A First Look at Unstable Mobility Management in Cellular Networks

Yuanjie Li
Dept. Computer Science
University of California, Los Angeles
Los Angeles, CA, USA
yuanjie.li@cs.ucla.edu

Jiaqi Xu
Dept. Computer Science Engineering
The Ohio State University
Columbus, OH 43210
xu.1629@osu.edu

Chunyi Peng
Dept. Computer Science Engineering
The Ohio State University
Columbus, OH 43210
chunyi@cse.ohio-state.edu

Songwu Lu
Dept. Computer Science
University of California, Los Angeles
Los Angeles, CA, USA
slu@cs.ucla.edu

ABSTRACT

Mobility management is a prominent feature in cellular networks. In this paper, we examine the (in)stability of mobility management. We disclose that handoff may never converge in some real-world cases. We focus on *persistent* handoff oscillations, rather than those *transient* ones caused by dynamic networking environment and user mobility (*e.g.*, moving back and force between two base stations). Our study reveals that *persistent* handoff loops indeed exist in operational cellular networks. They not only violate their design goals, but also incur excessive signaling overhead and data performance degradation. To detect and validate instability in mobility management, we devise MMDIAG, an *in-device* diagnosis tool for cellular network operations. The core of MMDIAG is to build a handoff decision automata based on 3GPP standards, and detect possible loops by checking the structural property of stability. We first leverage device-network signaling exchanges to retrieve mobility management policies and configurations, and then feed them into MMDIAG, along with runtime measurements. MMDIAG further emulates various handoff scenarios and identifies possible violations (*i.e.*, loops) caused by the used policies and configurations. Finally, we validate the identified problems through real measurements over operational networks. Our preliminary results with a top-tier US carrier demonstrate that, unstable mobility management indeed occurs in reality and hurts both carriers and users. The proposed methodology is effective to identify persistent instabilities and pinpoint their root causes in problematic configurations and policy conflicts.

1. INTRODUCTION

Mobility management (MM) is widely regarded as a fundamental service to the evolving Internet. To support billions of mobile devices (including smartphones, tablets, wearables, IoT, *etc.*), the 4G/3G/2G cellular network plays a pivotal role. To date, it is the only deployed large-scale system that successfully offers wide-area, ubiquitous Internet access and mobility support.

A key MM function to 4G/3G/2G network is *handoff*, which migrates the device from one serving cell (also known as base station) to another new one *when necessary*. The necessity is defined to satisfy versatile (sometimes conflicting) demands such as sustaining pervasive network availability, offering seamless voice/data support, providing high-speed data service, balancing the traffic load between cells, to name few.

Stability is a desirable property in MM. It states that MM should converge to certain choice given an invariant setting. It is desirable because each handoff comes at a cost. Each handoff incurs multi-round signaling exchanges and causes data/voice suspension or degradation. The more frequent handoffs, the higher cost to carriers and users.

In this paper, we take the first effort to examine the structural property of stability in MM. We are particularly interested in whether MM in reality suffers from persistent loops and whether such loops are caused by fundamental conflicts (*e.g.*, inconsistent policies, uncoordinated configurations), rather than by transient factors such as radio dynamics and user behaviors [1, 2]. Our work is inspired by the observation that, while each individual handoff policy or procedure may be well justified, the interplay among multiple handoffs can be problematic. Note that, each individual cell or the mobile device may customize its local policy in determining the target cell. The handoff decision is thus affected by each other, and prudent coordination is required. Otherwise, policy conflicts or misconfigurations lead to unstable handoffs.

We start with a real-world persistent-loop example to motivate our study (§3). We disclose its causes and the potential damages. It turns out that, a user-deployed femtocell introduces *conflicting* preference settings with two existing 3G and 4G cells, thus causing persistent loops among these three cells. It incurs 3–8x signaling overhead and 10-fold or more slowdown in file downloading. We then formulate the (in)stability problem and derive the necessary and sufficient conditions for stability (§4). Based on these rules, we further devise MMDIAG, an in-device approach to detect and validate possible instability in MM (§5). We leverage signaling exchanges between the device and the serving cell in the standard specifications to tackle the challenge without requiring access to network-side information. We build an automatic detector which enumerates each possible scenario and examines its likelihood of

HotMobile '16, February 26-27, 2016, St. Augustine, FL, USA
© 2016 ACM. ISBN 978-1-4503-4145-5/16/02...$15.00
DOI: http://dx.doi.org/10.1145/2873587.2873599

violating the stability. Finally, we validate our identified findings through real experiments. Our preliminary study via a top-tier US carrier shows that instability indeed exists and our proposed approach is effective (§6).

