Dependent Misconfigurations in 5G/4.5G Radio Resource Control

ZHEHUI ZHANG*, University of California, Los Angeles, USA
YANBING LIU*, Purdue University, USA
QIANRU LI, University of California, Los Angeles, USA
ZIZHENG LIU, Purdue University, USA
CHUNYI PENG*, Purdue University, USA
SONGWU LU, University of California, Los Angeles, USA

In this paper, we study an important, yet unexplored problem of configuration dependencies in 5G/4.5G radio resource control (RRC). Different from the previous studies in 3G/4G networks, 5G/4.5G allows more than one cells to serve a mobile device, resulting in more configuration dynamics and complexity that vary with all the serving cells. We analyze inter-dependency among configurations, categorize dependent misconfigurations, uncover their root causes, and quantify negative performance impacts. Specifically, we formulate configuration updates into a delta state machine (DSM) and unveil two types of dependent misconfigurations among states (inter-state) and within a state (intra-state); They stem from structural dependency and cross-parameter dependency. We further show that such misconfigurations incur service disruption and performance degradation. Our findings have been largely validated with three US operators and one Chinese operator; Our study has uncovered 644 instances of problematic dependencies.

CCS Concepts: • **Networks** → **Mobile networks**; **Network resources allocation**; • **Computer systems organization** → *Availability*.

Additional Key Words and Phrases: 5G, Radio Resource Control, Configuration, Dependent Misconfiguration

ACM Reference Format:
Zhehui Zhang, Yanbing Liu, Qianru Li, Zizheng Liu, Chunyi Peng, and Songwu Lu. 2023. Dependent Misconfigurations in 5G/4.5G Radio Resource Control. *Proc. ACM Netw.* 1, 1, Article 2 (June 2023), 20 pages. https://doi.org/10.1145/3595288

1 INTRODUCTION

Radio resource control (RRC) is responsible for controlling and managing radio access in a cellular network, thus critical to network performance and mobility support that a mobile device gets. In a cellular network, a cell is a *basic* unit to offer radio access over a contiguous radio frequency spectrum block (say, a channel). For an active radio access, the core of RRC is to select or re-select a group of serving cells (out of many candidate cells deployed in close proximity) to establish or maintain seamless radio access to the mobile device, no matter where the device is or goes.

*Z. Zhang and Y. Liu are co-primary authors. C. Peng is the corresponding author.

Authors' addresses: Zhehui Zhang, University of California, Los Angeles, Los Angeles, California, USA, 90095; Yanbing Liu, Purdue University, West Lafayette, Indiana, USA, 47907; Qianru Li, University of California, Los Angeles, Los Angeles, California, USA, 90095; Zizheng Liu, Purdue University, West Lafayette, Indiana, USA, 47907; Chunyi Peng, chunyi@purdue.edu, Purdue University, 305 N University Street, West Lafayette, Indiana, USA, 47907; Songwu Lu, slu@cs.ucla.edu, University of California, Los Angeles, 404 Westwood Plaza, Engineering VI, Los Angeles, California, USA, 90095.

2834-5509/2023/6-ART2
https://doi.org/10.1145/3595288

Category	Description of dependent misconfiguration	Negative impacts	♯	Quantitative damage
Structural dependency (inter-state, §5)	**D1** Unattended dependency between transition conditions	Persistent loop	486	up to 99% throughput drop, 1.8s disruption, 10-13x more signaling
	D2 Unnecessary dependency between states	SCells missed	109	12.8% to 87.6% throughput drop
Cross-parameter (intra-state,§6)	**D3** Unnecessary dependency during reporting	SCell removal	4	31% to 64.9% throughput drop
	D4 Unattended dependency during decision	Problematic handoff	44	38.7% to 91.2% throughput drop
	D5 Unattended dependency during measurement	Handoff failures	1	34.3% failure ratio increase

Table 1. Summary of main findings on two categories of dependent misconfigurations.

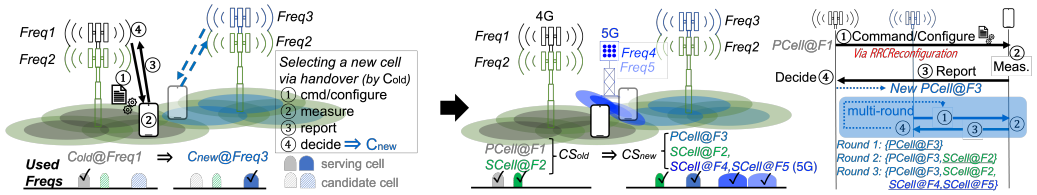
RRC configurations play an essential role in determining how and what cells are selected to serve mobile devices in practice. As a matter of fact, cell selection is realized through standard RRC procedures (e.g., handover) with the common mechanisms regulated by 3GPP specifications and the operator-specific policies configured with tunable parameters [10, 12]. These tunable parameters customize the subsequent operation criteria including whether to measure radio channel quality of neighboring cells, what cells to measure, whether and how to report the measurement results, and to name a few (detailed background in §**??**). If misconfigured, they result in degraded network performance in terms of reduced throughput, disrupted access, and oscillations in loops, as disclosed in prior 3G/4G studies [20, 30, 41, 42, 45–47].

In this work, we attempt to revisit RRC (mis)configurations in 5G/4.5G networks. Our study is fueled by one big advance from 3G/4G to 4.5G/5G, which allows *more than one cells* to simultaneously serve a mobile device while 3G/4G uses only *one* serving cell at a time. By increasing the number of serving cells from 1 to N ($N \geq 1$, mostly > 1), 5G/4.5G offers radio access over a much wider spectrum which aggregates all frequency channels used by all serving cells, thereby promising to enhance network performance [4, 7]. However, it is a double-edged sword because the complexity and dynamics of cell selection is multiplied with the increasing number of serving cells. Intuitively, cell selection in 3G/4G is often one-time effort which switches the serving cell from one to another (except uncommon handover loops). From 1 to N is a game changer. In 5G/4.5G networks, it might take several rounds, each of which determines a subset of serving cells (out of N cells selected eventually). A previous round impacts subsequent rounds as the configurations change with the serving cell(s) newly selected in the previous round. If ill-configured, radio access quality gets hurt.

In this work, we have identified a new type of misconfiguration, i.e., *dependent misconfiguration*, in networked systems that has never been reported in the literature. The key observation is that many 5G/4.5G configurations exhibit *inter-dependency*. For instance, before any 5G cell is used, RRC configurations are set to add 5G cells if their RSRPs (Reference Signal Received Power) are greater than -115 dBm; However, once 5G cells are added as the serving cells, the configurations are updated to remove 5G cells if their RSRPs are smaller than -96 dBm; There is no surprise that the radio access gets stuck into a loop where the RSRP of a 5G cell is between -96 dBm and -115 dBm (detailed in our real-world instance in §3.2). The loop repeatedly gets 5G and then quickly loses it. Such dependent configurations vary over time and with locations, which cannot be recognized by prior misconfiguration studies that assume static configurations per cell [20, 30, 41, 42, 45–47].

To examine problematic dependencies among varying configurations, we abstract the used configuration logic as a delta state machine (DSM), with a state transition denoting a one-time configuration update (§4). Each configuration update removes old configuration entries used at the previous state and adds new configuration entries at the new state. Such update is meaningful; It is explicitly associated with its transition trigger (mostly, measurement reports). By this means, DSM tackles the state explosion problem by decomposing a combination of all configuration entries into two parts: the dependent part impacted by the transition trigger and the independent one.

Given this modeling framework, we have deduced two categories of dependent misconfigurations, which stem from structural dependency (§5) and cross-parameter dependency (§6). These two categories are complementary to each other; they cover dependencies among states (inter-state) and



(a) 3G/4G radio access via one serving cell (b) 5G radio access via N serving cells (a cellset of N cells)

Fig. 1. The number of serving cells grows from one to N (≥ 1) during the advance from 3G/4G to 5G networks.

within a state (intra-state). We further find that there exist two types of misconfigured dependencies: necessary yet unattended dependency and unnecessary dependency. Together, they yield five subcategories of dependent misconfigurations D1-D5 (Table 1). Structural dependency arises among different configuration states. It has two subcategories. D1 exhibits necessary, yet unattended dependency between triggering conditions, thus incurring persistent loops. D2 shows unnecessary dependency between states and reduces data throughput. In contrast, cross-parameter dependency exhibits among various correlated parameters within a single configuration state. In this category, we uncover three subclasses: unnecessary dependency due to shared configurations (D3), and necessary yet unattended dependency due to optional configurations (D4 and D5).

While our primary goal is to better understand such dependent misconfigurations intellectually, we further seek to validate their practical existence, as well as their impacts and prevalence in operational 5G/4.5G networks. Our empirical study over three US carriers and one operator in China yields 5 subcategories and 644 instances of problematic dependencies. Such dependent misconfigurations lead to severe performance penalties including throughput drop up to two order of magnitude and access disruption for seconds.

Release. Datasets and codes used in this study are available at [1].

2 BACKGROUND: FROM 1 TO N CELLS

In this section, we introduce necessary background on how radio access and RRC evolve from 1 to N serving cells. Table 7 (Appendix A) lists all 5G/4G acronyms used in the paper.

Radio access from 1 to N cells. In a cellular network, a cell acts as a *basic* unit to offer radio access. Each cell runs one radio access technology (RAT, say, 5G, 4G or 3G) over one frequency channel (say, Freq1, Freq2, $\cdots$, Freq5 in Fig. 1). It physically resides in a cell tower which accommodates a number of cells over various frequency channels and directional antenna.

Fig. 1 illustrates how 5G advances the number of serving cells from 1 to N ($N \geq 1$), using two technologies: carrier aggregation [2, 3, 7] and dual connectivity [4]. Carrier aggregation combines multiple cells of the same RAT, which was first introduced for LTE-Advanced (4.5G) [7]. Dual connectivity uses two RATs (here, 5G and 4.5G), where one RAT acts as the master anchor to establish and manage the radio connection and the other RAT offers secondary radio access [4]. Note that 5G works with 4.5G. In this paper, 5G = 5G/4.5G unless specified.

As a result, 5G uses a *set* of serving cells, rather than an individual cell. A serving cellset consists of one primary cell (PCell) and several secondary cells (SCells)[1]. PCell is mandatory and SCells are optional; PCell is responsible for establishing and managing the radio connection in the control plane (via RRC), while both PCell and SCells are aggregated for data transmission in the user plane.

