Demystifying Secondary Radio Access Failures in 5G

Yanbing Liu
Department of Computer Science,
Purdue University
liu3098@purdue.edu

Junpeng Guo
Department of Computer Science,
Purdue University
guo567@purdue.edu

Chunyi Peng
Department of Computer Science,
Purdue University
chunyi@purdue.edu

ABSTRACT

In this work, we have conducted a measurement study with three US operators to reveal three types of problematic failure handling on secondary radio access which have not been reported before. Compared to primary radio access failures, secondary radio access failures do not hurt radio access availability but significantly impact data performance, particularly when 5G is used as secondary radio access to boost throughput. Improper failure handling results in significant throughput loss, which is unnecessary in most instances. *Datasets are available at https://github.com/mssn/scgfailure.*

CCS CONCEPTS

• **Networks** → **Mobile networks**; *Network management*; *Network measurement*; *Network reliability*.

KEYWORDS

Cellular Network, 5G, Secondary Cell Group (SCG), SCGFailure

1 INTRODUCTION

Handling radio access failures is essential to cellular network reliability, availability and performance. When the radio link (RL) between a mobile device and its serving cell (also known as a base station) fails to transmit packets in the air, the ongoing data/voice sessions are interrupted until this radio link failure (RLF) is recovered (e.g., by another RL that works).

Handling radio access failures is more complex and harder, as cellular networks advance from 3G/4G to 4.5G/5G and beyond. In a 3G/4G network, a radio access failure is a *YES-or-NO* problem; Radio access is available (or unavailable) when the used RL does not fail (or fails). This is because 3G/4G uses a single RL to serve a mobile device. The problem turns more complicated as 4.5G/5G increases the number of active RLs from 1 to N ($N \geq 1$, mostly $N \gg 1$), through two advanced radio access technologies: carrier aggregation [2, 3]
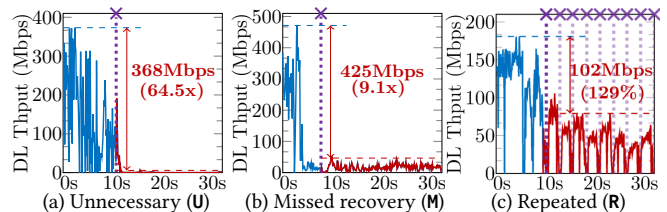
Figure 1: Real-world instances of three types of "problematic" SCG-Failure handling (U, M, R).

and dual connectivity [4, 5][1]. The former uses a group of serving cells, which was first adopted by 4.5G LTE-advanced [2]; The latter uses two cell groups, which was launched by 5G [5]. Specifically, each serving cell uses one RL over one frequency channel as a basic unit to offer radio access. All the serving cells are grouped into *Master Cell Group* (MCG) and *Secondary Cell Group* (SCG), based on their radio access technologies (RATs, here, 4G[2] and 5G) [5]. Each group uses carrier aggregation to combine one primary cell (PCell) and several secondary cells of the same RAT [3]. As a result, 5G aggregates radio frequency channels used by all active RLs over 5G and 4G/4.5G, thereby utilizing much wider radio frequency spectrum to boost network performance. Unsurprisingly, 5G is often much faster than 4G/4.5G, up to several hundreds of Mbps [10].

Radio access failures are handled at two levels: *logic* and *physical*. The above logic level is managed by radio resource control (RRC), which is responsible for establishing and maintaining a logic channel (namely, an active RRC connection) to transfer user traffic. Its connection state is still *YES-or-NO*, say, active/connected or idle. This logic channel is provisioned through *physical* RLs. 5G uses only MCG[3] to manage the logic RRC connection, and both MCG and SCG for physical radio access to mobile devices: MCG for primary radio access and SCG for secondary and opportunistic radio access.

In this work, we examine how 5G handles secondary radio access failures. Such failure is officially termed as *SCGFailure*, which was introduced in 3GPP Release-15, the first set of 5G standards [7]. An SCGFailure occurs when one or more RLs used by SCG fail but the RLs by MCG not. Therefore, the device, upon detecting an SCGFailure, is still able to report the detected failure to MCG and invoke a RRC procedure to recover the SCGFailure. In principle, SCGFailures do not harm access availability but impact data performance.

In this work, we are particularly interested in characterizing and understanding SCGFailure handling in operational cellular networks and demystifying *"problematic"* failure handling. We indeed

---

[1]Dual-connectivity has been extended to multi-connectivity in the recent 3GPP standard specification [4]. In this work, we focus on dual-connectivity because multi-connectivity has not been observed in operational cellular networks and all the findings are conceptually applicable to multi-connectivity.

[2]In the rest of the paper, 4G is used to represent all 4G variants, including LTE (4G), LTE-Advanced (4.5G) and LTE-Advanced Pro (4.75G). All US operators support 4.5G/4.75G.

[3]Actually, the PCell of MCG manages the RRC connection.