The paper makes three contributions.

- We present the first work to uncover persistent instability caused by misconfigurations and policy conflicts in mobility management, to the best of our knowledge.
- We devise MMDIAG, a device-based solution to identifying mobility instability in cellular networks.
- We conduct real experiments and validate the identified problems in an operational carrier network. We find that inconsistent mobility management between (macro)cells and femotcells are the main source of many handoff loops.

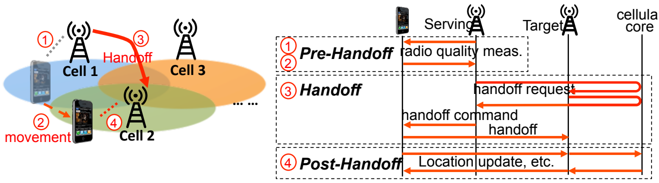
2. UNDERSTANDING MOBILITY MANAGEMENT IN CELLULAR NETWORKS

We first introduce necessary concepts on mobility management.

Handoff procedure flow. To depict the handoff procedure flow, we use a typical scenario: the user is about to move out of the coverage of the current serving cell. Figure 1 gives an illustrative example. Initially, the phone is served by Cell 1. As it moves toward Cell 2 (away from Cell 1), the serving cell switches from Cell 1 to Cell 2 via handoff (③).

The handoff procedure can be divided into three phases: *pre-handoff, handoff* and *post-handoff*. The *pre-handoff* phase decides whether to trigger a handoff, depending on user mobility, radio quality variation, load balancing, *etc.*. In the above example, the serving cell asks the phone to measure radio quality (defining measurement parameters and criteria that trigger reports) and invokes a handoff decision upon receiving the radio quality report from the phone. Afterwards, the serving cell requests a handoff to the target cell and performs admission control. Once accepted, the handoff request is acknowledged by the target cell. The serving cell executes the handoff by sending a handoff command to the phone. The phone changes its radio configuration (matching with the target cell) accordingly and responds with a handoff confirmation message to the target cell. In this process, the user traffic is still delivered via the original cell (likely with poor performance) until the handoff completes. In the final *post-handoff* phase, it performs *location update* and reconfigures the data/voice forwarding path if the new cell belongs to a different location area. This is to let the cellular network learn the current location of the phone. During this phase, it may also release resources at the source cell, update QoS profiles and the IP address, perform authentication *etc.*, depending on the handoff type. Finally, the phone continues its traffic delivery through the new cell. In reality, various handoffs take place (see Table 1) and their detailed procedures might vary. However, they all require the trigger-and-decision process to prepare for a handoff and perform multiple-round signaling message exchanges to execute the handoff.

MM-related procedures. There are several procedures related to MM. Table 1 lists the main procedures and their standard specifications, covering *initial attach, cell (re)selection, active handoff, voice support via CSFB and SRVCC, offloading, load balancing* (*e.g.*, via self-organizing networks). Each works with certain radio access technology (RAT, say, 4G/3G/2G), and/or various service types (say, active data/voice/both or idle).



(a) A handoff example (b) A basic handoff flow

Figure 1: Illustration of a handoff and its procedure flow.

Procedure	Standard	RAT	Service
Initial attach	23.401 [3]	all	idle
Cell (re)selection	25.304 [4],36.304 [5]	all	idle
Active handoff	23.009 [6]	all	active
CSFB and SRVCC	23.272 [7],23.216 [8]	4G	active(voice)
Femtocell offloading	25.367 [9]	3G,4G	active & idle
WLAN offloading	23.261 [10]	3G,4G	active & idle
Load balancing	32.500 [11]	all	active

Table 1: Main procedures (related to MM) in 3GPP standards.

Specifically, the *initial attach* and *cell-(re)selection* procedures are used to look for a serving cell or another better cell when the device has no active association with the serving one (idle). They are performed regardless of whether mobility is involved or not. The *initial attach* procedure is used to establish an association with a serving cell when the device just powers on or recovers from the out-of-service state (*e.g.*, the airplane mode). The *cell reselection* is used to switch its association when the device camps on a serving cell but has no active connectivity. In both idle cases, the handoff decision and execution are made by the user device. The decision is mainly based on the measured radio quality from different cells, the cell preference and radio evaluation criteria preconfigured by the device or reconfigured by the associating cell (*cell-reselection* only). The device receives configurations and commands over the broadcast channel in the current cell.