RRC from 1 to N cells. Serving cells are selected or re-selected through standard RRC procedures (e.g., handover) with the mechanisms regulated by 3GPP specifications and the policies configured with tunable parameters [10, 12].

[1]In this paper, PCell is the primary cell of the master RAT; All the other serving cells are called SCells, including all the secondary cells of the master RAT and all the cells of the secondary RAT if dual-connectivity is used.

	Criteria	Description		Criteria	Description
M	$\{F_C\}$	Measurement over frequencies $\{F_C\}$: intra-freq, inter-freq and inter-RAT			
A1	$R_s > \Theta_{A1}$	Serving cell becomes better than a threshold Θ_{A1}	A2	$R_s < \Theta_{A2}$	Serving cell becomes worse than a threshold Θ_{A2}
A3	$R_c > R_s + \Delta_{A3}$	Candidate becomes Δ_{A3} offset better than PCell	A4	$R_c > \Theta_{A4}$	Candidate becomes better than a threshold Θ_{A4}
A5	$R_s < \Theta^s_{A5}$ & $R_c > \Theta^c_{A5}$	PCell becomes worse than a threshold Θ^s_{A5} and candidate becomes better than Θ^c_{A5}	A6	$R_c > R_s + \Delta_{A6}$	Candidate becomes Δ_{A6} offset better than SCell
B1	$R_c > \Theta_{B1}$	Inter-RAT candidate cell becomes better than Θ_{B1}	B2	$R_s < \Theta^s_{B2}$ & $R_c > \Theta^c_{B2}$	PCell becomes worse than a threshold Θ^s_{B2} and inter-RAT candidate becomes better than Θ^c_{B2}

Table 2. Main RRC configurations. R denotes RSRP or RSRQ; Θ for thresholds and Δ for offsets.

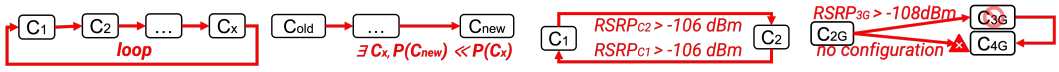
A significant change from 1 to N cells is that it might take multiple RRC rounds to switch to a new cellset (say, from CS_{old} to CS_{new} in Fig. 1b). Each round performs a basic *configuration-measurement-reporting-decision* procedure, similar to a handover procedure which switches the serving cell from one to another in 3G/4G networks (Fig. 1a). Each round relies on radio quality measurements to determine new serving cells (one, some, all, or zero) and configurations to update. Specifically, the PCell first sends its configurations to the device (step ①), mostly via `RRCReconfiguration`, a RRC signaling command to establish, modify or release the radio connection. It configures a few parameters (detailed in Table 2) to customize cell-specific operation criteria including whether and what to measure and report. When the configured criteria are satisfied, the device measures nearby cells (②) and reports their radio quality measurements (③). The PCell then runs its local policy to decide whether and how to change the serving cell(s) and/or update configurations (④).

Each round makes three possible decisions: (1) changing the PCell (with/without new SCells), (2) changing some or all SCells (without changing the PCell); (3) changing no cells. Configurations are updated accordingly. This process repeats until the serving cellset ends with a stable choice (here, from CS_{old} to CS_{new}). Stability will be formally defined by prior studies (§3.1). In this example, it takes three rounds to switch from CS_{old} to CS_{new}; Round 1 changes the PCell (but all SCells are removed), Round 2 adds 4G SCells (here, SCell@Freq2), and the last round adds 5G SCells (here, SCell@Freq4 and SCell@Freq5). RRC configurations also impact how many rounds are needed.

The maximum number of serving cells N_{max} is constrained by both network and device capabilities [3]. For instance, 3GPP Release-15 (the first 5G standard) initially supported up to four component cells over 5G mmWave and now up to eight 5G cells (8×100 = 800MHz); Additionally, Release-15 supports up to eight cells over 4G (total: 16 = 8+8). High-end 5G phones support advanced technologies and more serving cells. In this work, we use several phone models and support up to 8 serving cells (4+4 for 4G+5G, by Google Pixel 5).

The above process differs for an active or idle radio connection. When active, it repeats the above four steps at each round until it converges to a stable cellset or gets stuck into a loop. At the idle state, the device does not report the measurements (no PCell, step ③ skipped) and makes a decision locally. The idle-state device receives configurations broadcast by candidate cells (via `System Information Block`, RRC signaling messages specified in [10, 12]).

RRC configurations. Table 2 lists main configurations for measurement and reporting, as regulated by 3GPP [10, 12]. Each measurement that the device will perform is defined as *one measurement object*, which is added, removed or modified at the configuration step (①). Each object usually defines a list of cells or a list of frequency channels (all cells over the selected frequencies) to measure; This corresponds to one out of three measurement types: intra-freq (over the same frequencies used by the serving cells), inter-freq (over the different frequencies but using the same RAT) and inter-RAT (over frequencies for a different RAT). Each measurement object is associated with one or several report triggering events (say, `A1`–`A6`, `B1`-`B2`). Basically, reporting is triggered as long as the measured RSRPs/RSRQs of the serving/candidate cells satisfy the pre-configured event conditions. These conditions are customized by tunable thresholds (Θ_*) and offsets (Δ_*), which are used to compare RSRP/RSRQ measurements.



(a) Handover instability (b) Worse handover (c) An instance of (a) [30] (d) An instance of (b) [41]

Fig. 2. Two misconfiguration problems and their real-world instances reported in prior studies.

We note that RRC configurations are not static. With the same PCell, configurations are updated *dynamically* based on the serving cellset and environmental factors. For instance, it is configured to measure more cells (over more frequency channels) when the PCell's radio quality is much worse or no satisfactory measurements are reported. Once the PCell changes, configurations are all updated.

3 MISCONFIGURATION: FROM 1 TO N

In this section, we first introduce misconfigurations reported in prior 3G/4G studies, and then use a real-world instance to motivate our study on emerging misconfigurations in 5G networks.

3.1 Misconfigurations Reported for 3G/4G

Prior studies [20, 30, 41, 42, 45–47] have investigated the impacts of RRC misconfigurations in 3G/4G networks, where one serving cell is used at a time. They result in two problems: (1) *handover instability* and (2) *worse handover*, as illustrated in Fig. 2.

Handover instability means that given invariant measurements, the handover process fails to converge to one target cell. Instead, the serving cell persistently oscillates in a loop, say, $C_1 \Rightarrow C_2 \cdots \Rightarrow C_x \Rightarrow C_1$. It was first disclosed in a 3G/4G study with two US operators [30] and later extended to more use scenarios [42], more operators (operators in China) [47] and across multiple operators [46]. It is because each cell has its own configurations and makes its handover decision locally in a distributed manner. However, if improperly configured, all the transitions in the loop $C_1 \Rightarrow C_2 \cdots \Rightarrow C_x \Rightarrow C_1$ can happen when measurements do not change (more precisely, slight measurement fluctuations are allowed as long as they do not affect the measurement/reporting/decision criteria). For example, two cells switch to a target cell when A4 ($\text{RSRP}_c >$ -106dBm) is satisfied, as illustrated in Fig. 2c [30]. Such configurations result in a ping-pong loop when their RSRPs both are greater than -106dBm. Handover instability is harmful, because the serving cell frequently switches back and forth, resulting in huge signaling overhead and severe performance degradation [30, 42, 46, 47].

A worse handover is the one which does converge but the new serving cell performs much worse than another candidate cell available [20, 41, 45]. In particular, [45] discloses a special case where the new serving cell after handover performs even worse than the original one, namely, $P(C_{new}) < P(C_{old})$, where $P(\cdot)$ is a performance utility of the given serving cell; [20, 41] reveal a generic case where the new cell performs much worse than another candidate cell which is available but not chosen, namely, $\exists C_x, s.t. P(C_{new}) \ll P(C_x)$. This is because RRC configurations limit how a handover is performed; If ill-configured, they result in worse handovers which miss better cells. For example, the handover misses a 4G cell but uses a much slower 2G cell, as illustrated in Fig. 2d [41]. It turns out that this 2G cell does not configure any measurement of 4G cells; There is no way to directly switch from a 2G cell to a 4G cell. Instead, it must switch to a 3G cell first (when $\text{RSRP}_{3G} >$ -108dBm) and then to a 4G cell (when $\text{RSRP}_{4G} >$ -108dBm). This handover thus ends with a 2G cell when there are no such 3G cells, no matter how strong 4G cells are.

3.2 A Motivating Example in 5G

There is no surprise that the problems identified for 3G/4G are conceptually applicable to 5G/4.5G where a serving cell ($N = 1$) changes into a serving cellset ($N \geq 1$). Given a sequence of $CS_1 \Rightarrow CS_2 \cdots \Rightarrow CS_x$, it is prone to two same problems: (1) *instability* if the sequence fails to converge

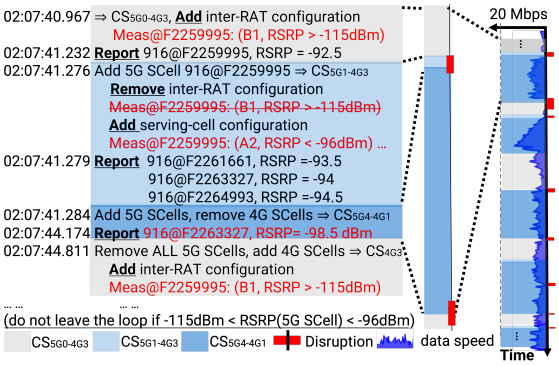


Fig. 3. Trace of a 4G-5G loop in a stationary test (US-I).

	ID@FChan	Downlink Freq	RSRP (dBm) Med (Min,Max)	RSRQ (dB) Med (Min,Max)
$PCell_{4G}$	21@F5110	739 MHz	**-80** (-84,-77)	**-9.5** (-14.5,-7)
$SCell_{4G1}$	323@F66486	2115 MHz	**-97** (-99, -93.5)	**-9** (-15, -7.5)
$SCell_{4G2}$	196@F66936	2160 MHz	**-93** (-96,-91)	**-10.5** (-11, -8.5)
$SCell_{5G1}$	916@F2259995	38.85 GHz	**-86.5** (-94.5, -83.5)	**-10.5** (-10.5, -10)
$SCell_{5G2}$	916@F2261661	38.95 GHz	**-95.5** (-98, -93)	**-10.5** (-10.5, -10)
$SCell_{5G3}$	916@F2263327	39.05 GHz	**-94.5** (-101,-92.5)	**-10.5** (-10.5, -10)
$SCell_{4G4}$	916@F2264993	39.15 GHz	**-92.5** (-97, -90.5)	**-10.5** (-10.5, -10)

$CS_{5G0-4G3}$:= { $\mathbf{PCell}_{4G}$, $SCell_{4G1}$, $SCell_{4G2}$ }
$CS_{5G1-4G3}$:= { $\mathbf{PCell}_{4G}$, $SCell_{4G1}$, $SCell_{4G2}$, $SCell_{5G1}$ }
$CS_{5G4-4G1}$:= { $\mathbf{PCell}_{4G}$, $SCell_{5G1}$, $SCell_{5G2}$, $SCell_{5G3}$, $SCell_{5G4}$ }

Table 3. Serving cells/cellsets seen in the trace.