	Description	Performance Impact	Root Cause	OPs	Main Results
U	Unnecessary handling	Performance unnecessarily drops	Retransmission triggering event improperly configured	A, T, V	Figure 1a, 3, 7, 9
M	Missed recovery	Long-time poor performance	No piggybacked measurement report to recover the failure	A, T, V	Figure 1b, 11
R	Repeated failures	Long-time performance fluctuation	Random access failures due to too low SCG Addition threshold	A, T, V	Figure 1c, 8, 10

Table 1: Summary of our main findings on three types of "problematic" SCGFailure handling in our reality check.

observe *"problematic"* failure handling with three major US operators (AT&T, Verizon and T-Mobile, short as **A**, **V** and **T** afterwards). Figure 1 gives three real-world instances per type observed in our study (each SCGFailure marked as ×), which all result in substantial performance loss. In this work, we have unveiled three types of *"problematic"* failure handling:

• **Unnecessary failure handling**. A SCGFailure is falsely detected and reported, resulting in the unnecessary removal of SCG RLs which can offer good performance (Figure 1a). Here, the peak downlink throughput shrinks from 373.2 Mbps to 5.7 Mbps, losing 368 Mbps (368/5.7 = 64.5×).

• **Missed failure recovery**. A true SCGFailure is detected but not recovered in presence of suitable RLs, which results in significant performance degradation (Figure 1b). Here, the peak rate decreases by one order of magnitude (471.3 Mbps → 46.5 Mbps).

• **Repeated failures**. A true SCGFailure is recovered but the recovery does not last long (Figure 1c); SCGFailures are frequently repeated every a few seconds because the failed RL is used again for the recovery. Data throughput not only oscillates but also greatly declines by 129% (about 102 Mbps).

More details are elaborated in §3.1. Most importantly, we notice that such failure handling is "*problematic*", not because current practice does not follow the standard procedures. Instead, current practice conforms to 3GPP standards but still suffers *significant-but-unnecessary* performance degradation. We analyze their root causes and quantitatively characterize their prevalence and performance impacts (§3). Table 1 summarizes our main findings.

2 SCGFAILURE PRIMER

In this section, we introduce necessary background on SCGFailure handling, which is regulated in 3GPP standards [4, 5].

Figure 2 depicts a typical flow of handling a SCGFailure when primary radio access through MCG works at all time. It starts with detecting a radio link failure (RLF) (❶). 3GPP defines six triggering events to detect a RLF [5]. We find that problematic handling is mainly associated with two events: RAF and RTMAX. An RAF event occurs when the device fails its random access to the SCG cell; More precisely, it fails after multiple random access attempts within a time period, say, timer T304. An RTMAX event means that the maximum number of retransmissions (rlc-MaxNumRetx) is reached with a small time interval (t-PollRetransmit). Evidently, SCGFailure detection is impacted by these tunable parameters.

Upon detecting a triggering event, the device reports the detected RLF (via a signaling message called SCGFailureInformation) to the network with the detected event as its failure type (❷). The network immediately releases the "failed" RL and invokes a standard procedure to recover this RLF (❸). Specifically, recovery is realized by RRC Reconfiguration Procedure [5], which is used to find and add RLs available and suitable (say, meeting the RSRP/RSRQ requirement). It performs four steps: *configuration, measurement,*
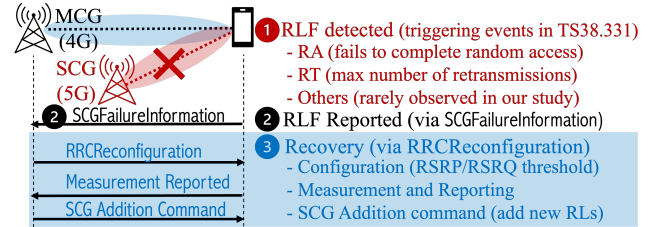

Figure 2: A typical flow of SCGFailure handling.

	OPs	Period	Region	Hours	#Conn	#SCGF
D1	A, V, T	04/21–01/22	19.8 Km²	705 hr	129.8K	4,275
D2	A, V, T	09/23–10/23	5.1 Km²	44 hr	7.5K	284

Table 2: Two datasets used in this study (SCGF= SCGFailure).

reporting and *command* (here, SCG Addition). The criteria for measurement and reporting are configured through several tunable parameters, particularly those RSRP/RSRQ thresholds used to compare radio quality of serving RLs and candidate RLs. The device needs to measure RSRP/RSRQ of available RLs (namely, candidate cells) and report new cells if found. Finally, the network sends a command to add the new SCG cell(s), namely, adding the new RL(s) to replace the "failed" RL(s) which were released before.

3 REALITY CHECK IN THE US

We now present our measurement study of SCGFailure handling with three US operators: **A**, **V** and **T**.