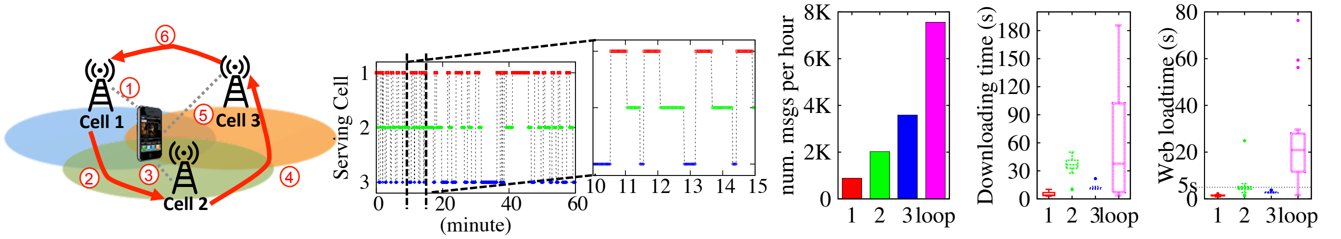
The *active handoff* procedure[1] regulates the cell switching for ongoing services, and its primary goal is to ensure seamless services. It exhibits many forms, including inter-RAT handoff (*e.g.*, 4G↔3G) and intra-RAT handoff (*e.g.*, within 4G), soft handoff (with simultaneous connectivities to multiple cells) and hard handoff (disconnect-and-connect). Moreover, cellular networks also support handoff for different purposes. For instance, 4G LTE leverages 3G/2G systems to carry voice through CSFB (Circuit Switched Fallback) and SRVCC (Single Radio Voice Call Continuity), thus invoking 4G↔3G/2G handoffs.

To enable opportunistic wireless access, the cellular network may offload traffic to small cells or user-deployed femtocells. It also allows for traffic redirection to different cells for load balancing or other carrier-specific optimizations. In these cases, both the user and the network are involved. They use different decision criteria based on many factors, such as radio quality evaluation threshold and cell preference, runtime traffic load, service type, and so on. These criteria and factors are not necessarily regulated by standards, but can be customized by carriers. However, the active handoff decision is fully controlled by the network, particularly via the serving cell.

3. A MOTIVATING EXAMPLE

We motivate our work with a real-world example. The discovered persistent loop differs from the transient ping-pong effect, which oscillates between cells due to frequent movement and wireless channel dynamics. As a matter of fact, the instability prob-

[1] We use "*active handoff*" to differentiate it from the case of switching the serving cell without active services.



(a) A handoff loop example (b) 1-hr log of serving cells (c) Signaling load (d) 5MB download (e) Web loadtime

Figure 2: A persistent handoff loop among three cells and the consequent overhead and performance degradation.

lem is caused by policy and configuration conflicts rooted in MM. That is, even given an invariant setting, the handoff process still never converges. Instead, inconsistent handoff decisions are made in turn, and the serving cell consequently oscillates among multiple cells under the invariant setting.

Figure 2a illustrates the example. The phone is placed at a spot covered by three cells, but repetitive handoffs (Steps 2, 4 and 6) are triggered in turn once the phone switches the serving cell. As a result, the phone oscillates among three cells. In our experiment, we place the phone in the idle mode (no voice/data) for 40 hours at this location. We record the network status (serving cell ID and RAT) per second. The loop repeats every several minutes (see 1-hour trace in Figure 2b), and it does not converge during the 40-hour test. We further test different phone models (Samsung S4/S5 and LG Optimus G), and verify that the finding is independent of phone models. Note that such oscillations are not caused by radio signal variations. The loop still exists even in an ideal scenario without any channel or traffic dynamics.

It turns out that, this persistent loop is caused by conflicting handoff configurations among different cells. In this example, Cells 1, 2, 3 are a 4G cell, a 3G femtocell and a 3G cell, respectively. Cell 2 is deployed by users while the other two cells are deployed by carriers. The carrier aims to offer high-speed access *and* balance the traffic load. This can be realized through configuring the MM preferences at different cells. In the example, Cell 1 believes that Cell 2 has a higher preference to itself for the offloading purpose. Cell 2 configures equal preference to all its neighboring cells and selects the one with strongest radio coverage. Cell 3 (3G) always prefers Cell 1 (4G) for its high-speed data service, as long as the Cell 1's radio signal is not weak. Unfortunately, these *independent* preference settings at different cells lead to inconsistent results. When Cell 3 has stronger coverage than Cell 2, it results in the persistent loop $c_1 \rightarrow c_2 \rightarrow c_3 \rightarrow c_1 \rightarrow \cdots$.

