How Voice Call Technology Poses Security Threats in 4G LTE Networks

Guan-Hua Tu, Chi-Yu Li
Department of Computer Science
University of California, Los Angeles
Los Angeles, CA 90095
{ghtu,lichiyu}@cs.ucla.edu

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

Songwu Lu
Department of Computer Science
University of California, Los Angeles
Los Angeles, CA 90095
slu@cs.ucla.edu

***Abstract*—To support voice calls vital to mobile users and carriers, 4G LTE cellular networks adopt two solutions: VoLTE (Voice Over LTE) and CSFB (Circuit-Switched FallBack). In this paper, we disclose that both schemes are harmful to mobile users from a security perspective. The adoption of the latest VoLTE allows an attacker to manipulate the radio resource states of the victim's device in a silent call attack, thereby draining the victim's battery 5-8 times faster. CSFB exhibits two vulnerabilities of exposing 4G↔3G network switch to adversaries. This can be further exploited to launch ping-pong attacks where mobile users may suffer from up to 91.5% performance downgrade, or 4G denial-of-service (DoS) attacks where mobile users are deprived of 4G LTE connectivity without their consent. We devise two proof-of-concept attacks as showcases, and demonstrate their viability over operational LTE networks. We analyze their root causes and uncover that the problems lie in seemingly sound design decisions for functional correctness but such choices bear unexpected and intriguing implications for security design. We finally propose remedies to mitigate the attack damage.**

I. Introduction

4G LTE (Long Term Evolution) is the latest cellular network technology to offer universal mobile and wireless access to smartphones and tablets. As December of 2014, there have been 367 commercial LTE networks in 121 countries [12]. By 2017, the number of LTE connections worldwide is expected to exceed one billion, with 5.7-fold increase up from 176 millions in 2013 [4].

The LTE network adopts an all-IP, Internet based design, offering much higher access speed (*e.g.*, 100–300 Mbps). Unlike its legacy 3G system, which supports dual modes of circuit-switched (CS) and packet-switched (PS) operations, LTE uses PS only. This decision is partly inspired by the great success of the Internet technology, and partly driven by the explosive demands for mobile broadband services. Mobile Internet data traffic is projected to explode by 10-fold from 2014 to 2019, reaching 24.3 exabytes per month by 2019 [14].

While PS is good for data, it does not well support voice, which is still a killer service vital to cellular subscribers. Historically, a prominent feature of the cellular network has been its carrier-grade voice service. In LTE, two voice solutions are proposed accordingly: CSFB (Circuit-Switched FallBack) [8] and VoLTE (Voice over LTE) [3]. CSFB leverages the CS domain in the legacy 3G systems[1] to support voice calls for LTE users. Whenever a call is made, CSFB transfers the call request from the 4G network to the 3G system. Once the call completes, CSFB moves the phone back to the 4G network. In contrast, VoLTE supports voice calls directly in the 4G system. It leverages the Voice-over-IP (VoIP) solution over the Internet, and still offers guaranteed Quality-of-Service (QoS) through resource reservation in LTE networks.

Both voice solutions are foreseen to coexist in the long run [5]. CSFB leverages the deployed legacy system and works with most current phone models (whereas VoLTE requires new phones). It thus offers a cost-effective, readily-accessible solution. As the most popular voice solution to date, CSFB has been widely deployed or endorsed by most LTE carriers such as top global carriers (China Mobile, Vodafone, Bharti Airtel, Telefonica, AT&T, T-Mobile, to name a few). On the other hand, VoLTE promises to be the ultimate solution. Due to its higher cost of upgrading mobile networks and phones, its current deployment is not as popular as CSFB. In US, a leading VoLTE market, three major operators (AT&T, T-Mobile and Verizon) have started to launch VoLTE until late 2014. In a nutshell, both are projected to survive. CSFB is the prevalent solution now and continues to be appealing in developing countries. Meanwhile, VoLTE will gain its widespread usage in the long run.

In this work, we uncover that both VoLTE and CSFB might be considered harmful from the security perspective. These vulnerabilities are not due to engineering glitches or implementation bugs, but rooted in the technology fundamentals. In VoLTE, its PS design changes its conventional voice call signaling flow and allows an attacker to remotely manipulate Radio Resource Control (RRC) [10] state at the callee's device via delivering certain call signaling messages. Consequently, it can trap the victim device into a high-power RRC state and drains its battery fast. In CSFB, any third party, including a malicious user, may trigger a switch from 4G→3G at the callee device any time without the callee's consent. Such an inter-system migration not only disrupts ongoing data sessions, but also degrades to the slower 3G access afterwards. Moreover, PS and CS have unexpected coupling effects in CSFB. The complex signaling operations in the CS domain

[1] 2G network can be used in the absence of 3G. We use 3G in the paper since most carriers have advanced to 3G/4G.

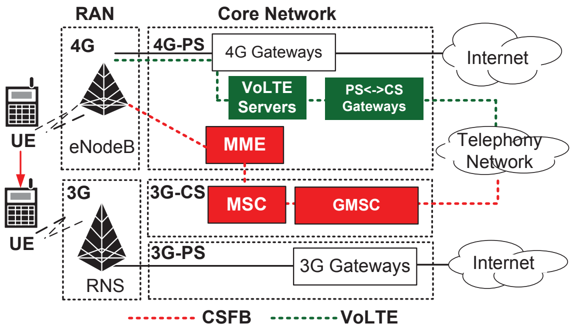


Fig. 1: 4G/3G network architecture supports CSFB and VoLTE.

bear unexpected consequences to the PS domain and even deprive mobile users of 4G access under certain conditions.

As a consequence, 4G users are vulnerable to two attacks that exploit VoLTE and CSFB: silent call attack and coercive ping-pong attack. In the silent call attack, an attacker sends certain VoLTE call signaling messages towards the victim and keep it staying in the high-power RRC state (*i.e.*, `CONNECTED`). In the ping-pong attack, a malicious hacker repetitively dials the victim's phone and hangs up before the ringtone is played. The victim suffers from frequent inter-system switches and oscillates between 4G and 3G networks. We further devise another attack variant where 4G access is eventually denied due to frequent ping-pong attacks. We implement and assess the proof-of-concept attacks over three carriers: two in the US and one in Japan. We find that the former attack leads to 5-8x battery drain and the latter incurs 49.8% - 91.5% throughput slump or even mobile application aborts in the worse case. Our analysis shows that current security mechanisms (*e.g.*, call blacklist, firewalls) are insufficient to defend against such attacks. One thing worth noticing is that these attacks do not require extra capability but a commodity, programmable smartphone. They are ready to launch, imposing real threats to mobile users. So carriers and vendors should take immediate actions in both of VoLTE and CSFB. Otherwise, billion of LTE mobile users will suffer from malicious attacks.