Methodology and Datasets. We first characterize and analyze SCGFailure instances using a public dataset **D1** from our recent 5G measurement study [9]. **D1** was collected over 5G experiments (primarily driving tests) in two US cities (Chicago and Indianapolis, total area: 19.8 Km²) from April 2021 to January 2022 (705 hours in total). With new findings and insights gained, we conduct a 2-month measurement study in West Lafayette, IN, focusing on problematic failure handling. In **D2**, we run speedtest through bulky file downloading as we did in [9]. Table 2 summarizes basic statistics of both datasets. In **D1** and **D2**, there are 129.8K and 7.5K RRC connections each with one or more RLs changed; We observe 4,275 and 284 SCGFailure instances, which account for 3.3% and 3.8% of all instances. We notice that the SCGFailure rate is not very high; It matches with our expectation; Operational cellular networks are largely successful and radio access failures should not be common. The astonishing finding is that more than half SCGFailure handling instances are problematic, which will be elaborated next.

3.1 Illustrative Examples

We start with three real-world instances (Figure 1) to unveil how SCGFailures are exactly (and improperly) handled in reality.

Unnecessary SCGFailure Handling. Figure 1a shows a stationary instance observed in West Lafayette (**D2**) with **T**, which runs 5G over sub-6GHz (< 6GHz). In our study, we see that all three
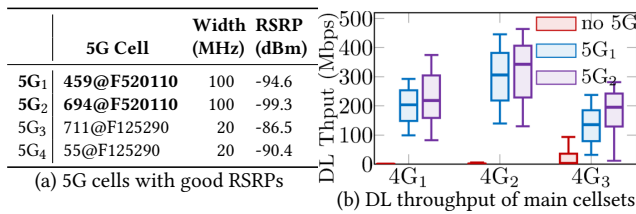
	5G Cell	Width (MHz)	RSRP (dBm)
5G$_1$	459@F520110	100	-94.6
5G$_2$	694@F520110	100	-99.3
5G$_3$	711@F125290	20	-86.5
5G$_4$	55@F125290	20	-90.4

(a) 5G cells with good RSRPs

(b) DL throughput of main cellsets

Figure 3: 5G SCG cells and main cellsets observed at the same location in the example instance of unnecessary failure handling (Figure 1a).

US operators use 4G MCG + 5G SCG, 4G for MCG and 5G for SCG. It is not hard to understand that problematic SCGFailure handling significantly hurts performance as 5G RLs are not properly utilized.

Initially, the device is served by 4G MCG + 5G SCG, achieving high throughput (median: 204 Mbps). The 5G SCG RL is used by a cell 5G$_1$ (459@F520110). Here, 459 is its cell ID and F520110 is its channel number, as specified in [3]; F520110 is a 5G channel centered on 2600 MHz with its channel width of 100 MHz. At 10s, a SCGFailure is detected with an RTMAX event, where the number of continuous retransmissions reaches its maximum (rlc-MaxNumRetx: 32) within a short interval (t-PollRetransmit: 45 ms). The device reports this detected RLF and the 5G RL is released immediately (though it is still able to offer high data speed). In order to expedite failure recovery, 3GPP recommends piggybacking the RSRP/RSRQ measurements of neighboring SCG cells while reporting the detected RLF [5]. In this instance, no measurement results of available 5G cells are piggybacked. Later at 17s, the device receives a new message RRCReconfiguration which configures the device to measure nearby cells over other 5G frequencies (here, F520110 and F125290). Surprisingly, no measurement results of 5G cells are reported despite the presence of four good 5G cells (Figure 3a). As a result, the device loses 5G as its secondary radio access and uses 4G only; Throughput shrinks below 10 Mbps.

This instance is *"problematic"* because 5G cells with good radio quality and high data throughput (hundreds of Mbps) are present but not used. We run extensive experiments at the same location and observe such 5G cells. Figure 3a lists four 5G cells with good radio quality (medium RSRP > -100 dBm). As long as 5G$_1$ or 5G$_2$ is used, data throughput is much higher (than no 5G), as shown in Figure 3b. 4G$_1$, 4G$_2$ and 4G$_3$ are three 4G cells used by the MCG.

We further examine why problematic failure handling occurs in this instance. It is attributed to two issues: *(1) improper* RLF *detection*, and *(2) no failure recovery*.

First, RLF is falsely detected by an RTMAX event which uses a threshold (rlc-MaxNumRetx: 32) and cannot effectively distinguish a SCGFailure under light/normal traffic and a normal use under heavy traffic. Ironically, when 5G SCG provides very high throughput (hundreds of Mbps), it highly likely experiences more than 32 continuous retransmissions even though many more packets are successfully transmitted over the used RL. Evidently, the higher throughput provided by 5G SCG, the higher the likelihood of a false trigger (an RTMAX event indicating the failure of the used RL [5])(or the higher the likelihood of losing high throughput provided by 5G SCG). Later, we will show that it is the dominant source to false RLF detection , which is commonly observed with all instances with unnecessary failure handling with all three US operators in §3.2.