Such a loop is undesirable, and it does hurt both the carrier and the user. Without converging to any cell, the mobile carrier fails to achieve the expected goals. Our 40-hour test further shows that more than 90% of loops can be formed within 200 seconds. That is, three handoffs approximately take place every three minutes. With such high-frequent switches, it fails to offer high-speed 4G access or achieve cost-effective offloading to the Femtocell. Even worse, it incurs a large amount of signaling overhead between the device and the network. Figure 2c compares the incurred signaling messages per hour with the case using each cell only. On average, this loop incurs 7555 signaling messages per hour, 8.5, 3.5, 2.2 times over those only using Cells 1, 2, 3 respectively. Finally, this results in data performance degradation. Data transfer speed decreases and the response is delayed. In addition to the experiment in the idle mode, we load a small webpage (www.cnn.com, interactive) and download a 5MB file to assess the negative impacts. We choose these two representative apps, since they take both the access speed and the response time into account. The results are similar to those running other apps and speedTest. Figures 2d and 2e show the boxplots of their (down)loading times. In the worst case, it takes at most 12 seconds to download a 5MB file and about 3 seconds to load the web page using 4G; However, the current practice takes 180 seconds and 76 seconds. It suffers the 10-fold slowdown (15 fold in the worst case) in file downloading, and large performance slump in web browsing (11x on average, 33x in the worst case).

We further examine why handoff instability incurs these negative effects. As described in §2, each handoff execution requires signaling message exchanges and it takes time to get ready to serve the mobile device using the new cell. Even worse, *location update* is mandatory in most cases. It has to add multi-round message exchanges related to radio resource allocation, data forwarding-path reconfiguration and authentication (see the standard TS24.008) We find that a *location update* typically takes 3–6 seconds in the absence of failures. This matches previous studies (*e.g.*, [12, 13]). Frequent location updates intermittently suspend traffic delivery, thus incurring significant delay, loss and throughput slump.

4. THE INSTABILITY PROBLEM IN MM

We now formulate the instability problem of mobility management in cellular networks. We look into under what conditions unstable handoffs occur in reality, particularly those caused by uncoordinated policy conflicts. We focus on the persistent loops rather than those transient ones (*e.g.*, ping-pong effects), because they have lasting negative impacts and can be prevented with appropriate mobility management. We examine the trigger-and-decision phase, while assuming that the handoff execution exactly follows its decision.

We model a handoff procedure as a transition from the serving cell s to the target cell t out of the available candidate set C: $s \rightarrow t, t \in C$. We define the decision function as $t = F_s(s, C)$. The handoff decision depends on the criteria used by the serving cell or the mobile device, as well as the neighboring cell measurement performed at the mobile device (but configured by the serving cell). For simplicity, we assume the environment is invariant. Consider the same device is used in all the decision functions, we simplify it as $t = F_s(s)$. Once the serving cell switches, the decision criteria and measurement will change accordingly. Finally, it can be expressed as a deterministic process

$$s \rightarrow F_s(s) \rightarrow \cdots c_i \rightarrow [c_{i+1} = F_{c_i}(c_i)] \rightarrow \cdots, c_i \in C.$$

We claim that, stability is guaranteed if the handoff process always converges to the target t, regardless of its initial value s. If this property is violated, a *persistent loop* happens within a subset of candidate cells. Note that, our work focuses on whether it converges, and does not discuss how long it takes to converge. We now give necessary and sufficient conditions for stability.

THEOREM 1. **[Necessary condition]** *There exists at least one cell who allows a handoff decision to itself, namely,* $\exists t \in C$, $t = F_t(t)$.

PROOF. This is proved by contradiction. Assume that it converges to t when no cells satisfy $c = F_c(c), c \in C$. Given the serving cell t, its next cell is not t. It leads to contradiction. □

THEOREM 2. **[Necessary and sufficient condition]** *It converges, if and only if (1) there exists at least one cell specified in Theorem 1:* $\exists t \in C$, $t = F_t(t)$*; (2) there exists a handoff path from the initial cell s to the desirable cell t.*

PROOF. We only need to prove that they are sufficient. Assume there exists a handoff path from $s \rightarrow \cdots \rightarrow t$. Follow this path, it converges to t since $t = F_t(t)$. So the handoff process converges. □

Bearing these stability conditions in mind, we next devise an automatic detector to infer possible instability.