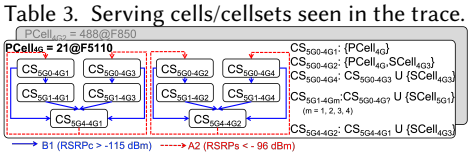


Fig. 4. Cellsets and loop variants in this example.

but oscillates in a persistent loop, and (2) *worse reachability* if the sequence does converge but ends with a poorly-performed choice in presence of much better choices.

We find that the problems are worse in 5G. In 3G/4G, misconfigurations exist mainly among different serving cells (aka, PCells). In 5G, new misconfigurations arise even when the serving PCell does not change. It is because 5G configurations are more complex and dynamic. A PCell updates its configurations based on the measurement reports from the device, which further impacts the measurement reports to generate in the next round. Such dependency, if ill-configured, raises negative impacts which have not been reported before.

We use a real-world instance to illustrate such dependence and emerging misconfigurations (Fig. 3). It is observed at a stationary test where the mobile device is placed at one fixed location and served by a US operator (US-I). The serving cellset gets stuck into a 4G-5G loop which repeatedly gets 5G and then quickly loses it. Fig. 3 shows a 4-second trace with $CS_{5G0-4G3} \Rightarrow CS_{5G1-4G3} \Rightarrow CS_{5G4-4G1} \Rightarrow CS_{5G0-4G3}$. Each cellset is denoted as $CS_{5Gn-4Gm}$, where n is the number of 5G cells and m is the number of 4G cells. Table 3 lists the serving cellsets and cells observed in this trace. All the cellsets use the same PCell, 21@F5110. In this work, each cell is represented by *ID@Frequency*; *Frequency* is marked by its frequency channel number (NR-ARFCN for 5G [11] and EARFCN for 4G [7]). Here, a frequency channel F5110 corresponds to a downlink frequency centered on 739 MHz for 4G; F2259995 is a 100 MHz channel centered on 38.85 GHz for 5G (more precisely, 5G mmWave). In this work, we observe that all three US operators launch 5G along with their existing 4G networks, primarily in a Non-Standalone (NSA) option where 4G acts as a master RAT and 5G offers secondary radio access [6]. NSA is indeed recommended to quickly deploy 5G at the start [13]. In this work, we study NSA for 5G unless specified. The implications to 5G Standalone (SA) are discussed in Appendix B. Table 3 shows measured RSRPs/RSRQs per cell.

We observe this loop repeats itself (with some variants explained later); There is no sign to stop in our 5-minute stationary test. Evidently, this loop is problematic because it hurts device performance and wastes resources. We run a speed test by downloading bulky files from Google Cloud. File download speed would reach 459 Mbps (median), if served by $CS_{5G4-4G1}$ only. However, the actual download speed during the loop is remarkably slower (two orders of magnitude); the median speed shrinks to 4.7 Mbps (all below 15 Mbps). We next unveil how this loop occurs.

- When the serving cellset switches to $CS_{5G0-4G3}$, the PCell updates its configurations by adding an inter-RAT measurement object over F2259995 (5G) and using B1 as its report triggering event (report B1 if $RSRP_c$ > -115dBm). This makes sense because 5G is not used and B1 is often used to report an inter-RAT candidate cell measurement (Table 2, [12]).
- Shortly, the device measures cells over F2259995 (here, 916@F2259995) and reports its RSRP.

• The PCell decides to switch the serving cellset to $CS_{5G1-4G3}$, which adds this reported 5G cell as a new SCell. It accordingly updates its configurations by removing the previous B1 event (for 5G candidate cells as 5G is not used) and adding new events (for 5G serving cells as 5G is used now). It actually adds three events (one A3 and two A2 events, shown in Fig. 5a) and Fig. 3 shows the only one which results in the loop: event A2 ($RSRP_s$ < -96dBm).

We note that 5G supports measurements over multiple adjacent channels although only one 5G frequency is explicitly configured for measurement [5, 12]. Here, the device measures over F2259995 first and then over three other channels (F2261661, F2263327, F2264993). It results in the following two transitions: $CS_{5G1-4G3} \Rightarrow CS_{5G4-4G1}$ and $CS_{5G4-4G1} \Rightarrow CS_{5G0-4G3}$.

• The serving cellset quickly switches from $CS_{5G1-4G3}$ to $CS_{5G4-4G1}$ as soon as the RSRPs of 5G candidate cells over other three channels are reported (within 10 ms).

• Event A2 is later triggered when the RSRP of one 5G SCell (here, 916@F2263327) is below -96 dBm; It takes several seconds (here, about three seconds) because the median RSRPs of all 5G cells are above -96dBm but they sometimes go below -96 dBm (Table 3). Upon receiving A2 ($RSRP_{5G}$ < -96 dBm), the PCell removes all 5G SCells and adds 4G SCells back, which switches back to $CS_{5G0-4G3}$. As 5G is not used, PCell updates its configuration by adding B1 for the inter-RAT measurement over F2259995. This way, the loop continually repeats itself as long as the RSRP of one 5G SCell goes below -96 dBm (here, the RSRPs of all 5G cells > -115 dBm).

We notice that the reality is more complex than this example. Several loop variants and more cellsets are observed in the same stationary test, as shown in Fig. 4. In total, we see ten cellsets using the same PCell (21@F5110), including four 4G-only cellsets ($CS_{5G0-4Gm}$, m = 1, 2, 3, 4) and four 5G+4G cellsets using one 5G SCell ($CS_{5G1-4Gm}$, m = 1, 2, 3, 4) and two 5G+4G cellsets using four 5G SCells. These cellset variants mainly depend on whether another 4G cell (155@F850) is used or not. Fig. 4 also shows the cellset transitions between the 4G-only cellsets and the 5G+4G cellsets. For simplicity, we do not plot all the cellset transitions such as those among $CS_{5G0-4Gm}$ (m = 1, 2, 3, 4) and among $CS_{5G1-4Gm}$ (m = 1, 2, 3, 4). We see that $CS_{5G1-4Gm}$ is sometimes skipped when the RSRPs of all four 5G cells are reported before a new decision is made. Note that $CS_{5G1-4Gm}$ is short-lived (for at most tens of milliseconds) in all the instances as the RSRPs of these four 5G cells are much larger than -115dBm. We see that this 4G-5G loop repeats itself every a few seconds, up to tens of seconds (about 4 seconds in Fig. 3). This actually depends on when the RSRP of one 5G cell goes below -96dBm. Unsurprisingly, the loop happens much more often when the test location moves to a nearby one with weaker 5G coverage (say, most RSRPs below -96 dBm and all above -115 dBm). All these cellset transitions not only depend on the configured criteria but also runtime measurements which fluctuate over time. Moreover, we see another 4G-5G loop with a different PCell (488@F850), as well as different SCells. This loop is also caused by the similar B1 and A2 events (with the same thresholds for distinct frequencies to measure).

Despite these variants, two essential problems remain the same. *First, it is unstable.* The device gets stuck into a persistent 4G-5G loop in a stationary test. *Second, this loop is caused by dynamic-yet-conflicting configurations by the same PCell* (switching between B1 and A2 events with two distinct thresholds). Such configurations make it possible to simultaneously satisfy multiple transitions needed for a loop in the unchanged environment with the same or similar measurements.

We want to highlight that this loop differs from the persistent handover loop reported in prior studies [30, 42, 47]. 3G/4G uses static configurations per PCell and the conflicting configurations occur among different serving PCells. From 1 to N ($N \geq 1$) is the game changer in 5G. There are many more serving cellsets with the same PCell. As the serving cellset changes, the PCell dynamically adjusts its configurations, which unfortunately opens up new room for various dependent misconfigurations if such dependencies are not well managed.

4 DEPENDENT MISCONFIGURATIONS

We now present our approach to uncovering dependent misconfigurations in 5G/4.5G networks.

4.1 Modeling Dependency with DSM

We first develop a formal model, termed as DSM, to examine dependencies among dynamic configurations. We use the above example to illustrate why and how we develop DSM.

CSM: Configuration State Machine. A naive approach is to follow the previous method proposed in [30] and use a finite state machine (FSM) to model all the possible cellset transitions. Each cellset transition from CS_k to CS_{k+1} is triggered by certain events R_k (mostly reporting events), which are associated with runtime configurations Ω_k.

$$\cdots CS_k \xrightarrow{\Omega_k \mapsto R_k} CS_{k+1} \xrightarrow{\Omega_{k+1} \mapsto R_{k+1}} CS_{k+2} \cdots \tag{1}$$

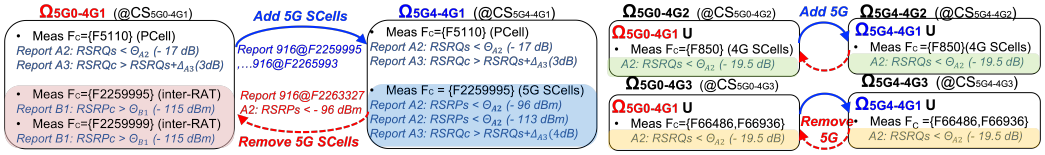
A cellset sequence can be converted into an equivalent chain,

$$\cdots \xrightarrow{CS_k} [\Omega_k \mapsto R_k] \xrightarrow{CS_{k+1}} [\Omega_{k+1} \mapsto R_{k+1}] \xrightarrow{CS_{k+2}} \cdots \tag{2}$$

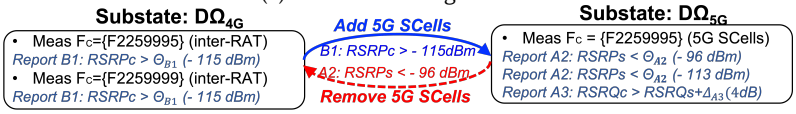
Note that $\mapsto$ and $\Rightarrow$ stand for two different types of causality. $\Omega_k \mapsto R_k$ means that configuration Ω_k results in RSRP/RSRQ measurements R_k; It also depends on channel conditions and runtime environments. $(\Omega_k \mapsto) R_k \Rightarrow \Omega_{k+1}$ is determined by the re-configuration logic used by the current PCell. Consequently, we generate a new CSM where each state is a combination of all configurations at a time $\{\Omega_k | k = 1, 2, \cdots\}$. We use the CSM to generate possible cellset sequences and check whether desirable properties (say, stability and reachability) are violated.