The rest of the paper is organized as follows. §II introduces both voice solutions and then gives an overview of our security study. §III analyzes VoLTE security and presents silent call attack design and validation. §IV analyzes CSFB security, presents and assesses coercive ping-pong attacks. §V proposes remedies. §VI compares with the related work, and §VII concludes the paper.

II. Voice Solutions in 4G LTE

Despite increasing popularity of mobile data services, voice is still a killer application to mobile users and carriers. To ensure guaranteed service quality, voice has been traditionally supported via the CS technology, which establishes a virtual circuit and reserves resource for each call. In contrast, data has been delivered in packets through IP-based technology in a best-effort fashion. In this section, we introduce 4G/3G network architecture and two voice solutions to 4G LTE networks, and then give a brief overview of our security study.

A. 4G/3G Network Architecture

Figure 1 depicts the 4G/3G network architecture. We focus on the most widely deployed 4G and 3G networks, *i.e.*, LTE and HSPA (High-Speed Packet Access)[2] [6]. Both infrastructures have two subsystems: the RAN (radio access network) and the core network. The major RAN components are base stations, *i.e.*, eNodeB (Evolved Node B) in 4G or RNS (Radio Network Subsystem) in 3G, which offer wireless radio access to user equipments (UEs), *e.g.*, phones. The core network bridges the RAN and the external networks, *e.g.*, the Internet or the telephone network.

The 4G LTE network supports PS only. It offers higher speed, thus being preferred by users when available. The PS core has two critical components: gateways and MME (Mobility Management Entity). 4G gateways route data packets between RAN and the Internet, akin to edge routers in the Internet. MME is the key control node for 4G RAN and manages radio access and user mobility, *e.g.*, tracking and paging each phone. To properly support VoLTE, two more elements are deployed. VoLTE servers are used to forward call traffic in the PS domain from/to 4G gateways, whereas PS↔CS gateways translate VoIP traffic and its control signals to/from the telephone network.

In contrast, 3G uses the lower-speed, dual-mode infrastructure. Its core network supports both PS and CS, for data and voice, respectively. Its PS gateways are similar to those in LTE. The CS core supports voice service via two elements: MSC (Mobile Switching Center) and GMSC (Gateway MSC). The former is responsible for paging the UEs in the CS domain and establishing voice calls, whereas the latter routes calls between MSC and the telephone network.

B. CSFB Primer

The core idea of CSFB is to *on-demand* leverage the existing CS domain in 3G to serve CS-based voice calls for 4G users. This way, carrier-grade call quality is ensured. By default, mobile users stay in 4G networks. Upon a call request, either inbound or outbound, CSFB immediately migrates the user from 4G to 3G and then serves the subsequent call. Once completed, it moves back to 4G if still available.

Figure 2 presents a simplified procedure on how an inbound voice call is handled in CSFB. When a 4G phone is called, the incoming call request is first routed to GMSC/MSC in 3G networks. MSC subsequently requests MME to page the phone in 4G. Once the phone is found, MME migrates it from 4G LTE networks to 3G networks via triggering an *inter-system switch* (from 4G to 3G). After the phone successfully connects to 3G RANs, MSC establishes the voice call with the phone, through the standard procedure of call establishment (Steps 3 – 5) [9]. Specifically, the phone responds to the paging request and attempts to set up the call with MSC. As long as all resource needed by the call is reserved, it alerts MSC, which in turn alerts the caller (*e.g.*, generates a ringtone). In the meantime, the callee's phone rings to notify the user about the incoming call. Once the call is answered/accepted (Step 6),

[2] The 3G architecture is also applicable to other 3G technologies — HSPA+ (enhanced HSPA) and UMTS (Universal Mobile Telecommunications System), its predecessor.

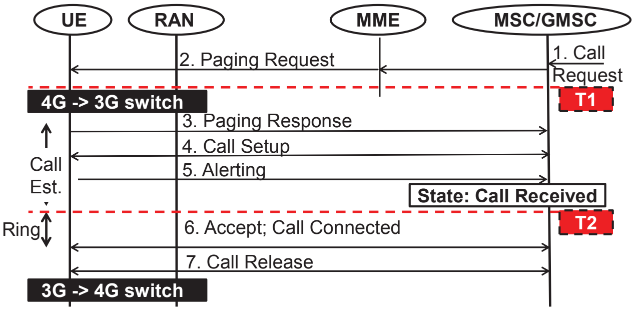


Fig. 2: The simplified procedure of an inbound call in CSFB.

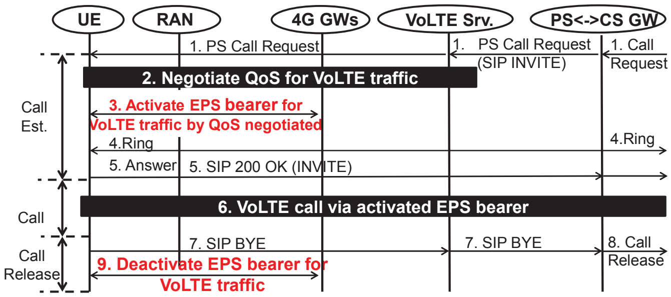

UE
RAN
4G GWs
VoLTE Srv.
PS<->CS GW
1. PS Call Request
1. PS Call Request (SIP INVITE)
1. Call Request
2. Negotiate QoS for VoLTE traffic
Call Est.
3. Activate EPS bearer for VoLTE traffic by QoS negotiated
4.Ring
4.Ring
5. Answer
5. SIP 200 OK (INVITE)
6. VoLTE call via activated EPS bearer
Call
7. SIP BYE
7. SIP BYE
8. Call Release
Call Release
9. Deactivate EPS bearer for VoLTE traffic

Fig. 3: The simplified procedure of an inbound call in VoLTE.

the call is connected and MSC helps to deliver voice messages between the caller and the callee. When the call ends, the call connection is released (Step 7), followed by the *second inter-system switch* which migrates the phone from 3G back to 4G. For an outgoing call, the procedure is similar, except that the request is directly sent to MSC via MME (Steps 1–3 skipped) which initiates 4G→3G switch for a CSFB call.

In a nutshell, two inter-system switch events (*i.e.*, 4G→3G, and 3G→4G) are triggered during the lifetime of each CSFB voice call. In practice, the switch is realized via a handoff or a cell reselection procedure [8]. The handoff changes the serving network during ongoing calls/data sessions, while the cell reselection is performed during idle. When the phone also has ongoing data sessions (*e.g.*, background services) during the voice call, all data packets will be delivered through the 3G network until the second switch from 3G to 4G completes. This is because most phones only have one set of radio hardware, which can only work in one network (either 3G or 4G) at a time. During this period, data delivery traverses 3G networks, instead of the faster 4G LTE networks, thereby slowing down the access for data services. For few phones support two sets of radio hardware, they can concurrent access two networks (3G or 4G) and might not experience the low-rate data services during call conversation at the cost of extra hardware and radio resource.