5. MMDIAG DESIGN

We design MMDIAG, an in-device diagnosis tool to detect and validate instability in MM. We take the device-based approach, since the carriers are reluctant to provide public access to their mobility management configurations and runtime information for handoff decisions. Our approach is deemed a viable solution, because we can leverage the signaling exchanges to bypass this major constraint. The underlying premise is that, the serving cell has to send their main parameters and decision logics to the device.

Inspired by this, we design MMDIAG as follows. Figure 3 plots its architecture, which is divided into two phases: detection and validation. In the detection phase (left), the core is an MM automata, which explores possible instability cases through an instability analyzer and reports counterexamples if found. It models the MM decision logic based on the 3GPP standards and feeds this model with real configurations collected directly from the device and indirectly from the serving cell, as well as dynamic environment settings created for various scenarios. The instability is inferred through examining two instability conditions given in Theorems 1 and 2. Once they are found, we move to the device-based validation phase (right). For each counterexample, we set up the corresponding experimental scenario and conduct measurements in operational networks for validation. We next elaborate on each component.

Instability analyzer. The key is to model the decision process in MM. This model determines the target cell using three factors: the *decision logic*, the *configurable parameters* and the *runtime observations*.

The decision logic is the algorithm to select the target cell, represented by F_s. We support the standard procedures specified in Table 1 and extract their logic engines from their specifications. For instance, *cell reselection* selects the one with the strongest radio coverage among those most preferable cells. That is,

$$F_s(s, C) = \arg\max_{c \in C'} radio(c),$$

$$C' = \{c | prefer(c) = \max_{i \in C} perfer(i), c \in C\}.$$

The configurable parameters are used to feed the logic. They are defined in the standards but can be customized by carriers and vendors. Table 2 summarizes related configurations specified in the standards. In the above example, cell preference $prefer(c)$ maps to a priority value, CELLRESELPRIORITY, which can be directly obtained from the handoff request message.

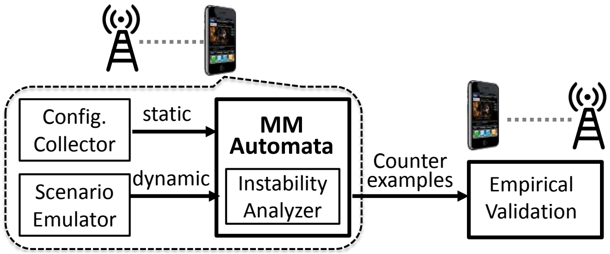


Figure 3: The MMDIAG architecture.

	Category	Parameter	State	Description
Net-work	Candidate cells	$MeasObj$	Active	Cells to be monitored
		$carrierFreqList$	Idle	Frequencies to be monitored
	Access control	*Handover restriction list*	Active	List of forbidden target cells
		Closed Subscriber Group	Idle	List of users with the cell access
	Radio evaluation	$eventA1 \sim eventA5$ $eventB1 \sim eventB2$	Active	Thresholds and report event criteria
		$Thresh1 \sim Thresh3$	Idle	Thresholds for cell re-selection
		$TimeToTrigger$	Active	Measurement duration for each cell in the active mode
		$T_{reselection}$	Idle	Measurement duration for each cell in the idle mode
	Traffic eval.	$event4A$ and $event4B$	Active	threshold for users' traffic volume report
	Cell preference	$cellReselPriority$	Idle	Cell reselection priority
		$SPID$	Active	Subscriber ID for RAT/Freq. priority
	Mobility method	$InactivityTimer$	Active	Timer for active→idle state transition
Device	Radio	Network mode	both	Frequency bands to be enabled
	Operation mode	Usage setting	both	Voice-centric or data-centric
		Voice preference	both	indicate if preferring PS or CS voice

Table 2: Summary of standardized configurations related to mobility management.

The runtime observations serve as the input to the handoff decision. The idle-state handoff adopts the cell-radio-quality assessments as the input. The active-state handoff uses both the radio quality and customizable observations (*e.g.* cell loads), which are fed through the *configuration collector* and the *scenario emulator*.

To infer whether the handoff converges, the Instability Analyzer first checks the necessary condition (Theorem 1). If no cell satisfies Theorem 1, it directly reports an instability counterexample with all the configurations and runtime measurements. Otherwise, we proceed to check the sufficient condition (Theorem 2). For each cell, we enumerate the possible paths. Note that this simple scheme may not be scalable; it can be improved as part of our future work.