Second, it is indeed hard to use this single instance to figure out why the failure is not recovered. We thus examine many instances with and without failure recovery to learn what makes a difference. We find that no recovery in presence of suitable 5G RLs is highly correlated with another operation: *no piggybacked measurement reports while reporting the detected* RLF *(via* SCGFailureInformation*)*. Interestingly, we find that the subsequent RRC Reconfiguration Procedure becomes ineffective if no measurement report is piggybacked. As long as at least one measurement report of any neighboring SCG cells is piggybacked, the recovery procedure can proceed: immediately add the qualified SCG cells in the piggybacked report or use the subsequent RRC Reconfiguration Procedure to later add qualified SCG cells which are not included in the piggybacked report (more details in Figure 4).

We also notice that unnecessary failure handling is not rare at the same location. It is true that SCGFailures are not common; The failure rate is 3.3% (4275 out of 129.8K) observed in **D1**, which ran a field test in two US cities [9]. Otherwise, cellular networks would not be largely successful. However, unnecessary failure handling is not rare where it occurs. At this location, more than 85% of SCGFailures are unnecessary, resulting in huge throughput loss (> 280 Mbps), in spite of various SCG and MCG cells involved.

Missed SCGFailure Recovery. Figure 1b is observed also with **T** in West Lafayette (**D2**), but on a walk route. The main difference from the above instance is that the detected SCGFailure is true. At 3.8s, data throughput drops below 30 Mbps from ~300 Mbps. The involved 5G SCG cell is 523@F520110 (a different cell but over the same channel). The RL indeed fails to complete random access to the 5G SCG cell, indicating that the uplink to the network does not work. As a result, the SCGFailure is detected with an RAF event, not with an RTMAX event. The problem lies in no recovery in presence of good RLs. We see two 5G cells (700@F125290, 700@F126270) with good radio quality (RSRP > -93 dBm), each of which can yield 100 − 150 Mbps once used. The plot is skipped due to space limit. However, recovery is missed for the same reason: *no measurement reports are piggybacked*, which blocks the reporting of qualified cells. It holds true for most instances with missed recovery.

Repeated SCGFailures. Figure 1c shows a stationary instance with **A**, another US operator observed in **D1**. At the start, the device does get high data throughput of 100 − 180 Mbps. At 9.2s, the device attempts to add a new SCG cell (634@F174270). F174270 is a 5G channel centered on 871 MHz (sub-6GHz). This 5G cell is measured with RSRP = -105 dBm, which is higher than the threshold (-110 dBm) needed for SCG addition. Then the device adds this cell but at 9.6s (400 ms later), it detects the RLF with an RAF event (similar to the second instance in §3.1). As a result, this SCG cell is released. It should not be a problem when the RL with a high RSRP value might fail (here, random access failure). The real problem is that the above process is frequently repeated. At 11.8s, the device attempts to reconnect to the same SCG cell but at 12.2s, the same SCGFailure happens again due to another RAF event. It is repeated for nine times within 20 seconds (9.2s, 30s). It keeps oscillating with two operations: SCG Addition and SCG Removal (due to SCGFailures). As a consequence, the overall throughput drops from 100 −180 Mbps to 0 − 80 Mbps. Clearly, repeated failures can be avoided if the network avoids the same mistake again and again.
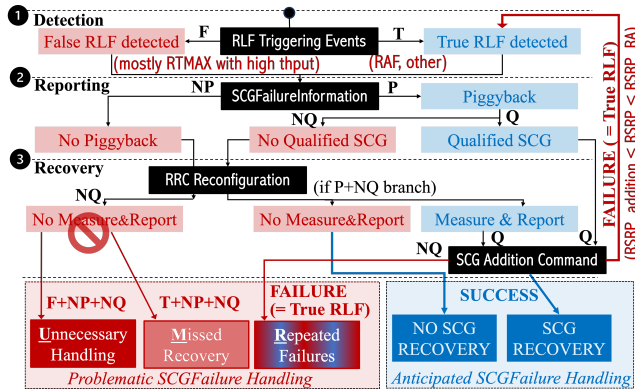
Figure 4: FSM for problematic SCGFailure handling.

(a) Dataset **D1**   (b) Dataset **D2**

Figure 5: Breakdown per SCGFailure type in **D1** and **D2**.

(a) **U**   (b) **M**   (c) **R**

Figure 6: Breakdown per triggering event.

Figure 7: Comparison of throughput before SCGFailures are triggered by **RTMAX** and other events.

3.2 Breakdown and Cause Analysis

We analyze root causes of problematic failure handling using all SCGFailure instances observed in both **D1** and **D2**. Figure 4 shows the results for three types of problematic failure handling, as well as two types of anticipated failure handling. We identify the problems at all three phases: ❶ detection , ❷ reporting and ❸ recovery.

Ideally, SCGFailures should be handled as follows. When an SCGFailure truly happens, this RLF should be quickly and correctly detected (true RLF detected), and immediately reported to the network for a prompt recovery (by piggybacking the measurement reports of candidate SCG cells). It should be recovered by proper RLs in presence of qualified SCG cells with acceptable radio quality and performance; Otherwise, if all candidate SCG cells are not acceptable, it should end with no SCG recovery.