Fig. 5a plots the resulting CSM for the example presented in §3.2. For sake of simplicity, we do not plot all the configuration states but show six states to illustrate the downside of the CSM. These configuration states are associated with the serving cellsets, $\Omega_{5Gn-4Gm}$ for $\text{CS}_{5Gn-4Gm}$. Each configuration state is a set of multiple configuration entries. For instance, $\Omega_{5G0-4G1}$ (PCell only) has three measurement objects over F5110, F2259995 and F2259999. The intra-freq measurement over F5110 is associated with two reporting events: A2 (RSRQ_s <-17 dB) and A3 ($\text{RSRP}_c > \text{RSRP}_s$ + 3 dBm). These two events are used to monitor if the serving PCell is too weak (if A2 is satisfied) or there exists another candidate cell over F5110 3dB stronger than the current PCell (if A3 is satisfied). Because no 5G cells are used, it also configures two inter-RAT measurement objects over F2259995 and F2259999, both using B1 reporting events: B1 (RSRQ_c > - 115 dBm).

Once 5G cells are added (upon reporting B1), the configuration state is updated to $\Omega_{5G4-4G1}$, which removes both inter-RAT measurement objects and adds an intra-freq measurement over F2259995. The reporting events are updated accordingly with three new events: A2 (RSRP_s < -96 dBm), A2 (RSRP_s < -113 dBm), and A3 ($\text{RSRQ}_c > \text{RSRQ}_s$ + 4 dB). More configuration entries are added when 4G SCells are in use (see, $\Omega_{5G0-4G2}$, $\Omega_{5G4-4G2}$, $\Omega_{5G0-4G3}$ and $\Omega_{5G4-4G3}$).



(a) CSM over configurations



(b) DSM over configuration updates

Fig. 5. An illustration from CSM to DSM (using the example in §3.2).

Its downside is the state explosion problem with the resulting CSM when the number of serving cells grows. It is because each configuration state (a combination of multiple configuration entries) can be partly updated, thereby generating a large number of possible configuration states. There are many cellsets and loop variants, but the fundamental problem is caused by two configuration updates and their associated reporting events: (1) 5G SCells are added due to B1 ($RSRP_c$ > -115 dBm, blue solid lines), and (2) 5G SCells are removed due to A2 ($RSRP_s$ < -96 dBm, red dashed lines).

DSM. To overcome the state explosion problem, we model configuration updates into a delta-graph called DSM. We convert each configuration state transition

$$\Omega_k \overset{R_k}{\Longrightarrow} \Omega_{k+1} \quad \text{into} \quad D\Omega_k \overset{R_k}{\Longrightarrow} D\Omega_{k+1}, \tag{3}$$

where $D\Omega_k = \Omega_k - \Omega_{k+1}$ and $D\Omega_{k+1} = \Omega_{k+1} - \Omega_k$. It indicates that measurement reports R_k result in removing some configuration entries (here, $D\Omega_k$) and adding new configuration entries (here, $D\Omega_{k+1}$). By this way, we can decompose a configuration state Ω_k into two parts: the part impacted by the reporting events (here, $D\Omega_k$) and the irrelevant one ($\Omega_k \cap \Omega_{k+1}$). More importantly, such decomposition is meaningful, linking the delta-configuration substate directly with the reporting events. Fig. 5b plots the resulting DSM for the above example. We find that all the transitions from a 4G-only cellset to its corresponding 5G+4G cellset share the same configuration update; $D\Omega_{5G} = \Omega_{5Gn-4Gm} - \Omega_{5G0-4Gm}$ and $D\Omega_{4G} = \Omega_{5G0-4Gm} - \Omega_{5Gn-4Gm}$ are the same when n = 1,2 and m = 1, 2, 3, 4. We thus derive the following sequence,

$$D\Omega_{4G} \underset{RSRP_{5G}>-115dBm}{\overset{\text{B1}}{\Longrightarrow}} D\Omega_{5G} \underset{RSRP_{5G}<-96dBm}{\overset{\text{A2}}{\Longrightarrow}} D\Omega_{4G}. \tag{4}$$

The loop occurs when both triggers can be satisfied with -115 dBm < $RSRP_{5G}$ < -96 dBm.

DSM is effective to combat the configuration state explosion problem because the current configuration practice does combine multiple configuration entries, each of which is associated with one or some serving cells in use or the candidate cells to be considered. The compound configuration state is partly updated and its dependent configuration entries are completely updated. DSM decomposes a configuration state into multiple independent or dependent updates, greatly compressing redundancy among dynamic configurations.

With DSM, we can deduce two categories of dependency. The first category lies in the dependency between states, i.e. $D\Omega_i \Longrightarrow D\Omega_j$, which we term as *structural dependency*. The second is intra-state dependency, where parameters at any single configuration state affect the same output with $D\Omega_i \mapsto R_i$, which we term as *cross-parameter dependency*. Note that we generate a DSM per PCell (though some PCells have similar DSMs using the same events and parameters). We use CSM to model configuration transitions among the cellsets with distinct PCells.

4.2 Misconfiguration Analysis

We use the obtained DSM and CSM to proceed our dependency analysis for misconfigurations. We next examine their impacts on *stability* and *reachability* to identify misconfigurations among such dependencies. It is treated as one dependent misconfiguration instance when such dependencies in RRC configurations result in handover instability or worse handover.

Structural dependency analysis. For structural dependency, we examine how inter-state dependency affects transitions with DSM. We define misconfigurations induced by inter-state dependency in two forms: loops and worse reachability. For loops, we check whether the device could get stuck in a persistent loop, which would induce service disruption. With DSM, we search for loops and derive the conditions of making a loop. Note each transition in a DSM is accompanied with a triggering condition that uses one of the configuration types listed in Table 2. We thus record the sequence of triggering conditions and terminate the search whenever the criteria cannot hold

true simultaneously. For example, if we find event A1 with $\text{RSRP}_s > \Theta$ followed by event A2 with $\text{RSRP}_s < \Theta$, we will stop exploring any walk from A2, terminating the search for the loop.

For reachability, we examine whether the device reaches a non-preferred configuration state that causes nontrivial performance loss. Specifically, we define the preferred state based on the subset of SCells (namely, all the serving cells except PCell). State $D\Omega_i$ is *preferred* but $D\Omega_j$ is *non-preferred*, if the group of SCell(s) in the former state $D\Omega_i$ becomes a superset of the latter. The device should not consequently reach the non-preferred state whenever a preferred state exists in DSM; Otherwise, it poses a reachability issue that would cause non-trivial throughput drop and waste the potential of aggregating multiple serving SCells. For each state $D\Omega_i$, we take a transition walk from it and examine all the reachable states. We enumerate all possible transition walks that connect state $D\Omega_i$ and other states, and examine misconfigurations. We check whether different transition walks ensure reachability and whether any transition walk contains loop.

We admit that we take a heuristic approach to detecting misconfigurations that leads to worse reachability. As a matter of fact, $CS_i \supset CS_j$ is not a sufficient condition for $P(CS_i) > P(CS_j)$ (CS_i performs better than CS_j). We see that CS_i (a superset with more cells) performs even worse than CS_j (fewer cells) in unanticipated-yet-possible cases. This is because runtime performance of a cellset is impacted by many factors other than the cellset itself. Large performance variance is also observed in recent measurement studies [23, 32, 44]. Evidently, $CS_i \supset CS_j$ is not a necessary condition. In fact, we do observe that a cellset CS_i largely (statistically) outperforms another CS_j when $CS_i \supset CS_j$ does not hold. It matches with the example in [28]. If the PCell changes, we consider a simple rule to compare the cellsets: a cellset without 5G cells is worse than another with 5G cells. Ideally, we should compare the received performance of the given cellsets; But it is not practical as the performance is highly dynamic and measuring performance of multiple possible cellsets at the same time is impossible. Nevertheless, the good news is that current practice over two heuristics is empirically effective (e.g., Fig. 4).

Cross-parameter dependency analysis. For a single state, we examine how the dependency between parameters affects the output (i.e., cell selection result) and check whether the output is desirable or not. Note that "desirable" is a relative definition; In this work, we consider two desirable properties: loop-free and good performance. It turns out that cross-parameter dependency does not generate persistent loops but result in worse reachability in two forms: worse performance and failures. We take the same approach in the above reachability analysis to examine whether the selected cellset performs worse and manually check if it results in failures.

Specifically, we first need to understand how the output is affected by the interactions among parameters, device capability, and runtime radio quality. Such interactions have been standardized in 3GPP specifications [8–10, 12]. The specifications are written in the informal language, and each parameter can be referred in its aliases. We thus extract cross-parameter dependency from standards manually. We first map each parameter across specifications with its aliases. These parameters interact with each other by affecting the same variables. For example, measurement parameters affect the measurement results, which are checked with reporting parameters. We then track these parameters and their affected variables. We finally model a sequence of functions about parameter, UE capabilities, and radio quality measurements of the cells.

The challenge for the above analysis lies in the nondeterministic nature of the output, since it depends on dynamic radio channel quality. The naive idea is to examine how the output changes for each possible radio quality value given the current parameters. However, the space of possible dynamic radio quality for all cells may explode. We leverage the insight that cells can be grouped into classes based on their frequency types. For each cell class, the radio quality is monitored by the same parameters. For example (Fig. 5a), all 4G SCells over the serving frequencies (say, F850,

	Los Angeles (2019 – 2021)			Chicago (2021)			Indianapolis (2021 – 2022)			West Lafayette (2019 – 2023)			China (2019)	Total
Operator	US-I	US-II	US-III	US-I	US-II	US-III	US-I	US-II	US-III	US-I	US-II	US-III	CN-I	
♯ serving cells (4G)	2,450	1,642	2,869	2,001	1,970	767	964	1,143	235	2,562	2,184	1,866	4,352	25,005
♯ serving cells (5G)	21	N/A	31	254	381	198	435	154	87	N/A	N/A	52	0	1,613
♯ cellset instances	24K	18K	33K	41K	50K	10K	25K	39K	13K	50K	39K	23K	51K	416K
♯ cellset (w/o 5G)	8,913	5,248	9,696	1,838	2,221	710	1,426	1,493	212	10,629	8,166	3,567	7,489	61,608
♯ cellset (with 5G)	570	N/A	411	5,198	4,025	1,332	2,674	2,691	1,046	N/A	N/A	562	N/A	18,509

Table 4. Overview of the datasets in the US and China.