C. VoLTE Primer

The alternative solution is VoLTE (voice-over-LTE) [3]. It directly serves voice calls over PS-based IP core networks, akin to VoIP over the Internet. It has two distinctive features from CSFB. First, it uses PS to carry voice (see Figure 1). It sets up an EPS (Evolved Packet System) bearer (also used for data transfer) to deliver voice traffic between the user device and the 4G gateways. After that, voice traffic is further routed by VoLTE servers and PS↔CS gateways to reach the telephone network. Second, its control is done through Session Initiation Protocol (SIP), but not by CS signaling (*i.e.*, SS7 [2]). SIP is widely used for voice and video call signaling over the Internet [24]. SIP messages are exchanged within LTE networks (*e.g.*, between VoLTE servers and PS↔CS gateways). PS↔CS gateway acts as the bridge connecting two networks over SIP and CS.

We still use the example of an inbound call to illustrate how VoLTE works. Figure 3 shows a simplified procedure to establish and release a VoLTE call. The call request from the telephone network (using CS signaling) is first routed to the PS↔CS gateways. After being translated into a PS call request (*i.e.*, `SIP INVITE`), it is delivered to the VoLTE server. The server further initiates the VoLTE call setup with the callee. It thus negotiates the Quality of Service (QoS) contract with the callee (Step 2), and an EPS bearer for VoLTE is then activated by 4G gateways (Step 3). Afterwards, the call rings on both sides. Once the callee answers the call, an OK message is delivered, which eventually establishes this call. During the call, all voice traffic is carried by the activated EPS bearer, similar to normal PS traffic, but at higher priority (Step 6). When the call is about to end (here, the callee hangs up), BYE signaling is invoked and the call is released. The EPS bearer for VoLTE is deactivated thereafter (Step 9).

VoLTE seeks to provide comparable or better quality than typical CS calls. This is a key difference from the popular VoIP over the Internet. To this end, VoLTE does reserve resource for each voice call within cellular networks, and release it once the call ends. This is done through the *EPS Bearer Context Activation/Deactivation* procedures [11]. Moreover, QoS is configured during the bearer activation. VoLTE signaling and data are set to higher priorities over other non-VoLTE traffic (*i.e.*, conventional PS data like web browsing).

D. Vulnerabilities in Both Voice Solutions

We find that both VoLTE and CSFB are vulnerable to security breaches unanticipated by their designers. In VoLTE, the latest practice of VoLTE provides attackers with a possible exploit to launch silent call attack towards mobile users. While malicious SIP call signaling messages are exchanged between the caller and the callee, the callee's phone stays busy and consumes 5-8 times power consumption than the standby one. The victim is unaware of the attacks because no tone rings or no incoming call pops up. Fundamentally, VoLTE is vulnerable to fine-grained manipulation of signaling messages, which is almost impossible in the traditional CS network. Its operations which rely on SIP, not CS signaling, expose voice calls to remote manipulation without consent from the callee.

In CSFB, mobile users are exposed to security threats of losing control on their own network access. Their access can be hijacked by an adversary without hacking into their devices (i.e., no Trojan or malware is required). Specifically, two vulnerabilities can be exploited for attacks: (1) a 4G→3G switch can be triggered by anyone without the callee's consent, and (2) a subsequent 3G→4G switch can be deferred or disabled by others. We devise a new ping-pong attack against a chosen victim user (given a phone number) in 4G LTE networks. This attack starts with an unnoticed dialing , which exploits the first vulnerability and forces the callee (victim) to switch from 4G to 3G without awareness. By repeating of the unnoticed dialing (dials→hangs up→dials) with sophisticated calling interval, attackers can make victims to either downgrade transmission rate up to 91.5%, tear down all TCP connections in few minutes. We further identify an attack variant where users will be even deprived of 4G LTE connectivity.

Threat Model We assume a modest adversary model without giving too much attack capability. The adversary only has full control over its own smartphone and/or a remote server (outside the victim's cellular network). The phone or the server is a commodity device, but it is programmable. The adversary has no access to the cellular core infrastructure or other devices. It solely relies on the public available information when launching attacks. We assume that neither the victim phone nor other components in the cellular infrastructure are compromised. This model seeks to investigate the vulnerabilities in CSFB and VoLTE, rather than aggravating their damage with more powerful attacks.

Responsible Experiment Settings We perform experiments in operational carriers to only validate identified vulnerabilities and the feasibility of possible attacks. Our study is conducted in a responsible manner. We carry on tests strictly in a controlled setting and confine our experiments to a small scale. In case of certain feasibility study and attack evaluation detrimental to other users, we only use our own phones as the victims. We seek to limit, rather than aggravate, the damage on operational networks as much as we can (albeit it is viable by reasoning).

We test with three LTE operators, including two Tier-1 US carriers and a major Japanese operator. They are denoted as `US-1`, `US-2` and `JP-1` for privacy concerns. All support CSFB, whereas only `US-2` supports VoLTE in our tested area. We use seven smartphone models: HTC One, LG Optimus G, Samsung Galaxy S3, S4 and S5, and Apple iPhone5 and 5S, covering both Android and iOS systems. All seven models support CSFB, while only S5 supports VoLTE. Each experiment has 10 runs unless explicitly specified.

III. Silent Call Attack

We first disclose that mobile users are exposed to silent call attack, where the victims are unaware of abnormal call states manipulated by an adversary but suffer from high energy waste incurred. This is due to one VoLTE vulnerability where the signaling messages for call establishment, as well as its corresponding RRC state, can be deliberately manipulated by another VoLTE caller without the callee's consent or even awareness. In this section, we elaborate on its vulnerability analysis, attack validation and assessment.

A. *Vulnerability: Fine-grained Manipulation of Call Signaling*

In typical 2G/3G CS voice call setup, the callee's ringtone plays immediately as long as the first message of the call setup, `IAM` (Initial Address Message, Call signaling), is received. During the delivery of `IAM`, the resource (e.g., a voice tunnel) is reserved en route at the intermediate switch centers (*e.g.*, MSC) between the caller and the callee. However, in VoLTE, the call setup procedure is different since it adopts PS, not CS.