Figure 4 plots a finite state machine (FSM) based on the outcomes at each phase of all the SCGFailure instances. We use three key signaling messages (SCGFailureInformation, RRCReconfiguration and SCG Addition) which are used at the reporting and recovery phases. We extract the failure type (say, the RLF triggering event) reported in SCGFailureInformation to analyze detection outcomes. At the detection phase (❶), there are two possible states: false RLF detected (**F**) and true RLF detected (**T**). At the reporting phase (❷), the measurement reports of candidate cells might be piggybacked (**P**) or not piggybacked (**NP**). In the **P** branch, the reported cells might be qualified (**Q**) or not qualified (**NQ**). If there exists at least one qualified cell, the recovery procedure (❸) skips the subsequent configuration and measurement steps, and directly uses the SCG Addition command to add new SCG cells; This ends with SCG recovery if success. In all other cases (in the **NP** or **P+NQ** branch), the recovery phase (❸) starts with RRCReconfiguration to run a complete 4-step procedure to add qualified SCG cells. Through analyzing the signaling messages in all the SCGFailure instances, we find that no qualified cells will be reported as long as there are no piggybacked reports at the reporting phase (in the **NP** branch). In the **P+NQ** branch, RRC Reconfiguration Procedure is performed as anticipated: cells will be measured and reported as configured in RRCReconfiguration. Specifically, the device reports the measurement reports of candidate cells if their RSRP/RSRQ is stronger than the given RSRP/RSRQ threshold (B1 event [5]). It ends with two anticipated failure handling: no SCG recovery (without qualified SCG cells) and SCG recovery (with qualified cells).
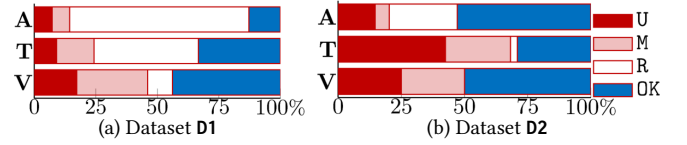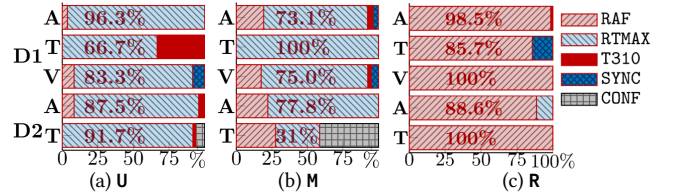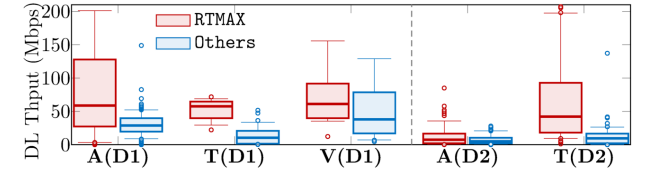
Finally, we find that problematic failure handling comes in three forms: *unnecessary handling* (**U**), *missed recovery* (**M**) and *repeated failures* (**R**). **U** and **M** share the same problem of no recovery in presence of qualified cells. That is, we share the error path **NP+NQ** in the left of Figure 4. Their difference is that **U** starts with a false RLF while **M** starts with a true RLF. **R** occurs when the newly added SCG cell suffers with the same failure which results in failure recovery.

Breakdown.   Before we dive into the root causes of problematic failure handling, we present the breakdown of all SCGFailure instances per type in Figure 5. We have three observations.

First, problematic SCGFailure handling are quite common out of all failure instances. We notice that SCGFailures are not common ($3.3\% \approx 4{,}275/129.8\text{K}$ in **D1**); However, once a SCGFailure occurs, it is likely handled in a problematic manner; For all three US operators, their ratios of normal failure handling all are below 50% in both **D1** and **D2**. *Second, the breakdown does vary with operators and test regions.* Operator **V** seems to do a better job while **A** and **T** suffer with more problematic SCGFailure handling in our study. The breakdown differs in **D1** and **D2** because 5G experiments are conducted in different cities. We admit that **D2** might be more biased as we intend to run more experiments at several locations of our interests. *Third, problematic failure handling also varies at locations.* In **D1**, 72.8% and 42.4% of SCGFailure instances are repeated failures with **A** and **T**. However, we notice that most instances with repeated failures take place in one or two small regions rather than evenly scatter at many locations. In contrast, unnecessary handling (**U**) and missed recovery (**M**) are observed at more places.

Root causes.   We next reveal how problematic failure handling occurs, namely, these key state transitions shown in Figure 4. Figure 6 shows the breakdown of all three types of problematic failure handling per trigger event. We skip **V** in **D2** because we do not see sufficient SCGFailure instances.