F66486, F66396) are examined with the same criterion A2 ($RSRQ_s < -19.5dB$). In the meanwhile, the 4G PCell (over F5110) is examined with two criteria: (1) A2 ($RSRQ_s < -17dB$) and (2) A3 ($RSRQ_c > RSRQ_s + 3dB$). Moreover, we find that DSM is helpful to group cells into different classes. For any state Ω_k in the CSM, we check all its possible delta states, namely, $D\Omega_{k,i} = \Omega_k - \Omega_{k+1,i}$, where $\Omega_{k+1,i}$ is the i-th configuration state directly derived from Ω_k. Interestingly, we find that a configuration state Ω_k is made of multiple delta states. Take $\Omega_{5G0-4G3}$ as an example; It consists of configurations for three classes: (1) 4G SCells (over F66486, F66936), namely, $\Omega_{5G0-4G3}$ - $\Omega_{5G0-4G1}$, (2) 5G cells (not 5G SCells), namely, $\Omega_{5G0-4G1}$ - $\Omega_{5G4-4G1}$ (actually, $\Omega_{5Gn-4G1}, n \geq 1$), (3) 4G PCell (over F5110), the remaining part of $\Omega_{5G0-4G1}$ except class (2), namely, $\Omega_{5G0-4G1} \cap \Omega_{5G4-4G1}$. Note the configuration decomposition remains the same when we use $\Omega_{5G0-4G2}$, not $\Omega_{5G0-4G1}$ to decouple configuration entries used by $\Omega_{5G0-4G3}$. Such decomposition is meaningful, which is associated with the role of the involved cells in the used cellset. For each cell class, we examine whether the output is undesirable by checking whether it is an *unconditional* checking of cell radio quality. An *unconditional* checking means 1) the radio quality is not checked so the result is always true, or 2) no matter what the radio quality is, the result is always false.

4.3 Overview of Our Reality Check

We have conducted a reality check on dependent misconfigurations with three 5G operators in the US (denoted as US-I, US-II and US-III) and one 4G operator in China. We use 4G to name all 4G family technologies which evolve from LTE (4G) to LTE-advanced (4.5G) and LTE-advanced Pro (4.75G). All the operators support carrier aggregation and run at least 4.5G. Table 4 lists the statistics of all the datasets. We have conducted several measurement studies in four metropolitan areas in the past years: Los Angeles (2019 – 2021), Chicago (2021), Indianapolis (2021-2022), West Lafayette (2019 – 2023). In Los Angeles, we used Google Pixel 4a, an early 5G phone model, to conduct 5G experiments when 5G was launched by US-I and US-III [28]. In Chicago and Indianapolis, we primarily used Google Pixel 5 (and rarely Pixel 4a) to measure 5G performance [32]. In West Lafayette, we used our past 4G/4.5G measurement studies before 2020 [19] and added new 5G traces after US-III launched 5G in late 2022. In China, we have extended our previous measurement study [20] and collected 4G traces from one operator (CN-I) in 2019[2]. All the phone models are rooted to use MobileInsight [36], an open-source tool to collect signaling messages on test smartphones.

This reality check has covered about 26.6K serving cells (4G: 25,003 and 5G: 1,613) and 416K cellset instances. There are 80,117 unique cellsets, including 18,509 cellsets with 5G. Interestingly, we do not see many 5G cells (4G cells are still dominant), but we do see that 5G is used more thanks to dual-connectivity (both 5G and 4G). In this study, all three US operators run 5G in a non-standard-alone (NSA) mode, which requires 4G as the master RAT and uses 5G as a secondary RAT. It implies that all the PCells are 4G cells. We have collected data speed results through file downloading experiments in the US. We thus use the US datasets only to quantify the impacts of misconfigurations on network performance. We then develop an offline misconfiguration checker to examine real-world traces, generate the DSMs and CSMs, and detect dependent misconfigurations. Table 1 summarizes our main findings in two categories and five sub-categories. All incur performance degradation.

[2]It was done before the covid-19 pandemic and no more experiments are possible after 2020 due to travel constraints.

5 STRUCTURAL DEPENDENCY

We present two sub-categories of misconfigurations (D1 and D2) in structural dependencies.

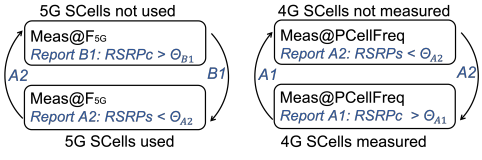
5.1 Necessary but Unattended Dependency (D1)

We check unattended dependency between state transitions using the learned DSM. For each possible transition walk between a pair of delta configuration substates, it might formulate a loop $D\Omega_1 \Rightarrow D\Omega_2 \Rightarrow \cdots D\Omega_x \Rightarrow D\Omega_1$. This is similar to the generic cyclic property in a directed graph. Note that each transition in DSM is conditional, so a loop only exists when all the conditions needed for the loop can hold simultaneously. We next present more instances where the loops indeed happen and cause disruption.

Issues. Fig. 6 shows two unattended dependencies identified in our study, each of which causes a persistent loop. The first is regarding the addition/removal of 5G SCells using events B1 and A2, as illustrated in§3.2. As long as $\Theta_{\text{B1}} < \Theta_{\text{A2}}$, the loop becomes possible. The second is related to the addition/removal of 4G SCells using events A1 and A2. We find that some 4G PCells configure an A2 event for its intra-freq measurement (over the same frequency used by the PCell) to decide whether to run an inter-freq measurement. When A2 is satisfied (say, $\text{RSRP}_s < \Theta_{\text{A2}}$, e.g., -104 dBm), the PCells add new measurement objects over other 4G frequencies and then add these 4G cells as SCells (when their RSRPs/RSRQs meet the configured criteria like event A3 or A5). Afterwards, the PCell updates its configurations by removing the A2 event and adding an A1 event (say, $\text{RSRP}_s > \Theta_{\text{A1}}$, e.g., -110 dBm). When A1 is satisfied, the PCell removes all the 4G SCells and gets back the original state (without any 4G SCell). Clearly, this loop might happen when $\Theta_{\text{A1}} < \Theta_{\text{A2}}$.

Root causes. We want to highlight that they are not dumb errors. Instead, they stem from necessary tradeoffs between flexibility and dependency. Using the same thresholds for both triggering conditions (or a smaller threshold for A2) can avoid loops since these two conditions can never be satisfied at the same time. Nevertheless, it would significantly limit the flexibility of adjusting conditions to adapt to complex wireless environments. Unfortunately, the flexibility of tuning individual thresholds increases the risk of unattended dependency. It is not easy to detect at the design and planning phase. Loop detection is a combinatorial problem by examining all the possible configuration states and their transitions. Each state is a set of configurations and it faces with an explosion problem well known in formal model checking [17]. It is also hard to detect by limited tests and field trials, because the problem appears only when certain conditions are satisfied (e.g., the device stays at places with 5G Cell's RSRP in [-115 dBm, -96 dBm] or at places with the PCell's RSRP in [-110dBm, -104dBm].) Driving tests or short static tests are unable to reveal anomaly.

Existence and prevalence in reality. We have examined all the configuration states per cell in our datasets and list misconfiguration instances in Table 5. In US-I, we find that 336 PCells (114 in Chicago and 192 in Indianapolis) have experienced these harmful B1-A2 loops. They take up about 5.7% (114/2001) and 19.9% (192/964) of all the PCells. A lower ratio was observed in Chicago because we run more experiments in Indianapolis and collected more configuration traces per PCell. The problematic PCells include cells over almost all 4G frequencies like F675, F850, F1150, F5110, F9820,

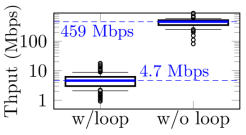


(a) B1-A2 ($\Theta_{\text{B1}} < \Theta_{\text{A2}}$) (b) A1-A2 ($\Theta_{\text{A1}} < \Theta_{\text{A2}}$)

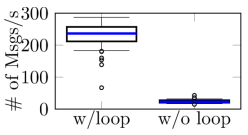
Fig. 6. Two persistent loops caused by necessary yet unattended structural dependencies (D1).

B1-A2	♯ PCell	F_{PCell}	Θ_{B1} (dBm)	Θ_{A2} (dBm)
US-I	336	F675, F850, F5110 $\cdots$ (♯:15)	-113, -110	-96
A1-A2	♯ PCell	F_{PCell}	Θ_{A1} (dBm)	Θ_{A2} (dBm)
US-I	68	F850, F5110, $\cdots$ (♯:12)	-19.5 (dB)	-17 (dB)
US-II	6	F1050, F1550, F5230 (♯:3)	-136, -130	[-126,-116]
US-III	76	F40978, F2300, $\cdots$ (♯:5)	-111 (-140)	-105 (-118)

Table 5. D1 instances observed in our study and common threshold values: RSRP (dBm), RSRQ (dB).



(a) Throughput drop

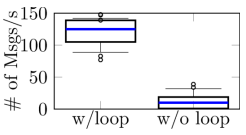


(b) Signaling overhead

Fig. 7. Negative impacts due to a B1-A2 loop.

	Loop (s)	Dis.(s)	D/L(%)
Min	0.6	0.07	3.1%
Max	213.3	14.7	12.5%
Med	3.4	0.2	6.8%
Avg	26.2	1.8	7.1%

(a) Disruption in driving tests



(b) Signaling overhead

Fig. 8. Negative impacts due to a A1-A2 loop.

F66686 (15 frequencies observed). The B1 threshold Θ_{B1} is mostly set as -113dBm, -112dBm or -110dBm, all smaller than the A2 threshold (-96dBm). We notice that this problematic configuration happens with 5G cells over both mmWave (say, F2259995, F2256663) and sub-6G channels (say, F174270 and F174300). In total, 11 frequencies for 5G are involved. We have observed A2-A1 loops with 144 PCells (US-I: 68, US-II: 6 and US-III: 70). US-I configures A1/A2 thresholds in RSRQ and other two operators in RSRP. It is observed over almost all possible frequencies in US-I and limited to certain frequencies in US-II and US-III.