Configuration collector. We collect surrounding cells' handoff policies and configurations from the signaling messages sent by the serving cell. We retrieve configuration parameters through the mapping defined by the standards. To collect signaling messages, we enable the diagnostic mode (*e.g.*, dialing secret code *#0808# for Sumsang Galaxy S5) at the mobile device. We log signaling messages through MobileInsight [14], an in-device cellular signaling collector developed by us. This acts like QXDM [15] and XCAL [16], proprietary software used by professionals to record message exchanges over the air. To collect a complete set from all cells, the device proactively switches to every 3G/4G cell at each location. Given each cell, we collect handoff parameters (see Table 2) from the radio resource control (RRC [5, 17]) layer and mobility management layer (MM [18, 19]).

Scenario emulator. Based on the handoff decision logic, we create runtime scenario parameters (*e.g.*, radio signal strength and traffic loads) and feed them into the MM automata. This is not easy because these parameters are not formally defined by the 3GPP standards, but largely dependent on carrier requirements and user demands. Testing all combinations is neither feasible nor necessary. For scenarios with an unlimited number of options (*e.g.*, user mobility at various speeds, traffic arrival patterns), we take the random sampling approach. We first retrieve the configuration values and then divide the runtime value range into several ranges. We assign each usage scenario with certain probability, and randomly sample values within these ranges. For the scenarios with a limited number of options (e.g., device switch on/off, data/voice service), we enumerate all possible combinations. This approach is similar to our prior studies [12, 20]. However, it differs from the previous work in that we examine configurations and policies on the management plane, whereas the prior work examines protocol interactions on the control plane.

Empirical validation. Using counterexamples as the input, the validation phase needs to construct test scenarios, run experiments, collect real traces, and confirm whether a loop appears. The real challenge is to *precisely* re-create the counterexample scenario. For example, MMDIAG infers that one persistent loop would incur as long as Cell 1 is stronger than -108 dbm and 3 dbm stronger than Cell 2 (see the example later). However, in reality, it is not easy for us to find such a location. To this end, we pre-collect a radio map through extensive measurements in indoor and outdoor testbeds. We further use them as hints to approximately locate the spots of our interests.

6. PRELIMINARY RESULTS

We conduct experiments in two metropolitan areas in Los Angeles and Columbus using a top-tier US carrier. We run both outdoor and indoor experiments. The outdoor experiments cover 63 different locations over 240 km^2 in LA and 260 km^2 in Columbus. Each location is separated by at least two kilometers apart, to obtain diverse cell coverage. We also collect information on indoor experiments at 50 spots in an 8-floor office buildings and an apartment, respectively. In this indoor setting, 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 the office and at home for indoor tests. We use four Android phone models: Samsung Galaxy S4, S5 and Note 3, and LG Optimus G. The results are similar for all phone models. Our dataset confirms that today's cell deployment is dense and hybrid. At most locations, there are about 8–16 cells available (11 cells on average).

MMDIAG reports 17 types of conflicts that might cause loops at the idle state and one type of loop at the active state. They are all validated in real experiments. Figure 4 summarizes the loops at the idle state. The smallest loop involves 3 cells, while the largest one has 7 cells. These happen when they use various RATs (4G, 3G, 2G) or different frequency bands[2]. Furthermore, they can be classified into three categories: 4G-Femtocell-3G loops (8 types), 4G-Femtocell-3G-2G loops (8 types), and 4G-only loop (1 type). Note that the the specific deployment location of the femtocell does not affect the discovery of loops involving the femtocells. Our outdoor tests confirm that, all 2G/3G/4G Macrocells have the problematic configurations, and a potential loop might exist as long as a Femtocell was deployed at the spots. We further test with femtocell deployment in campus buildings, and conduct indoor experiments at all viable locations. Among the tests, 25% of locations incur loops. Based on the root causes, these loops can be further classified in three categories:

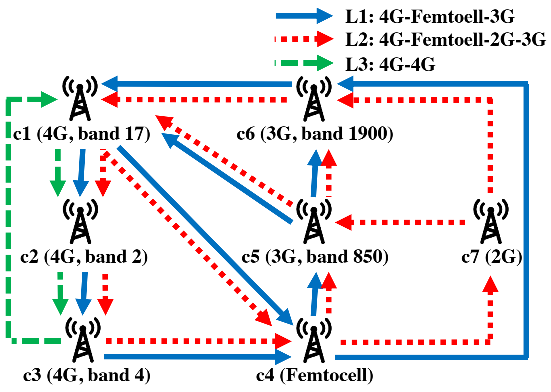


Figure 4: Idle-state persistent loops detected in US-I.