First, we have one interesting finding: RTMAX *is the only dominant trigger to unnecessary failure handling while* RAF *is almost the sole*
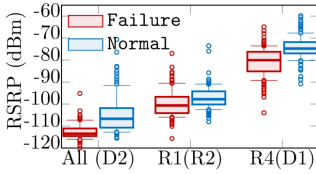
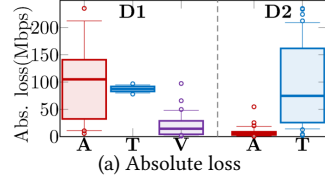**Figure 8: RSRP of SCG cells in repeated failures and normal cases.**

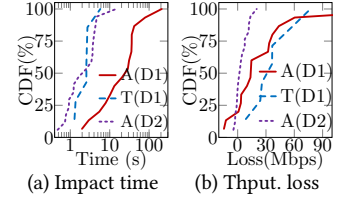**Figure 9: Absolute and relative throughput loss of unnecessary SCG-Failure handling (U).**

Figure 10: Negative impacts of repeated SCGFailures (R).

one to repeated failures. This matches with our illustrative instances (§3.1). RTMAX contributes to 83% – 96% of U instances for all three operators. Each operator sets rlc-MaxNumRetx as a constant so that an RTMAX event only considers the absolute count of continuous retransmissions but not the relative ratio of retransmissions. Unfortunately, more false alarms occur with heavier data traffic and higher throughput; A false SCGFailure is detected and reported even when the SCG is still working well and offering high throughput; As a matter of fact, we do observe that more RTMAX events are triggered with higher throughput. Figure 7 compares data throughput before an SCGFailure in two cases: triggered by RTMAX or other five events (defined in [5]). The median throughput is 42 – 62 Mbps if the SCGFailure is later triggered by RTMAX; It is much higher than those triggered by other events (16.6 – 38.0 Mbps). It implies that the "SCGFailures" triggered by RTMAX events might not be true failures. We do see that a false RLF can be linked with an RTMAX event at high throughput (Figure 4). It holds true in all operators except A in D2. This exception is due to various 5G deployment: A deploys 5G over mmWave and Sub-6GHz in D1, but operates 5G only over Sub-6GHz in D2. As a result, data throughput using 4G only or over 4G+5G is close in D2.

Second, RAF is the dominating event for repeated failures. 85.7%-100% of repeated failures are triggered by RAF (Figure 6c). RAF is designated to release the poor SCG when the device fails to complete random access to this SCG cell. It often happens when the RSRP of this SCG cell is below a certain threshold. However, the operator might set the RSRP threshold for SCG Addition below the threshold needed for random acces. More precisely, when the actual RSRP is larger than the threshold for SCG Addition (but smaller than the threshold for random access ($RSRP_{SCGAddition} <$ $RSRP < RSRP_{RA}$), the device re-connects to the failed cell and the SCGFailure is persistently repeated.

We further examine repeated failures in three representative regions: R1 and R4 in D1 [9] and West Lafayette in D2. We choose them because most instances with A were observed in these three regions. Figure 8 compares RSRPs of 5G cells with repeated failures and those with normal SCG Addition. We find that A sets a lower threshold (-114 dBm) for SCG Addition but a higher threshold for random access (RA): -110 dBm in D2, -100 dBm in R1 (D1) and -76dBm in R4 (D1). This helps us to understand why there are so many repeated failures observed in R4 (D1). The RSRP threshold in R4 is too high; Random access is hard, if not impossible, to succeed. For the same reason, the ratio of repeated failures is significantly higher in D1 (72.8%) than D2 (27.0%).

Third, apart from RAF and RTMAX, we observe three other triggering events of SCGFailures in our study: (1) SYNC (Cell synchronization fails), (2) T310 (Timer T310 expires with out-of-sync indications), (3) CONF (RRC reconfiguration fails). All events are regulated

by 3GPP [5]. Figure 6 shows that for all three kinds of problematic SCGFailure, the dominant triggering event is either RAF or RTMAX, while other events are rarely observed. For missed recovery failures, the dominant triggering event (RTMAX) is same to unnecessary failures. Figure 4 shows that the only difference between unnecessary failure handling (U) and missed recovery (M) is that a M instance is triggered correctly with a true RLF, while U with a false one. When a SCGFailure is correctly triggered but without piggybacked measurement reports of neighboring SCG cells, no qualified SCG cells will be reported; Consequently, MCG cannot send the SCG Addition command to the UE without candidate SCG cell information, and the UE thus misses the chance of recovery to good SCG cell(s).

3.3 Performance Impacts

We next present negative performance impacts of problematic SCG-Failure handling per type. Improper failure handling result in substantial throughput loss; We observe that download speed drops by half or more in most instances; It even declines by one order of magnitude (up to two orders of magnitude) in a few instances.