We want to point out that there might be more cells with such dependent misconfigurations. We cannot crawl all the configuration states used by each PCell because they are also impacted by runtime radio measurement. We cannot test all the possible RSRP/RSRQ measurements and thus fail to observe all configuration states. Such misconfigurations are not common; however, loops are highly likely to happen (likelihood from 66.7% to 100% with 93.3% on average).

Damage assessment. We further assess their performance impacts. We use the B1-A2 loop instance presented in §3.2 to evaluate the negative performance impacts. Fig. 7 shows that data throughput drops by two order of magnitude (from 459Mbps to 4.7Mbps) with the drop rate of 99%. Loops also incur excessive signaling overhead, which wastes resources of both user devices and operators. Here, the number of signaling messages grows by one order of magnitude (from 23 to 239 messages/second). We choose a A1-A2 loop instance in US-III and run driving tests passing by this misconfigured cell. We calculate loop durations, loop-induced disruptions, and signaling overhead. Fig. 8 shows that the loops last 26.2 seconds on average (up to 213.3s) with 1.8s disruption (up to 14.7s) due to inter-freq measurement and increase the number of signaling messages by 13-fold (median). In one drive test, 15.7 seconds (7.4%) of disruption is observed in this 211.3-second loop. In more tests, the disruption ratio goes up to 12.5%. Note that the loops do not last long because we drive and eventually leave these cells. In the static tests, the loop persists as long as the radio quality of PCells satisfies the loop conditions.

5.2 Unnecessary Dependency (D2)

Dependency among states should guarantee that the device reaches good states. We reveal all possible SCell combinations observed in the study, and use two heuristic rules (presented in §4.2) to locate unnecessary dependency which likely goes wrong: (1) a cellset CS_1 is better than CS_2 if $CS_1 \supset CS_2$, and (2) a cellset without 5G cells is worse than another that works with 5G cells.

Issues. We use an instance observed in US-I to exemplifies how unnecessary dependencies among states result in undesired reachability. To better understand this instance, Fig. 9 plots the involved cellsets and the key configuration entry, not the inferred DSM. The device is initially served by CS_{850} that uses a PCell at F9820 and an SCell at F850. When this SCell is becoming weak (upon an A2 reporting), the PCell considers other combinations of SCells. Based on the inferred DSM, the available choices include one SCell at F1975 and two SCells at F1975+F5110, corresponding to cellsets CS_{1975} and $CS_{1975+5110}$. However, in practice, the configuration only considers F1975, thereby missing the opportunity of better network performance using two SCells. Note that we cannot judge whether CS_{850} or $CS_{1975+5110}$ is better.

We find two types of D2 instances which miss 4G SCells or 5G SCells (Fig. 10). The first (D2A) is to miss some 4G SCells when updating frequencies to measure. In our study, we see that all the US

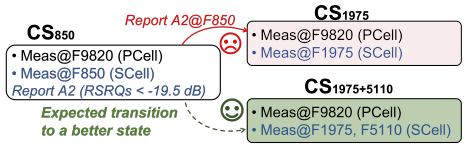


Fig. 9. An example of worse reachability caused by problematic (unnecessary) dependency (D2).

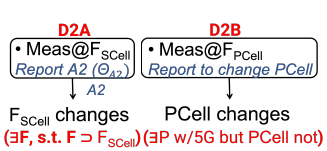


Fig. 10. Two types of worse reachability: D2A (4G SCells missed) and D2B (5G SCells missed).

	D2A	D2B
US-I	33	56
US-II	0	0
US-III	3	16
CN-I	1	N/A
♯ PCell	37	72

Table 6. D2 instances observed in our study.

operators use event A2. When A2 is reported, inter-freq frequencies F_{SCell} are updated. However, F_{SCell} may miss some frequencies available so that the cells over these frequencies will never be considered. The second (D2B) misses 5G SCells when the PCell changes. Several reporting events (say, A3, A5 and reportStrongestCell) are used. Note that reportStrongestCell is to report the cell with strongest RSRP/RSRQ over the same frequency of the original PCell. We see that some PCells never configure measurements over 5G frequencies so that they do not work with 5G. As a result, selecting such PCells will end with a 4G-only cellset, which is usually worse than 4G+5G choices.

Root causes. The problem is rooted in the inefficiency of 4G/5G SCell selection logic. The dependency between the current state and a bad state blocks the transition to a better state. The unnecessary dependency between states takes the blame. This is not a trivial fault since it is computationally hard to generate all possible configurations and examine all possible SCell combinations. It is hard, if not impossible, for the operators to consider all possible frequencies when configuring their SCell selection logic.

Existence and prevalence in reality. We have identified and validated 109 instances that miss 4G SCells and 5G SCells. In total, there are 33+56 PCells in US-I and 3+16 PCells in US-III with problematic configurations. We do not observe such instances with US-II. This is because US-II typically only has one or two frequencies available for SCells in our study. The limited choices reduce the risks of reachability issues. All three US operators disable 5G access at some 4G cells. Specifically, we see that all the cells over F5330 (US-I), F39874 (US-III) and F40072 (US-III) do not work with 5G. There are 218, 13, 7 cells observed in three US dataset except in West Lafayette. In West Lafayette, we observed that US-III used 95 cells over F39874 and 101 cells over F40072 before 2020 and these cells disappeared in the recent measurement in 2023 because these two 4G channels (over band 41) are repurposed to run 5G (over band n41).

Damage assessment. To assess performance impacts, we compared speed test throughput with and without reachability issues at the same location. The current practice selects the set of SCells, which is a subset of the available one. In the example (Fig. 9), the device is served by an SCell at F1975, missing another SCell at F5110. We run three comparison experiments and see that throughput drops by 28.3% on average (from 12.8% to 44.2%). We randomly pick five locations well covered by such misconfigured PCells to run comparison experiments. Throughput drops by 48.2 Mbps on average (up to 138 Mbps) with a drop rate from 18.3% to 87.6%. Figures are skipped due to space limit. These findings also match with recent measurement studies in US carriers [19, 28].

6 CROSS-PARAMETER DEPENDENCY

We check cross-parameter dependency within a single configuration state. Even for a single state where all device behaviors follow the specifications, unnecessary and unattended dependencies may result in unanticipated operations at the measurement, reporting and decision steps.

6.1 Unnecessary Dependency in Reporting (D3)

Issues. We discover that unnecessary dependency between parameters affects the reporting of candidate cells to change the PCell and SCells (D3). Such unnecessary dependency stems from shared

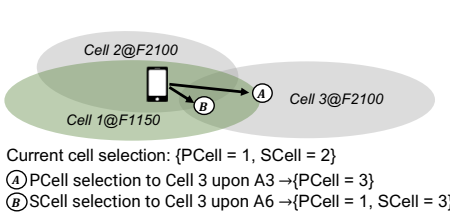
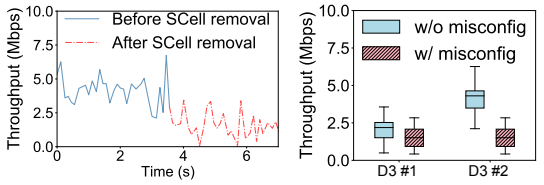


Fig. 11. Illustration of dependency in reporting.



(a) Showcase (b) Throughput drop

Fig. 12. Negative performance impacts in D3 instances.

reporting parameters for PCell selection and SCell selection. As noted in a technical report [48], the thresholds used to select PCell or SCells should be independently tuned to adapt to the environment. However, the same offset is used when a cell is considered as a PCell candidate (by A3) or an SCell candidate (by A6). It thus incurs unnecessary dependency between PCell/SCell selection.

Figure 11 exemplifies such dependent misconfigurations in one instance. There are three candidate cells near the device. At the start, the device is served by Cell 1 as PCell and Cell 2 as SCell. Cell 3 can be the target cell for PCell selection (PCell changes to Cell 3 via A3) or SCell selection (PCell is unchanged; SCell changes to Cell 3, via A6). However, the thresholds used by A3 and A6 are dependent because they use the same offsets, $\Delta_{A3} = \Delta_{A6}$ (Table 2). When the operator tunes Δ_{A3} to trigger PCell selection at given locations, the offset Δ_{A6} to trigger SCell selection also changes accordingly. We observe that such problematic offsets do exist in reality. The tuned offset results in a better PCell selection. However, it results in a negative threshold for A6, triggering a report for SCell removal or replacement by another one with worse radio signals. The triggered report falsely removes a good SCell and hurts performance.

Root causes. The offset parameters are shared for two procedures: PCell selection and SCell selection. But they were independent in the earlier releases of 3GPP standards where no SCells were considered in 3G/4G networks. Fixing these issues is also challenging. The operator has to trade-off between handoff robustness and aggregated performance.

Existence, prevalence and damage assessment. We find that only US-III has the identified reporting issues with four PCells. These PCells run over F5035 and F66786 while the involved SCells all run over F1150. Both A3 and A6 use the same offset 3dB. We further conduct validation experiments at misconfigured cells and observe problematic SCell removal in all the instances.

We use two instances to quantify negative impacts under similar experiment settings. Figure 12a shows an instance where the throughput drop happens right after removing the SCell. Removing SCell causes a throughput drop from 2.2 Mbps to 1.5 Mbps (31.0%) in the first instance and 4.3 Mbps to 1.5 Mbps (64.9%) in the second instance, as shown in Figure 12b. We observe that changing SCell does not change cell offset; however, changing PCell with the same SCell changes the offset values. It implies that the operators prioritize the optimization of PCell selection over SCell selection.

6.2 Necessary but Unattended Dependency (D4 and D5)

We also find necessary but unattended dependency. Parameters are mandatory or optional, as specified by 3GPP. It seems reasonable to give the operators flexibility to use optional configurations only when needed. However, we find that the absence of optional configurations leaves the dependency on them unattended.

D4: cross-parameter dependency in decision. At the decision step, RSRP thresholds are mandatory, but RSRQ thresholds are optional [10]. It is because RSRP was first required, and RSRQ was added later. For backward compatibility, RSRQ might not be supported by old-model devices or base stations. In practice, the PCell configures whether to use RSRP or RSRQ to make decisions, as illustrated in Figure 13. A RSRQ threshold, `NeighThreshHigh-RSRQ`, is optional, depending on the presence of another RSRQ parameter, `ServThresh-RSRQ`.