C1: Uncoordinated handoff goals. In this category, 8 variants of loops are reported, and all happen between 4G Macrocell, Femtocell and 3G Macrocells. The example in §3 illustrates the smallest loop, with c_1 = 4G, c_2 = Femtocell and c_3 = 3G. These loops are caused by conflicting preference settings for conflicting goals: the 4G Macrocells intend to offload user to the private Femtocells, but 3G Macrocells prefer to move the user to the high-speed 4G cell.

C2: Device-side preference misconfiguration. MMDIAG further reports 8 variants of loops between 4G Macrocells, Femtocell, 2G and 3G Macrocells. Compared with C1, when leaving the Femtocell, the mobile device hands off to 2G first, then switches 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 additional handoff is caused by improper preference configurations at the mobile device. With weak signal strength, the device may temporarily lose association to the Femtocell. According to the standards [5], the device resumes the service by scanning all cells and associating to the first available one. The order of the scanning cells is based on a pre-configured preference list stored at the phone's SIM card. For certain phones, 2G is listed as the highest preference, so the phone moves to 2G instead of 3G. Once associated with 2G, the device would immediately switch to 3G, 4G and 3G Femtocells. This way, the persistent loop continues.

C3: Imprudent 4G infrastructure upgrade. The last instance is a 4G-only loop. We observe that, US-I is upgrading its 4G infrastructure and deploying cells over a new frequency band (c_2 in Figure 4). Before the upgrade, existing 4G cells (c_1 and c_3) assign equal preferences to each other. US-I intends to migrate users to the new cells, which offer higher speed. To this end, some old cells (c_1) configure the new cells with higher preference. However, not all cells' preferences are updated in a timely fashion: preference ties still happen on some cells (c_2). Such partial upgrade fails to migrate the user to the new cells. This loop has no direct impact on users, because all cells belong to the same location area. However, it incurs larger 4G-Femtocell-3G and 4G-Femtocell-2G-3G loops, and indirectly amplifies their negative impacts.

C4: Uncoordinated load balancing. MMDIAG reports one loop between two 4G cells at one location. Both cells try to offload the user to each other when both signal strengths are higher than a threshold (here, -106 dBm). However, such load-balancing policies are not coordinated, so the user oscillates between cells when both cells' signal strengths are higher than -106 dBm. Fortunately, this loop is not commonly observed. Among all 4G cells we have collected, 67% of them use the same policy for the active-state handoff, but its neighboring cells are not observed to use the same rule

[2] https://en.wikipedia.org/wiki/List_of_LTE_networks

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.

7. RELATED WORK

In recent years, mobility management has been well examined in the context of cellular networks. These studies focus on handoff optimization [1,2,21], offloading [22,23], TCP/app performance [24] and cross-layer optimization [25,26]. In contrast, we demonstrate that improper interplay between handoff policies and misconfigurations can lead to instability. Our work is inspired by our prior efforts on the control-plane protocol verification [12, 20], but focuses on the management plane rather than the control plane.

Instability and policy misconfigurations have been studied in other problem contexts, including BGP routing divergence [27], DNS [28], home networks [29], and data center networks [30], *etc.*. Our work complements these efforts, but applies domain-specific analysis to study (in)stability in mobility management.

8. DISCUSSION AND CONCLUSION

Mobility support offers an indispensable utility function in 2G/3G/4G cellular networks. However, its management is more complex than expected. In practice, it allows for customizable policies and configurations at each cell and each device, to accommodate diverse demands from carriers and users. In this work, we conduct the first study to look into its persistent instability problem. We propose a device-based methodology to detect possible loops and validate them through real experiments. We show that, persistent loops may occur without proper parameter configurations or/and coordinated decision logics. They can result in heavy signaling overhead and significant performance degradation.

This work is still at its early stage. Several issues remain to be explored. First, we focus on the stability property only. It can not cover all desirable features; Other structural properties can be violated along with policy conflicts and misconfigurations. For example, the handoff converges to an undesirable choice (*e.g.*, 3G/2G even when 4G is available [13]). Second, we look into deterministic factors and do not take transient factors (*e.g.*, channel or traffic dynamics) into consideration. In reality, Non-deterministic factors need to be accounted. It is thus more challenging for stability analysis. Third, our approach may fail to identify all the unstable cases without all essential information on the network side. It thus calls for a holistic approach with cooperations from both parties. Fourth, the current work focuses on instability *only within* cellular technologies. Loops between different radio technologies (say, between WiFi and cellular networks) may exist, due to the offloading criteria between WiFi and cellular, similar to mobility management policies and configurations. Last but not the least, we focus on detecting the loops, but not fixing them. The future work is to sketch a solution that facilitates to both detect and fix the problems in MM. To this end, `MMDIAG` can be used to report identified problems to carriers. It can also assist end devices to break loops when they occur. Given hints of possible loops, we need to further check runtime measurements and detect whether a loop occurs. As long as the device confirms that the loop is caused by policy conflicts and/or misconfigurations, it take actions to intervene the loop; For instance, in the motivating example (§3), it can move to 4G and disable the path to the 3G Femtocell (thus hiding the existence of 3G femtocells). This prevents the further occurrence of loops.