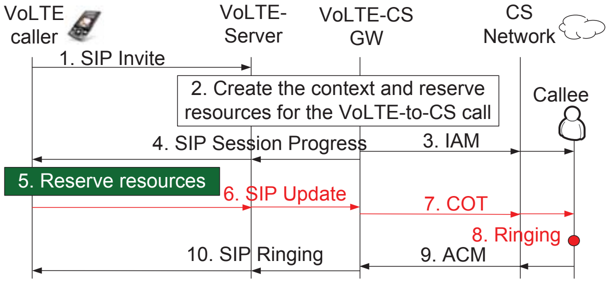
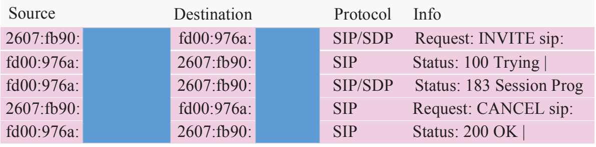


Fig. 4: The call setup procedure between a VoLTE caller and a CS callee. The callee's ringing event is controlled by the VoLTE caller.



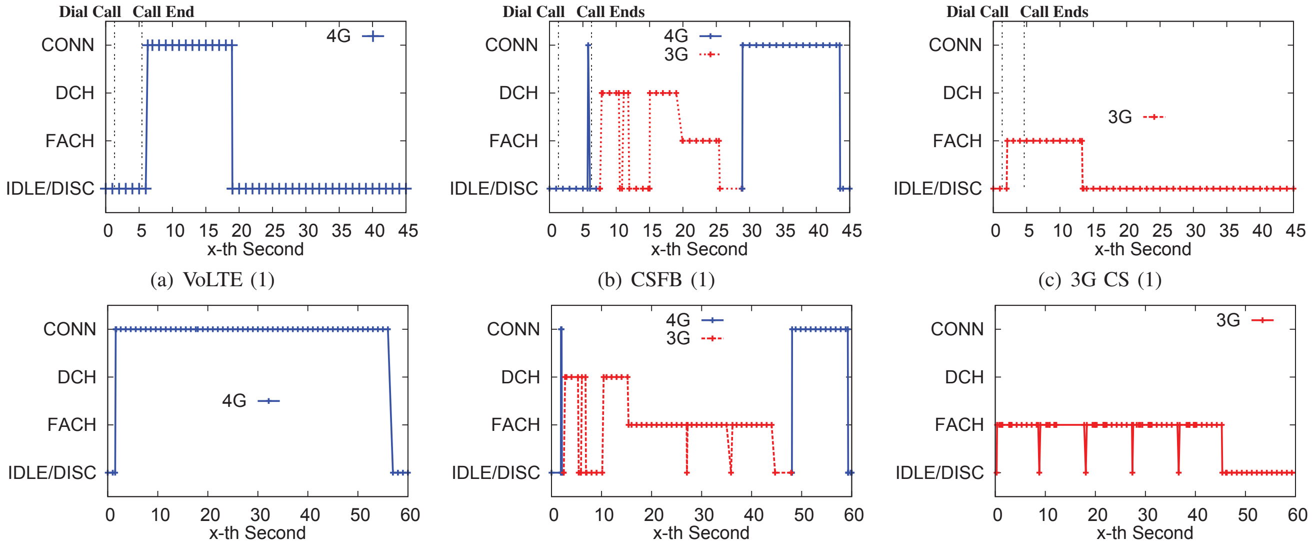
Source		Destination		Protocol	Info
2607:fb90:	[redacted]	fd00:976a:	[redacted]	SIP/SDP	Request: INVITE sip:
fd00:976a:	[redacted]	2607:fb90:	[redacted]	SIP	Status: 100 Trying |
fd00:976a:	[redacted]	2607:fb90:	[redacted]	SIP/SDP	Status: 183 Session Prog
2607:fb90:	[redacted]	fd00:976a:	[redacted]	SIP	Request: CANCEL sip:
fd00:976a:	[redacted]	2607:fb90:	[redacted]	SIP	Status: 200 OK |

Fig. 5: Wireshark traces obtained on the VoLTE caller device. The first 8 hexadecimal digits of IPv6 addresses are shown to distinguish the source and the destination while hiding their actual identities.

Figure 4 illustrates how a VoLTE user calls another 2G/3G CS voice user. Different from the conventional CS voice calls, receiving an `IAM` at the callee only indicates that the voice tunnel between VoLTE-CS GW and the callee has been established. However, the voice tunnel between the VoLTE user and VoLTE-CS GW is not ready yet. It will be established until VoLTE-CS GW receives `SIP Update` from the VoLTE caller (Step 6). As a consequence, VoLTE-CS GW enables the `COT` (Continuity Test) feature where the callee's device must wait for the `COT` message before playing the ring tone. This is not designed without rational. It is able to reduce the VoLTE-CS call setup duration (*i.e.*, resource reservation goes in parallel). However, it is thus exposed to one possible exploit where the adversary can make successive call attempts silently without the victim's awareness. S(he) can deliberately avoid sending the `Update` message out, thereby suppressing the ring tone at the victim.

Vulnerability validation We implement a testing tool *SilentAutoCall* which makes an outgoing VoLTE call, and immediately hangs up when the `Session Progress` message is received from the VoLTE server (Step 4 in Figure 4).We aim to prevent `SIP Update` from being delivered to VoLTE server. Figure 5 shows the VoLTE signalling messages exchanged between *SilentAutoCall* and the VoLTE server. Note that all VoLTE signallings are encrypted in the IPsec messages. To decrypt them, *SilentAutoCall* retrieves the keys from the Android OS using the command "`ip xfrm state`". We further run *SilentAutoCall* on the VoLTE-capable phone (here,, Samsung Galaxy S5) to make outgoing calls. On the callee's mobile phone, we activate the engineer mode[3] to collect the logs of the Radio Resource Control (RRC [7], [10]) layer, which mandates the radio connection between the user device and the base station. We monitor RRC state transitions and check whether call signaling messages are successfully delivered; For example, `DISCONNECTED→CONNECTED` (`DISC→CONN`) implies

[3] E.g, dials `*#197328640#` on Samsung Galaxy S5.



(a) VoLTE (1)

(b) CSFB (1)

(c) 3G CS (1)

(d) VoLTE (5)

(e) CSFB (5)

(f) 3G CS (5)

Fig. 6: The RRC state transitions at the VoLTE, CSGB and 3G CS callees when incoming silent call(s) are made from a VoLTE caller (upper: one single silent call, lower: five successive silent calls).

that the radio connection between the callee and the base station has been established to deliver packets or signaling message (it occurs at Step 3, where IAM is received.). During the experiments, we disable all background mobile data, which may also affect the RRC state transition.