Unnecessary SCGFailure handling (U). We define two metrics to assess the resulting throughput loss – (1) *absolute loss*: the gap of average throughputs in 10 seconds before and after a SCGFailure; (2) *relative loss*: the ratio between absolute loss and throughputs after a SCGFailure. Figure 9 plots the distributions of the absolute and relative throughput loss with three US operators; V in D2 is skipped without sufficient instances. We have two observations.

First, throughput loss greatly varies with operators. In terms of relative loss, **T** suffers more throughput degradation than **A** and **V**. For **T**, download speed drops by more than one order of magnitude in almost all instances in D1 and 41.7% of instances in D2; The worst instance was observed in D2, with a 111.5-fold decline from 142.2 Mbps to 1.3 Mbps (median). For **A**, download speed declines by more than half (namely, the relative loss > 100%) in 63% of instances in D1 and 50% of instances in D2. Compared to **T** and **A**, **V** does the best job with its median loss below 30%.

Second, throughput impacts are largely consistent in terms of absolute and relative loss and inconsistent patterns are caused by various data speed before failures occur. Interestingly, we see that **A** has distinct patterns in terms of the absolute and relative loss in both datasets. Although its relative loss is similar in D1 and D2, the absolute loss is much lower in D2. For **A**, the absolute loss is >100Mbps in 51.9% of instances in D1, but <10 Mbps in 79.2% of instances in D2. Specifically in D1, **A** has the median loss of 105 Mbps (25th/75th percentile: 32.4 Mbps/141 Mbps), which is even higher than 87.3 Mbps with **T**. We further examine why. It turns out that such distinct impacts are caused by various 5G deployment. **A** deploys mmWave
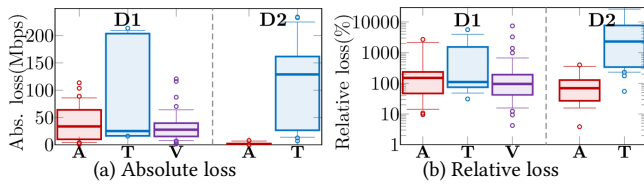
Figure 11: Absolute and relative throughput loss of missed SCGFailure recovery (M).

cells with much larger bandwidth (100 MHz) in **D1** but uses narrow channels over Sub-6GHz (10 MHz) in **D2**. The use of mmWave cells allows much higher throughput than 5G over Sub-6GHz. With much higher data throughout prior to SCGFailures, **A** thus loses much more absolute speed in **D1**; In **D2**, although the absolute loss is much smaller, negative impacts are not negligible; Data speed still reduces by half in more than 50% of instances. In contrast, **T** deploys 5G cells on the same sub-6GHz band (n41) in both datasets and the resulting impacts are consistent in these two datasets. Compared to **A**, **T** achieves higher data speed over sub-6GHz because it uses wider channels (bandwidth: 60/100 MHz).

Missed recovery SCGFailure handling (M).   We observe huge performance impacts when SCGFailure recovery is missed. To assess performance impacts of each **M** instance, we compare data throughput in two scenarios: (1) the reality without SCG cells being recovered, and (2) a what-if case with active SCG cells on the same location. We calculate the absolute and relative throughput loss between (1) the average throughput in a short time period (10 seconds) just after the SCGFailure occurs and (2) the median throughput with active SCG cells. Figure 11 plots the results.

In terms of absolute throughput loss, **T** performs worse than **A** (and **V**). **T** loses more than 100 Mbps in 40% (**D1**) and 65.5% (**D2**) of instances, even with the absolute throughput loss up to 326 Mbps (**D2**). In contrast, **A** loses much less than **T**, with its absolute loss below 70 Mbps in most instances; Note that **A** experiences distinct throughput loss in these two datasets: the median throughput loss is below 40 Mbps in **D1** and even below 5 Mbps (1.6 Mbps) in **D2**. This is also caused by various 5G deployment as explained above.

It is worth noting that the median throughput loss of unnecessary failure handling (**U**) is much larger than **M** in **D1**; For **A**, it declines from 105 Mbps to 34 Mbps; For **T**, the loss of missed recovery is more diverse but its median throughput loss also decreases from 87 Mbps to 26 Mbps. It implies that **U** poses more negative impacts than **M**, in terms of absolute throughput loss for **A** and **T** in **D1**.

In term of relative loss, data throughput reduces more than by half (say, relative loss = 100%) in more than 50% instances in **D1** (**A**: 57.7%, **T**: 60%, **V**: 50%). For **T** in **D2**, download speed declines by more than one order of magnitude in 96.6% of instances. We note that the relative loss due to missed recovery (**M**) is higher the one due to unnecessary handling (**U**), which is different from the conclusion in terms of absolute throughput loss. It is because that absolute data speed without failure recovery is smaller in **D2**. It is not hard to understand; **D2** is collected in West Lafayette, a much smaller city. Compared to **D1**, both **A** and **T** offer lower data speed in **D2**, regardless of the use of 5G.