Acknowledgments

We greatly appreciate the insightful and constructive comments from our shepherd, Dr. Aaron Schulman, and the anonymous reviewers. This work is supported in part by the National Science Foundation under Grants No. CNS-1423576, CNS-1421440, CNS-1526456 and CNS-1526985.

9. REFERENCES

[1] C. Brunner, A. Garavaglia, M. Mittal, M. Narang, and J. V. Bautista. Inter-system Handover Parameter Optimization. In *VTC Fall*, 2006.

[2] 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.

[3] 3GPP. TS23.401: GPRS Enhancements for E-UTRAN Access, 2011.

[4] 3GPP. TS25.304: User Equipment (UE) Procedures in Idle Mode and Procedures for Cell Reselection in Connected Mode, 2012.

[5] 3GPP. TS36.304: E-UTRA; User Equipment Procedures in Idle Mode, 2015.

[6] 3GPP. TS23.009: Handover Procedures, 2011.

[7] 3GPP. TS23.272: Circuit Switched (CS) fallback in Evolved Packet System (EPS), 2012.

[8] 3GPP. TS 23.216: Single Radio Voice Call Continuity (SRVCC), 2011.

[9] 3GPP. TS25.367: Mobility procedures for Home Node B, 2014.

[10] 3GPP. TS23.261: IP flow mobility and seamless Wireless Local Area Network (WLAN) offload; Stage 2, 2014.

[11] 3GPP. TS32.500: Self-Organizing Networks (SON); Concepts and requirements, 2014.

[12] G.-H. Tu, Y. Li, C. Peng, C.-Y. Li, H. Wang, and S. Lu. Control-Plane Protocol Interactions in Cellular Networks. In *SIGCOMM*, 2014.

[13] G. Tu, C. Peng, H. Wang, C. Li, and S. Lu. How Voice Calls Affect Data in Operational LTE Networks. In *MobiCom*, Oct. 2013.

[14] Mobileinsight project. http://metro.cs.ucla.edu/mobile_insight.

[15] QUALCOMM eXtensible Diagnostic Monitor. http://www.qualcomm.com/media/documents/tags/qxdm.

[16] Mediatek. Xcal-mobile. http://www.accuver.com.

[17] 3GPP. TS36.331: E-UTRA; Radio Resource Control (RRC), 2012.

[18] 3GPP. TS24.008: Mobile Radio Interface Layer 3, 2012.

[19] 3GPP. TS24.301: Non-Access-Stratum (NAS) for EPS; , Jun. 2013.

[20] G.-H. Tu, Y. Li, C. Peng, C.-Y. Li, and S. Lu. Detecting problematic control-plane protocol interactions in mobile networks. *IEEE Transactions on Networking (TON)*, pages 1–14, March 2015.

[21] 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.

[22] A. Balasubramanian, R. Mahajan, and A. Venkataramani. Augmenting mobile 3g using wifi. In *ACM MobiSys*, June 2010.

[23] 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.

[24] 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.

[25] N. Balasubramanian, A. Balasubramanian, and A. Venkataramani. Energy consumption in mobile phones: A measurement study and implications for network applications. In *IMC*, 2009.

[26] U. Javed, D. Han, R. Caceres, J. Pang, S. Seshan, and A. Varshavsky. Predicting handoffs in 3g networks. In *MobiHeld*, 2011.

[27] T. G. Griffin and G. Wilfong. An Analysis of BGP Convergence Properties. In *ACM SIGCOMM*, 1999.

[28] V. Pappas, Z. Xu, S. Lu, D. Massey, A. Terzis, and L. Zhang. Impact of Configuration Errors on DNS Robustness. In *SIGCOMM*, 2004.

[29] B. Aggarwal, R. Bhagwan, T. Das, S. Eswaran, V. N. Padmanabhan, and G. M. Voelker. NetPrints: Diagnosing Home Network Misconfigurations Using Shared Knowledge. In *NSDI*, 2009.

[30] P. Sun, R. Mahajan, J. Rexford, L. Yuan, M. Zhang, and A. Arefin. A Network-State Management Service. In *ACM SIGCOMM*, 2014.