For three types of callees including VoLTE, CSFB and 3G CS, we have 50 runs in each setting. We do not observe any ring tone or incoming call popup from all the test phone models. It shows that the callees do not detect our silent calls no matter what phones they are using. Figure 6(a), 6(b) and 6(c) plot the RRC state transitions of the VoLTE, CSFB and 3G users' devices, respectively. *SilentAutoCall* ends the outgoing VoLTE call around the 5^{th} second but this has already affected the RRC state at the victim. The RRC state is switched to a connected state (i.e., VoLTE:4G CONN, CSFB:4G CONN/3G DCH, CS:3G FACH), which consumes more power than usual.

B. Attack and Evaluation

With this vulnerability in mind, we devise a proof-of-concept attack to make silent calls to drain the battery of the victim's phone.

Attack design The attack idea is to force the victim's device stuck into a high-power RRC state for a long while without any awareness. It is done by sending successive SIP calling messages. To this end, we modify the *SilentAutoCall* to repeat the following three steps: (1) dials (sends INVITE to the VoLTE server), (2) waits (receives Session Progress from the VoLTE server) and (3) hangs the outgoing call up (sends Cancel to the VoLTE server) towards the victim.

Evaluation Figures 6(d), 6(e) and 6(f) compare the RRC state transitions at the VoLTE, CSFB and 3G CS callees when the VoLTE caller makes five successive call attempts. We observe that all mobile devices almost stays in RRC connected states (4G:CONN and 3G:DCH/FACH) during the attack period (the 5^{th} silent call attack finishes at around $40-45^{th}$ seconds). This implies that a malicious attacker is able to force any victim' device not to enter IDLE/DISC states to reduce energy consumption. This attack can drain the device battery fast. This is even worse for a VoLTE callee where RRC stays in 4G CONN; This consumes more power than CSFB/3G CS with FACH which consumes less power than DCH.

	VoLTE			CSFB/3G CS		
	Attack	No-attack	$\frac{attack}{no-attack}$	Attack	No-attack	$\frac{attack}{no-attack}$
Battery usage	**24%**	3%	**800%**	**20%**	4%	**500%**

TABLE I: Battery usage for VoLTE, CSFB and 3G CS callee using Samsung Galaxy S5 after 6-hour silent signaling attack. Battery drain is represented by the ratio of $\frac{attack}{no-attack}$.

Table I summarize the battery usage at three VoLTE, CSFB and 3G CS callees after a 6-hour silent call attack. For fair comparison, we use the same phone model (Samsung Galaxy S5) for all callee types. We configure the voice type at the Settings menu. We make two observations. First, the battery indeed drains *5-8* times faster under attack than that without attack. The VoLTE callee suffers more, with 8x energy consumption under attack. We measure radio power using a Monsoon power monitor [1]. In particular, radio power in non-IDLE is about 400-1200 mW, 40-240x larger than the IDLE one (5-10 mW). This result is consistent with other prior measurement [22]. Second, we do not observe obvious distinction between 4G CSFB and 3G CS callees in terms of battery usage. This is because the CSFB callee never returns to 4G LTE networks (stays in 3G) during the 6-hour attack. Thus, the CSFB callee has the same usage as the 3G CS callee.

IV. Ping-Pong Attack

In this section, we explore vulnerabilities in CSFB. We find that CSFB allows an adversary to force a victim to frequently switch its serving networks without consent and the current shield fails to work. We accordingly devise a Ping-Pong attack and its variant on 4G DoS, and assess their damage in reality.

A. Vulnerability: Intervening 4G↔3G switch in CSFB

There are two vulnerabilities that intervenes two inter-network switches in CSFB. Both 4G↔3G switches suspend users' data transmission for seconds.

First, with CSFB, it becomes possible to force any 4G user to downgrade to 3G networks without his/her consent. In fact, the user is even unable to block this involuntary switch. It can be manipulated by any party who dials a voice call toward a given number (Later, blacklist will be discussed). Even worse, such a switch happens without any user involvement. We analyze the root cause. This vulnerability is rooted in one sound design component to support carrier-grade voice. A remarkable feature of CS calls is to employ control signaling throughout the call procedure in order to meet the stringent voice quality requirement. For a user in 4G LTE, migrating to 3G is the prerequisite for leveraging CS signaling in the 3G network. Consequently, the 4G→3G switch is triggered upon a request, before the call is accepted by the callee. However, such involuntary design is questionable from the security perspective. The callee who experiences such a switch and its subsequent performance degrade, has no power to reject it.

Second, the 3G→4G switch after a CSFB call, can be postponed or even disabled under certain conditions. The problem lies in the concrete procedure used to realize the 3G→4G switch. This switch is mandated by RRC and use different switching conditions in different RRC states (`IDLE` or `CONN`). However, RRC states are together determined by both CSFB calls and data services. This coupling effect between CS calls and PS data builds a persistent loop so that the victim device cannot escape and return to 4G. The loop details can be found in our prior study [28].

In summary, given a phone number, any caller (via the Phone or VoIP) is able to manipulate the 3G/4G network state at the CSFB callee. This possesses a similar issue in VoLTE where RRC states are controlled even without asking for permission.

B. Attack and Evaluation

Attack Design The attack is launched by repeatedly dialing the victim user before the call is through. The attacker can use *SilentAutoCaller* described in Section III or resides outside cellular networks and leverages VoIP tools (e.g., Skype, Hangout) to dial the victim over the Internet. Figure 7 plots the attack procedure. It consists of successive attack calls (ACs) , each of which forces the victim to perform two switches: 4G→3G and 3G→4G. This is done in two steps: (1) dials the victim and hangs up, and (2) waits until the victim goes back to 4G.

To aggravate the damage, the attacker seeks to make malicious calls as frequently as possible. Performance penalty is incurred by network transitions, thereby growing with the frequency of attack calls. Therefore, two requirements must be met. First, the attack must remain silent so that the victim is unaware of it. Second, each AC has to be employed while the victim is being in 4G; otherwise, CSFB cannot be triggered.

To maximize the frequency of attack cycles, we configure the dialing time to be the minimum interval that successfully triggers 4G→3G but no ring tone is played at callee (i.e., T1 in Figure 2). Each dial thus lasts only 4s, 5s and 3s in US-1, US-2 and JP-1, respectively, by our measurement. For the second requirement, the interval between two attack calls has to be long enough to initiate the second dial once the victim switches back to 4G. This interval can be the upper bound of the sum of the dialing time and two network switch durations (4G→3G and 3G→4G). Figure 8 plots the CDFs of these two durations. Test cases with both strong and weak radio signal strengths are covered. `JP-1` is faster than US carriers. 90% of the 4G→3G and the 3G→4G switches finish within 3s and 1.5s, while taking 7s and 12s in `US-1`, and 17.5s and 19.7s in `US-2`.

