \Theta^{eq}_{k,j}$; (3) if $P_{k,j} < P_{k,k}$, c_j is pruned in DFS if $\gamma_{i,k} > \Theta^{serv}_k$, otherwise $\gamma_{i,j} \leftarrow \Theta^{low}_{k,j}$. Whenever c_i is visited again, a persistent loop is reported. Once all loops with c_i involved are found, we remove c_i and all the related decision rules from the cell list, and detect loops for the remaining cells. □

D. PROOF FOR PROPOSITION 4

We prove this with a concrete configuration scheme. For every active-state decision logic $c_i \rightarrow c_j$, we define an auxiliary "idle-state" handoff logic based on its radio evaluation part. If form (a) in Assumption 1 is used, then the auxiliary logic is defined as $P_{i,j} > P_{i,i}, \Theta^{high}_{i,j} = \Theta_{i,j}$. If (b) is used, it is defined as $P_{i,j} < P_{i,i}, \Theta^{serve}_i = \Theta 1_i, \Theta^{low}_{i,j} = \Theta 2_{i,j}$. Otherwise, it is defined as $P_{i,j} = P_{i,i}, \Theta^{eq}_{i,j} = \Theta_{i,j}$. This auxiliary logic has the following property: for any runtime observation that can trigger active-state handoff $c_i \rightarrow c_j$, it can also trigger this idle-state handoff. Then we replace all active-state handoff policies with auxiliary ones, and apply Propositions 1 on all policies for threshold coordination. Based on the relation between active-state handoff and the auxiliary "idle-state" handoff, this coordination ensures stability among idle and active-sate handoff decisions. □

E. PROOF FOR PROPOSITION 5

It is proven by contradiction. Note that, regardless of the preference and/or threshold setting on c_j, certain cell radio quality γ_i and γ_j can trigger idle-state handoff from c_j to c_i. Without evaluating radio quality, $c_i \rightarrow c_j$ can always concurrently happen with $c_j \rightarrow c_i$. □