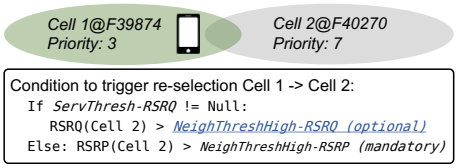


Fig. 13. Illustration of optional parameters.

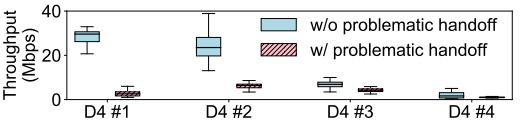


Fig. 14. Negative impacts of problematic handoffs (D4).

Issues. The operators using RSRQ may fail to deal with dependency between two thresholds. Configuring `ServThresh-RSRQ` without `NeighThreshHigh-RSRQ` is standard-compliant but results in handoff failure or throughput degradation. When the latter is absent, a negative infinity value would be applied. The switch from Cell C1 to C2 is triggered when $RSRQ_{C2} > -\infty$, which is always true. Since the target cell will be selected no matter how poor its radio quality is, the device might hand over to an inaccessible cell. Even if the target cell is accessible, the service might be much worse. The device performance will degrade after the handoff.

Root causes. The dependency between optional configurations is necessary to adapt configurations for different scenarios. However, it is unattended. The root cause is two-fold. First, the standard fails to address the dependency on optional configurations. The standard could enforce the dependency between parameters by adding necessary presence requirements (e.g., `NeighThreshHigh-RSRQ` is mandatory when `ServThresh-RSRQ` is present). Second, the operator fails to coordinate inter-dependent configurations across cells. In this case, the RSRQ thresholds for the target cell are decided by that cell while `ServThreshRSRQs` is decided by the serving cell.

Existence, prevalence, and damage assessment. We have found 11, 4, 22 and 7 instances (PCells) with US-I, US-II, US-III and CN-I, respectively. We randomly select 11 instances in the US and observe negative impacts in 4 of them. Problematic PCell selection does not happen in the other 7 instances because other cells of higher priority get selected. We do not see such instances in Chicago and Indianapolis probably because there are many fewer cells using RSRQ.

To assess impacts, we compare file downloading speed with and without problematic handoffs. We run back-to-back comparison experiments to minimize the impact of dynamic cell loads. In these four instances, handoffs happen with a probability of 50% - 100% and throughput drops by 38.7% - 91.2%. We observe that phones do not always trigger problematic handoff since they decide handoff targets based on the varying measurement results of all the cells.

D5: unattended dependency in measurement. In order to get valid measurement results over monitored frequencies, the parameters should be coordinated. Given a single state (set of configuration entries used together), the dependency exists not only between network configurations but also between network configurations and device capabilities.

Issues. We have identified one unattended dependency between configurations and device capabilities on an inter-freq measurement. Once an inter-freq measurement is configured, the device needs a measurement gap (`MeasGap`), if the device modem does not support no-gap monitoring [8, 10]. At hence, the need for a measurement gap depends on whether inter-freq measurement is configured and whether the device supports no-gap monitoring. In this case, if the dependency is unattended and the measurement gap is missing, the device cannot perform inter-freq measurement even if it is about to leave the coverage of the current serving cells and needs to explore more candidate cells over other frequencies. As a result, the device might miss inter-freq cells and thus get unconnected or miss potential serving cells. If failure, the device has to re-establish the connection, which disrupts its ongoing call and data services.

Root causes. The dependency between parameters and device capabilities is designed to support heterogeneous devices. However, it is not trivial to decide the presence of a measurement gap since the cell often updates inter-freq measurement. Without properly attending to the dependency, the measurement gap is problematic and will cause problems.

Existence, prevalence, damage assessment. We have identified and validated D5 with only one instance in China. The device is not configured with `MeasGap` when the inter-freq measurement is needed. Consider the limited phone models used in China, the problem might be underestimated. There are no such cases in the US. This implies more prudent engineering practice by US operators.

We take a trace-driven emulation[3] to assess the impacts of missing inter-freq measurement. We have emulated all the failure instances and found that 127 out of 458 observed handoff failures (out of 11,312 handoffs) are caused by missing inter-freq measurement. The handoff failure ratio increases by 37.9% (from 2.9% to 4.0%).

7 RELATED WORK

Misconfiguration in Internet systems. Misconfigurations have been extensively studied in various Internet systems, including BGP [21, 22, 27, 34], DNS [25, 40], SDN [22, 39], and data centers [14, 24, 43]. A few studies work on similar misconfiguration problems like instability, reachability and frequent misconfigurations, but the problem contexts are entirely different. Our problem context is 5G/4.5G RRC, which uses more complex, time-varying, location-dependent configurations at the device and the network.

Misconfiguration in cellular networks. Previous studies have reported a number of misconfigurations in 3G/4G networks which result in handover instability [30, 42, 46, 47] and undesired reachability with worse performance [20, 41, 45] (detailed in §3.1). Our work is inspired by them but investigates new misconfigurations in 5G/4.5G networks which use more than one cells to serve the device; Configurations are much more dynamic and complex per PCell. The findings also differ. We focus on inter-dependency among dynamic configuration and have identified new dependent misconfigurations which have not been reported before.

Configuration management and optimization in cellular networks. Another related but different topic is on optimizing cellular configurations to enhance performance [18, 35], improve robustness [15, 31, 33], and save energy [26]. Their approach is to formulate a mathematical optimization problem and apply various techniques to tune parameters. Ours is different. The inter-dependency among configurations cannot be captured with precise mathematical forms. We seek to conceptually categorize this new type of misconfigurations and develop a quick fix solution by detecting possible misconfigurations on the device. We further reveal the negative impacts in terms of persistent loops, reachability, or throughput degradation.

Recent measurement studies in 5G. Several 5G measurement studies have been conducted in the recent years [23, 28, 29, 32, 37, 38, 44]. Our findings somehow match with their measurement results: current configurations hurt 5G performance in some cases. But they do not work on RRC misconfiguration, which is our focus.

8 CONCLUSION

Proper configuration is critical yet challenging to the operation of 5G/4G cellular networks. On one hand, cellular configurations are highly dynamic, varying over time and locations. On the other hand, configurations must be coordinated at runtime and collectively done by the device and the infrastructure. The device sends runtime measurement reports, and the infrastructure makes decisions. The operators should have enough freedom to customize their cell-level configurations to better support mobility and boost network performance for mobile users. In this work, we study inter-dependent configurations in 5G/4.5G networks, and move beyond prior works on single-cell misconfiguration and configuration conflicts among multiple cells. We take a fresh view of new issues, devise new methodology, and uncover novel findings on dependent misconfigurations.

[3]No travel to China for empirical validation can be made due to COVID-19.

ACKNOWLEDGMENTS

We greatly appreciate our shepherd and all anonymous reviewers for their constructive feedback. The work has been partially supported by NSF grants CNS-1750953, CNS-1910150, CNS-2008026 and CNS-2112471.

APPENDIX A: ABBREVIATIONS

RAT	Radio Access Technology (say, 5G, 4G, 3G)
RRC	Radio Resource Control
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
MCG	Master Cell Group (using the master RAT)
SCG	Secondary Cell Group (using the secondary RAT)
PCell	Primary Cell (of MCG, using the Master RAT)
SCell	Secondary Cell (all serving cells except PCell)
NSA	Non-Standalone
SA	Standalone
SIB	System Information Block

Table 7. 5G/4G abbreviations used in the paper.

APPENDIX B: DISCUSSION ON OPEN ISSUES

We discuss open and remaining issues.

Implications to 5G SA. We focus on 5G NSA in this work because 5G SA is rarely observed in the used datasets. In our measurement study, we observed that 5G SA is used only around one official store of US-III in a big city. In several 5G SA instances collected in our study, each serving cellset consists of one 5G cell as PCell and zero/one 4G cell as SCell. The implications to 5G SA depends on how many RATs are used. For single-RAT 5G SA (namely, only 5G cells used), some misconfigurations (e.g., the `B1-A2` loop) likely disappear because there are no problematic dependencies between RATs. However, other misconfigurations (e.g., the `A1-A2` loop and missing certain cells due to improper reporting) may still exist with 5G SA, where the PCell changes from a 4G cell to a 5G one. For dual-RAT 5G SA (with dual connectivity and emerging multiple connectivity [16]), all dependent misconfigurations identified in this work seem applicable to 5G SA, where 5G and 4G swap their roles in the possible misconfigurations. However, we gauge that the performance impacts are less negative because the impacts of improper 4G SCells to a 5G PCell is smaller than the one of improper 5G cells to a 4G PCell. Finally, we would like to discuss a special 5G SA scenario where 5G uses two frequency ranges: sub-6GHz (< 6GHz) and mmWave (> 24 GHz). Although both are called as 5G New Radio, they operate like two RATs due to physical spectrum constraints. To this sense, this single-RAT but dual-range 5G SA (over sub-6GH and mmWave) is similar to a dual-RAT scenario, and is thus likely prone to the identified dependent misconfigurations. Interestingly, such dual-range 5G use is allowed by 3GPP but have not been observed in reality yet [32]. Essentially, whether 5G SA is prone to similar or even new misconfigurations, depends on the operators' policies as well as the vendors' implementation.

Recommended solutions. A quick fix we suggest is to convert the offline misconfiguration checker into runtime detection and migration at the device side. Once any misconfiguration detected, the device reacts to alleviate the damages. We suggest a device-centric solution because network upgrades are not open to the public, though our detection should conceptually work at the network side. This device-side solution has to tackle two challenges: (1) incomplete or even limited

configurations states as not all can be observed at runtime, and (2) dynamics induced by varying environment. We recommend fixes on the infrastructure side. Mobile operators have incentives to detect and fix misconfigurations. Moreover, they have complete knowledge about customized configuration logic. Therefore, they are capable of checking dependency among configurations proactively. Another long-term remedy is to improve 3GPP specifications. We recommend decoupling dependent configurations to avoid unnecessary dependency. We also propose to mandate the presence if the default value is error-prone.

Network-side Adoption. All the approaches to detecting and fixing misconfigurations in this work can be applied to the network side. Moreover, with a global view of nearby cells as well as their runtime work loads and performance, it would be better and easier to ensure proper RRC configurations at the network side. Detecting undesired reachability requires to predict runtime performance of every candidate cellset to determine whether the selected cellset is a worse choice. It is more technically feasible on the network side, rather than on the device side.