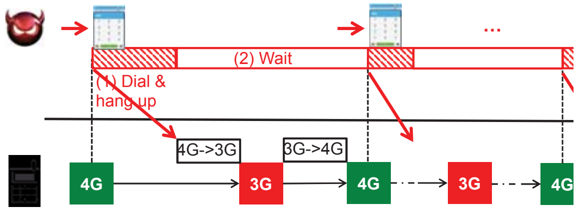


Fig. 7: Illustration of *Ping-Pong attack*. An entire attack that consists of multiple attack calls (ACs): (1) dials and hangs up, and (2) waits.

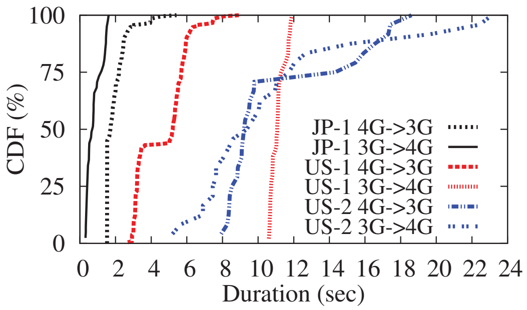


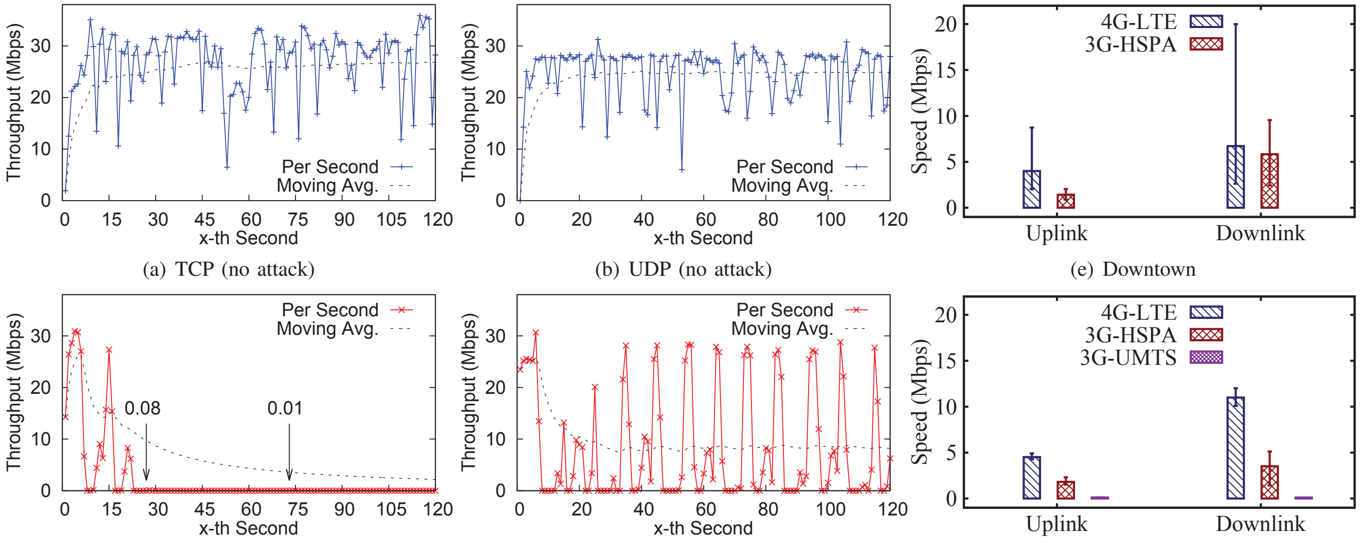
Fig. 8: CDFs of the 4G→3G and 3G→4G switch durations.

In summary, a simple attack is ready to launch. It dials, hangs out, waits, and then repeats. Compared with the two US carriers, the victims in `JP-1` could suffer even more. This is because the attacker can hang up earlier and wait for a shorter time without user awareness. Specifically, in `JP-1`, we set the dialing time as 3s and waiting time as 5s in our attack prototype. Note that, we do not intend to explore all attack options since there exist sufficient flexibility in configuring the attack. Instead, our goal is to verify the attack feasibility and unveil its impact factors and possible damages.

Evaluation We assess the damage from four aspects: TCP/UDP performance degradation, lost network connectivity, and impacts on applications.

○ *TCP Performance* Figure 9(a) and 9(c) compare the first 2-minute throughput of a TCP connection in non-attack and under-attack cases. Under the attack, data throughput initially oscillates between 0 Mbps and 30.9 Mbps and then freezes (0 Mbps) after the 22nd second (except the 27th and 73rd seconds). The average throughput shrinks from 26.7 Mbps to 2.0 Mbps, with a 91.5% decline. The TCP connection will be eventually terminated due to excessive retransmission attempts, once it is compelled to freeze for a while.

Such large performance slump in TCP matches with our



(a) TCP (no attack)

(b) UDP (no attack)

(e) Downtown

(c) TCP (under attack)

(d) UDP (under attack)

(f) Suburb

Fig. 9: Impact of Ping-Pong attacks. (a)(b) TCP/UDP throughput without attack; (c)(d) TCP/UDP throughput under attack; (e)(f) uplink and downlink throughput in the downtown or suburb.

expectation. It is mainly attributed to its congestion control and the exponential backoff mechanism in its retransmission timeout (RTO). Data suspension results in packet losses and triggers retransmissions; this further reduces the congestion window size. If the retransmission occurs when data suspends, no response from the victim will be received and it further doubles the RTO. With consecutive retransmission failures, the RTO dramatically increases. Even worse, the server cannot resume its transmission right after data suspension stops. It has to wait for the expiration of RTO. In our experiments, TCP connection is tore down within 30 minutes.

○ *UDP Performance* UDP also gets hurted but the harm is less than TCP. Figure 9(b) and 9(d) plot the UDP throughput in both cases, with a 40 Mbps constant-rate traffic source. Each session runs for 2 minutes. The average throughput drops from 24.9 Mbps in the no-attack case, to 8.7 Mbps under attack, leading to 65.1% performance reduction. This is because the instantaneous throughput decreases to 0 Mbps during the suspension periods. In fact, the per-second throughput dramatically oscillates between 0 Mbps and 29 Mbps. Although the throughput degradation is smaller than that of TCP (UDP: 49.8%-69.3%), the quality of UDP applications (*e.g.*, video streaming) may already deteriorate too much to be used. The aggregate suspension occupies 57.3% of the time.