Repeated SCGFailure handling (R).   In our study, most repeated SCGFailures are observed in three settings: **A** (**D1**), **T** (**D1**)

and **A** (**D2**), as shown in Figure 5. We thus use them to assess performance impacts of repeated failures (**R**). Figure 10 plots the cumulative distribution functions (CDFs) of impact time and throughput loss. In every **R** instance, we use the interval from the first failure to the last one as the impact time, which is actually a lower bound of the actual impact time; The throughput loss is calculated as the absolute gap between the average throughput during the impact time and the median throughput without SCGFailures at the same location. We have two observations.

First, the impact time lasts much longer in **A** (**D1**) than **A** (**D2**) and **T** (**D1**). In **A** (**D1**), repeated failures lasts more than 30s in 40% of instances and even goes up to >200s. In **A** (**D2**) and **T** (**D1**), most repeated failures are shorter than 5s. Second, **A** (**D2**) has the minimal throughput loss, which is somehow consistent to those observed in the **U** and **M** instances. It is because **A** offers low data speed even without failures in **D2**. In terms of throughput loss, **T** (**D1**) is worse than **A** (**D1**), despite shorter impact time. In **D1**, **T** loses more than 30 Mbps in 50% of instances and **A** in 33% of instances.

4 RELATED WORK

To our best knowledge, this is the *first* measurement study to reveal problematic SCGFailure handling in reality. This study was inspired by our recent work to examine misconfiguration in 5G networks as the number of serving cells advances from 1 to N [13]. Different from dependent configuration in the multi-round of serving cell selection studied in [13], we focus on SCGFailure handling, particularly when failure handling goes wrong. In the literature, SCGFailures have been studied in several studies but they are mainly on optimization algorithms in order to reduce interruption [6, 12] and save energy by blocking cells [8, 11].

5 CONCLUSION AND DISCUSSION

We have conducted the first measurement study to characterize how 5G networks handle secondary radio access failures in the US. Although such failures are not common, most failure instances are not handled properly in three forms (**U**, **M**, **R**), resulting in unnecessary and significant performance degradation.

This work is still at its early stage; There are many remaining issues, including but not limited to (1) measuring and understanding performance impacts on popular streaming and latency-sensitive applications other from file downloading, (2) quickly fixing the link layer, particularly avoiding repeated and unnecessary failures, and (3) designing cross-layer or higher-layer algorithms (on TCP congestion control and application) to tame improper SCGFailure handling and alleviate its negative impacts. Last but not least, we would like to highlight that problematic SCGFailure handling significantly hurts performance because 5G currently uses non-standalone (NSA) with 5G as secondary radio access. Performance impacts of problematic SCGFailure handling should be much smaller when 5G advances to standalone (SA) and serves as master radio access. However, problematic failure handling may occur with master radio access which will not only hurt data performance but access availability (access is interrupted with such failures).

REFERENCES

[1] 2024. 5G SCGFailure Datasets. https://github.com/mssn/scgfailure.

[2] 3GPP. 2022. Carrier Aggregation on Mobile Networks. https://www.3gpp.org/technologies/carrier-aggregation-on-mobile-networks.

[3] 3GPP. 2023. TS36.101: E-UTRA; User Equipment (UE) Radio Transmission and Reception. V16.16.0.

[4] 3GPP. 2023. TS37.340: NR; Multi-connectivity; Overall description; Stage-2. V16.12.0.

[5] 3GPP. 2023. TS38.331: NR; Radio Resource Control. V16.12.0.

[6] Xinran Ba. 2021. QoS-forecasting-based intelligent flow-control scheme for multi-connectivity in 5G heterogeneous networks. *IEEE Access* 9 (2021), 104304–104315.

[7] Ericsson. 2020. How to tackle fast recovery from radio link failure. https://www.ericsson.com/en/blog/2020/9/fast-recovery-from-radio-link-failure.

[8] Kailash Kumar Jha, Alok Kumar Jangid, et al. 2020. Efficient Algorithm to Reduce Power Consumption for EUTRA-New Radio Dual Connectivity RAN Parameter Measurements in 5G. In *5GWF*.

[9] Yanbing Liu and Chunyi Peng. 2023. A Close Look at 5G in the Wild: Unrealized Potentials and Implications. In *INFOCOM'23*.

[10] OpenSignal. 2023. 5G User Experience Report USA. https://www.opensignal.com/reports/2023/07/usa/mobile-network-experience-5g.

[11] Todd Ou, Qin Zhang, and Xiantao Sun. 2021. Dynamic Blocking of 5G Cells in Non-Standalone Networks. (2021).

[12] Carlos Pupiales, Daniela Laselva, and Ilker Demirkol. 2022. Fast Data Recovery for Improved Mobility Support in Multiradio Dual Connectivity. *IEEE access* 10 (2022), 93674–93691.

[13] Zhehui Zhang, Yanbing Liu, Qianru Li, Zizheng Liu, Chunyi Peng, and Songwu Lu. 2023. Dependent Misconfigurations in 5G/4.5G Radio Resource Control. In *CoNext*.