REFERENCES

[1] 2023. Artifact Release. https://github.com/mssn/5G_misconfig.
[2] 3GPP. 2022. Carrier Aggregation Explained. https://www.3gpp.org/technologies/101-carrier-aggregation-explained.
[3] 3GPP. 2022. Carrier Aggregation on Mobile Networks. https://www.3gpp.org/technologies/carrier-aggregation-on-mobile-networks.
[4] 3GPP. 2022. TS37.340: NR; Multi-connectivity; Overall description; Stage-2. V16.10.0.
[5] 3GPP. 2022. TS38.215: NR; Physical layer measurements. V16.5.0.
[6] 3GPP. 2023. TS23.501: System Architecture for the 5G System. V16.16.0.
[7] 3GPP. 2023. TS36.101: E-UTRA; User Equipment (UE) Radio Transmission and Reception. V16.16.0.
[8] 3GPP. 2023. TS36.133: E-UTRA; Requirements for support of radio resource management. V16.15.0.
[9] 3GPP. 2023. TS36.304: E-UTRA; User Equipment Procedures in Idle Mode. V16.8.0.
[10] 3GPP. 2023. TS36.331: E-UTRA; Radio Resource Control (RRC). V16.12.0.
[11] 3GPP. 2023. TS38.104: NR; Base Station (BS) radio transmission and reception. V16.15.0.
[12] 3GPP. 2023. TS38.331: NR; Radio Resource Control (RRC). V16.12.0.
[13] 5G America. Jan 2021. 3GPP Releases 16, 17 and Beyond. https://www.5gamericas.org/wp-content/uploads/2021/01/InDesign-3GPP-Rel-16-17-2021.pdf.
[14] Theophilus Benson, Aditya Akella, and David A Maltz. 2009. Unraveling the Complexity of Network Management.. In *NSDI*.
[15] Christopher Brunner, Andrea Garavaglia, Mukesh Mittal, Mohit Narang, and Jose Vargas Bautista. 2006. Inter-system Handover Parameter Optimization. In *VTC Fall*.
[16] S. A. Busari, R. Mumtaz, and J. Gonzalez. 2020. Multi-Connectivity in 5G New Radio Standards. https://www.standardsuniversity.org/e-magazine/december-2020/multi-connectivity-in-5g-new-radio-standards/.
[17] Edmund M Clarke, William Klieber, Miloš Nováček, and Paolo Zuliani. 2011. Model checking and the state explosion problem. In *LASER Summer School on Software Engineering*. Springer, 1–30.
[18] Estefanía Coronado, Shuaib Siddiqui, and Roberto Riggio. 2022. Roadrunner: O-RAN-based Cell Selection in Beyond 5G Networks. In *2022 IEEE/IFIP Network Operations and Management Symposium (NOMS)*. 1–7.
[19] Haotian Deng, Qianru Li, Jingqi Huang, and Chunyi Peng. 2020. iCellSpeed: Increasing Cellular Data Speed with Device-Assisted Cell Selection. In *ACM International Conference on Mobile Computing and Networking (MobiCom'20)*.
[20] Haotian Deng, Chunyi Peng, Ans Fida, Jiayi Meng, and Charlie Hu. 2018. Mobility Support in Cellular Networks: A Measurement Study on Its Configurations and Implications. In *ACM Internet Measurement Conference (IMC'18)*.
[21] Nick Feamster and Hari Balakrishnan. 2005. Detecting BGP configuration faults with static analysis. In *NSDI*. 43–56.
[22] Timothy G Griffin and Gordon Wilfong. 1999. An analysis of BGP convergence properties. In *ACM SIGCOMM Computer Communication Review*, Vol. 29. 277–288.
[23] Ahmad Hassan, Shuowei Jin, Arvind Narayanan, Ruiyang Zhu, Anlan Zhang, Wei Ye, Jason Carpenter, Z. Morley Mao, Zhi-Li Zhang, and Feng Qian. 2022. Vivisecting Mobility Management in 5G Cellular Networks. In *SIGCOMM'22*.
[24] Dilip A Joseph, Arsalan Tavakoli, and Ion Stoica. 2008. A policy-aware switching layer for data centers. In *ACM SIGCOMM Computer Communication Review*, Vol. 38. ACM, 51–62.
[25] Siva Kesava Reddy Kakarla, Ryan Beckett, Behnaz Arzani, Todd Millstein, and George Varghese. 2020. GRoot: Proactive Verification of DNS Configurations *(SIGCOMM'20)*. 310–328.

[26] Ali T Koc, Satish C Jha, Rath Vannithamby, and Murat Torlak. 2014. Device power saving and latency optimization in LTE-A networks through DRX configuration. *IEEE Transactions on wireless communications* 13, 5 (2014), 2614–2625.
[27] Craig Labovitz, Abha Ahuja, Abhijit Bose, and Farnam Jahanian. 2000. Delayed Internet routing convergence. *ACM SIGCOMM Computer Communication Review* 30, 4 (2000), 175–187.
[28] Qianru Li and Chunyi Peng. 2021. Reconfiguring Cell Selection in 4G/5G Networks. In *IEEE International Conference on Network Protocols (ICNP'21)*.
[29] Qianru Li, Zhehui Zhang, Yanbing Liu, Zhaowei Tan, Chunyi Peng, and Songwu Lu. 2023. CA++: Enhancing Carrier Aggregation Beyond 5G *(MobiCom'23)*.
[30] Yuanjie Li, Haotian Deng, Jiayao Li, Chunyi Peng, and Songwu Lu. 2016. Instability in Distributed Mobility Management: Revisiting Configuration Management in 3G/4G Mobile Networks. In *ACM International Conference on Measurement and Modeling of Computer Science (SIGMETRICS'16)*.
[31] Min Liu, Zhongcheng Li, Xiaobing Guo, and Eryk Dutkiewicz. 2008. Performance analysis and optimization of handoff algorithms in heterogeneous wireless networks. *IEEE Transactions on Mobile Computing* 7, 7 (2008), 846–857.
[32] Yanbing Liu and Chunyi Peng. 2023. A Close Look at 5G in the Wild: Unrealized Potentials and Implications. In *IEEE International Conference on Computer Communications (INFOCOM'23)*.
[33] Andreas Lobinger, Szymon Stefanski, Thomas Jansen, and Irina Balan. 2011. Coordinating handover parameter optimization and load balancing in LTE self-optimizing networks. In *2011 IEEE 73rd vehicular technology conference (VTC Spring)*. 1–5.
[34] Ratul Mahajan, David Wetherall, and Tom Anderson. 2002. Understanding BGP misconfiguration. *ACM SIGCOMM Computer Communication Review* 32, 4 (2002), 3–16.
[35] Ajay Mahimkar, Zihui Ge, Xuan Liu, Yusef Shaqalle, Yu Xiang, Jennifer Yates, Shomik Pathak, and Rick Reichel. 2022. Aurora: conformity-based configuration recommendation to improve LTE/5G service. In *Proceedings of the 22nd ACM Internet Measurement Conference*. 83–97.
[36] MobileInsight. 2022. http://www.mobileinsight.net.
[37] Arvind Narayanan, Eman Ramadan, Jason Carpenter, Qingxu Liu, Yu Liu, Feng Qian, and Zhi-Li Zhang. 2020. A First Look at Commercial 5G Performance on Smartphones *(WWW'20)*.
[38] Arvind Narayanan, Xumiao Zhang, Ruiyang Zhu, Ahmad Hassan, Shuowei Jin, Xiao Zhu, Xiaoxuan Zhang, Denis Rybkin, Zhengxuan Yang, Zhuoqing Morley Mao, Qian Qian, and Zhi-Li Zhang. 2021. A variegated look at 5G in the wild: performance, power, and QoE implications *(SIGCOMM'21)*. 610–625.
[39] Heng Pan, Zhenyu Li, Penghao Zhang, Kave Salamatian, and Gaogang Xie. 2020. Misconfiguration checking for SDN: data structure, theory and algorithms. In *IEEE 28th International Conference on Network Protocols (ICNP'20)*.
[40] Vasileios Pappas, Zhiguo Xu, Songwu Lu, Daniel Massey, Andreas Terzis, and Lixia Zhang. 2004. Impact of configuration errors on DNS robustness. In *ACM SIGCOMM Computer Communication Review*, Vol. 34. 319–330.
[41] Chunyi Peng and Yuanjie Li. 2016. Demystify Undesired Handoff in Cellular Networks. In *IEEE International Conference on Computer Communication and Networks* (Waikoloa, Hawaii, USA) *(ICCCN'16)*. 9 pages.
[42] Chunyi Peng, Yuanjie Li, Zhuoran Li, Jie Zhao, and Jiaqi Xu. 2016. Understanding and Diagnosing Real-World Femtocell Performance Problems. In *IEEE International Conference on Computer Communications (INFOCOM)*.
[43] Peng Sun, Ratul Mahajan, Jennifer Rexford, Lihua Yuan, Ming Zhang, and Ahsan Arefin. 2014. A Network-State Management Service. In *SIGCOMM'14*.
[44] Dongzhu Xu, Anfu Zhou, Xinyu Zhang, Guixian Wang, Xi Liu, Congkai An, Yiming Shi, Liang Liu, and Huadong Ma. 2020. Understanding Operational 5G: A First Measurement Study on Its Coverage, Performance and Energy Consumption *(SIGCOMM'20)*.
[45] Shichang Xu, Ashkan Nikravesh, and Z Morley Mao. 2019. Leveraging Context-Triggered Measurements to Characterize LTE Handover Performance. In *PAM*.
[46] Zengwen Yuan, Qianru Li, Yuanjie Li, Songwu Lu, Chunyi Peng, and George Varghese. 2018. Resolving Policy Conflicts in Multi-Carrier Cellular Access. In *ACM International Conference on Mobile Computing and Networking (MobiCom'18)*.
[47] Xiaohui Zhao, Hanyang Ma, Yuan Jin, and Jianguo Yao. 2018. Measuring instability of mobility management in cellular networks. *IEEE Network* 32, 5 (2018), 138–144.
[48] ZTE. 2013. UMTS Handover Description. http://www.slideshare.net/quyetnguyenhong/zte-umtshandoverdescription. Accessed on Oct 16, 2021.

Received November 2022; revised April 2023; accepted April 2023