○ *Lost network connectivity* Certain attack calls may even result in lost network connectivity at the victim. It may cause all ongoing UDP/TCP sessions to abort, as long as the victim's IP address changes after it reconnects. The loss lasts 5–20 seconds, during which the victim misses all incoming calls and is unable to access any data service. The probability of this CSFB-incurred connectivity loss is about 3%. Though this probability is not large, successive calls during this attack indeed increase the loss frequency over a short time window.

○ *Impact on popular applications* We launch this attack when each of eight popular applications is running. Table II shows the task tested on each application and the damage incurred. It can be divided into two cases, attacked without (Case-I) and with (Case-II) bearing network connectivity loss. In each experiment, we launch attacks right after the application task starts.

Apps	Task	TCP/UDP	Case-I	Case-II
Web	Access one CNN page	TCP	Abort	Abort
Gmail	Sending/receiving email	TCP	Fail & Multi-retry	Abort & Auto Recover
Facebook Messenger	Ongoing chat session	TCP	Slower	Abort
Whatsapp	Ongoing chat session	TCP	Slower	Abort & Auto Recover
AndFTP	file download	TCP	Abort	Abort
Youtube	video play	TCP	Freeze	Abort
PPStream	video play	UDP	Freeze	Abort
Skype	Ongoing video call	UDP	Freeze	Abort

TABLE II: Impact on popular applications under *ping-pong attack* in two cases without (Case-I) and with (Case-II) network connectivity loss.

We make four observations in Case-I. First, both the FTP client and the Web browsing encounter abort due to connection timeout. The FTP client freezes the download progress in about 30 seconds and then terminates. The browser fetching one CNN page may also abort due to requesting large objects, such as video clips. Second, Youtube, PPStream and Skype stop playing video and voice call when an attack call arrives at the victim. Third, for Gmail, the email send and receive may fail. Although Gmail will re-send/re-fetch emails periodically, users still suffer from longer time to send/fetch emails. Furthermore, for those emails containing pictures or attachments, users will experience more attempts to send/receive them. Fourth, the two popular instant messaging applications take longer time to transmit messages with images.

We have three observations in Case-II. First, all applications abort due to the change of IP addresses. This is because their ongoing sessions are always bound to the IP address. For example, the FTP client will receive the "Broken Pipe" error message from its Android OS. Moreover, we observe that Web browsing takes longer to fetch Web pages after recovering from abort, since all DNS caches on mobile phones are cleared upon network connectivity loss. Second, Whatsapp and Gmail are unable to transmit or receive any instant message during network connectivity loss. Message delivery is postponed until the user device reconnects to cellular networks. Third, the Facebook Messenger cannot send/receive messages as well. The users have to manually retransmit unsent messages.

C. Attack Variant: 4G DoS Attack

We next devise an attack variant that compels the victim to lose the preferred, higher-speed 4G access even when it is available. We exploit the second vulnerability that the second network switch, 3G→4G, can be deferred or disabled by others. In particular, we launch the ping-pong attack with a shorter calling interval (the time intervals for US-1, US-2, and JP-1 are configured to less than 12s, 19.7s and 1.5s, respectively). This aims to get that victim's device stuck in 3G, since there are always new incoming 3G CS calls toward the victim before 3G→4G inter-system switch is triggered.

We find that the victims will get stuck in 3G as long as 4G DoS attack is performed (it lasts for 7 hours in our test). Data throughput decreases and the slump depends on the 3G network technology. We measure downlink and uplink throughput in both downtown and suburb areas and plot the median (minimal and maximal) values in Figures 9(e) and 9(f). In the downtown area, 4G LTE and a high-speed 3G (HSPA)[4] are supported, whereas lower-speed UMTS may be adopted in the suburb areas in case of insufficient HSPA coverage. As a result, this attack can impose data throughput slump as large as 59-76% (uplink) and 13-53% (downlink) in the downtown areas (downgrading to 3G HSPA); The reduction can be even larger (up to 99%) in the suburb areas while it switches from 4G LTE to 3G UMTS.

D. Discussion on Current Shields

The current common shields fail to properly defend against the above attack. For the security mechanisms in cellular networks, network-based blacklist, has to be enabled and configured by the users themselves. However, the victim is still unable to defend this attack via the blacklist. This is because s(he) is unaware of the attack and does not know the attacker's phone number.

We admit that the operators can monitor high-volume call attempts toward victims and block all suspicious phone numbers or control the call dialing rate. However, such fixes might be circumvented by adversaries. For example, attackers can purchase hundreds of pre-paid SIM cards without personal IDs and do not need to always dial the same numbers. Besides, they only mitigate the damages of this attack but are not considered as the ultimate solution.

For the security mechanism at mobile device, the mobile user may install certain device-side blacklist tools (such as CallBlocker [17]) to block disturbing calls (e.g., all calls not on the contact list). Such blacklist tools are popular because they save users from annoying calls. However, we observe that, the blacklist stops disturbing the user but cannot prevent from triggering the CSFB procedure. In fact, if the caller's number is added onto the blacklist (done by harassing the UE with multiple courtesy calls for advertisement), it makes the attack even easier and stealthier. The blacklist automatically rejects the call request after call establishment. However, prior to that, 4G→3G has already been invoked and call has been canceled by attackers.

V. Suggested Remedies

In this section, we suggest immediate remedies. Our proposal seeks to mitigate the attack damage, but not to eliminate it due to practical constraints (e.g., CSFB phones has one radio hardware and cannot support 3G and 4G both).

Silent Call Attack Prevention The VoLTE-CS gateway should not send `IAM` signaling message to the CS callee until VoLTE caller completes its resource reservation for this call conversation to be established. This way, the COT (continuity test) procedure is no longer required. This is backward compatible with the existing CS-based call signaling procedures, at the cost of slightly longer call setup time; Our experimental results show that resource reservation takes less than a second in our tests and this should be tolerable. As a result, silent call attack is eliminated accordingly.

Ping-Pong Attack Prevention Once the operator detects an attack attempt (*i.e.*, consecutive calls are made and terminated within a short time), it should retain the victim inside 3G for certain period of time. The detection can be done at the proxy VoLTE server (i.e., P-CSCF [3]) which has similar functions to session border controllers deployed in VoIP. This way, one can significantly reduce the damage on his/her data services and the large amount of signaling induced by CSFB calls. However, the downside is that the victim suffers from performance degradation when (s)he stays in 3G networks. Moreover, since deterministic patterns can be easily exploited by the ping-pong attack, we suggest that operators randomize network switch timers. Upon timeout, the phone will be migrated back to the LTE network. By this approach, the attackers cannot accurately predict the timing to launch the 4G↔3G switches towards victim and fail to maximize data suspension time of Ping-Pong attacks. For example, the attacker's call request may arrive when victim is still in 3G network and no 4G→3G switch will be triggered.

Unawareness Prevention We advocate that mobile devices should get user's consent for the 4G→3G switch (downgrade). When an incoming call comes to the CSFB user, the MME sends a CS paging notification [8] with the number of the caller to the callee and asks whether the callee accepts to answer this call. If the user declines the call at this point, the phone should not be switched to the 3G network. Note that our proposal differs from the current practice. To the best of our knowledge, all CSFB phones by default respond with "YES" to the MME without user interaction and thus immediately switches to 3G to handle incoming calls.

[4] The max downlink/uplink rate for HSPA is 42/23 Mbps. UMTS supports up to 2 Mbps in both links.

VI. Related Work

Security of cellular networks has become an active research area due to the popularity of mobile devices and applications. In this work, we only review closely related ones. Racic *et al.* exploited MMS to drain the mobile device battery [23], Ricciato *et al.* discovered large-scale resource waste incurred by unwanted traffic in cellular networks [25]. Traynor *et al.* demonstrated how to launch DoS attacks in a target area by leveraging SMS to overload the 2G/3G control channels [26] or by using unwanted traffic from the Internet [27]. Arapinis *et al.* disclosed loopholes in user authentication in mobile telephony systems [18]. Peng *et al.* uncovered overcharging and undercharging threats in mobile data billing [20], [21], whereas Go *et al.* exploited TCP retransmissions to attack the accounting system [15]. Our work differs from all the above by addressing a different problem. We investigate how the voice service poses security threats to 4G LTE networks.

Several recent studies have assessed CSFB and VoLTE solutions, but from the performance viewpoint. Koshimizu *et al.* proposed a mechanism to improve transition from VoLTE to CS call services [16]. Ozturk *et al.* studied the VoLTE performance in heterogeneous LTE networks [19] and Bautista *et al.* assessed the CSFB performance [13]. Our previous work examined mutual interference between voice and data in CSFB and assess their impact on data/voice performance [28]. They all focused on VoLTE or CSFB under common yet non-malicious usage settings. Differ from them, our study explores both of VoLTE and CSFB from the security perspective.

VII. Conclusion

The 4G LTE network is a relatively unexplored area for security evaluation. In this work, we disclose that CSFB and VoLTE may not be a sound voice solution to 4G LTE from the security standpoint. CSFB exposes 3G↔4G network switches to any adversary without any consent from the victim.VoLTE makes the similar mistake, allowing any VoLTE adversary to manipulating call and RRC states stealthily to threaten all types of callees (VoLTE, CSFB, and 3G CS).

Our study also yields two insights. First, the ultimate root cause for vulnerabilities in VoLTE and CSFB lies in seemingly sound design decisions from the functional correctness standpoint. However, such choices may bear unexpected, yet intriguing implications for security. In the worst case, they are prone to attacks. Proper design thus needs to take into account both functions and security at the first place. Second, the control/signaling plane in cellular networks is much more complex than the Internet counterpart. Information sharing and state transition on the control plane have to be carefully crafted. Otherwise, they may lead to more severe attacks than the data-plane loopholes. A state change in the CS domain may impose unanticipated effect in the PS domain. The security implication is that CS can be exploited to degrade the performance of PS.

Acknowledgements

We would like to thank the anonymous reviewers for their valuable comments. This work is supported in part by the National Science Foundation under Grants No. CNS-1421440, CNS-1423576, CNS-1421933 and CNS-1422835. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] Monsoon. https://www.msoon.com/LabEquipment/PowerMonitor/.

[2] Signalling System No.7. http://www.itu.int/rec/T-REC-Q.700/en.

[3] Voice over LTE. http://www.gsma.com/technicalprojects/volte.

[4] Global LTE Network Forecasts and Assumptions, 2013-2017, 2013. https://gsmaintelligence.com/analysis/2013/11/global-lte-network-forecasts-and-assumptions-201317/408/.

[5] VoLTE: Changing the Conversation, 2013. http://www.telecoms.com/160552/changing-the-conversation-2/.

[6] 4G America:3G/4G Deployment Status, July 2014.

[7] 3GPP. TS25.331: Radio Resource Control (RRC), 2006.

[8] 3GPP. TS23.272: CSFB in EPS, 2012.

[9] 3GPP. TS24.008: Core Network Protocols, 2012.

[10] 3GPP. TS44.018: obile radio interface layer 3 specification; Radio Resource Control (RRC) protocol, Nov. 2012.

[11] 3GPP. TS24.301: Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3, Jun. 2013.

[12] G. America. Year-end 2014: Nearly half a billion lte connections worldwide, March 2015. http://www.4gamericas.org/en/newsroom/press-releases/year-end-2014-nearly-half-billion-lte-connections-worldwide/.

[13] J. Bautista, S. Sawhney, M. Shukair, I. Singh, and etc. Performance of CS Fallback from LTE to UMTS. *IEEE Communications Magazine*, 51:136–143, 2013.

[14] Cisco Visual Networking Index. Global Mobile Data Traffic Forecast Update, 2013–2018, 2014.

[15] Y. Go, J. Won, D. F. Kune, E. Jeong, Y. Kim, and K. Park. Gaining Control of Cellular Traffic Accounting by Spurious TCP Retransmission. In *NDSS*, 2014.

[16] T. Koshimizu, I. Tanaka, and K. Nishida. Improvement on the VoLTE (Voice over LTE) Domain Handover with Operator's Vision. In *IEEE WTC*, 2012.

[17] V. Lee. Calls Blacklist - Call Blocker. available at Google Play.

[18] E. R. M. R. Myrto Arapinis, Loretta Ilaria Mancini. Privacy through pseudonymity in mobile telephony systems. In *NDSS*, 2014.

[19] O. Ozturk and M. Vajapeyam. Performance of VoLTE and data traffic in LTE heterogeneous networks. In *IEEE GLOBECOM*, 2013.

[20] C. Peng, C. Li, G. Tu, S. Lu, and L. Zhang. Mobile Data Charging: New Attacks and Countermeasures. In *CCS*, Oct. 2012.

[21] C. Peng, C.-Y. Li, H. Wang, G.-H. Tu, , and S. Lu. Real Threats to Your Data Bills: Security Loopholes and Defenses in Mobile Data Charging. In *ACM CCS*, 2014.

[22] F. Qian, Z. Wang, A. Gerber, Z. Mao, S. Sen, and O. Spatscheck. Characterizing Radio Resource Allocation for 3G Networks. In *IMC*, 2010.

[23] R. Racic, D. Ma, and H. Chen. Exploiting MMS Vulnerabilities to Stealthily Exhaust Mobile Phone's Battery. In *SecureComm'06*.

[24] RFC3261: SIP: Session Initiation Protocol, 2002. RFC 3261.

[25] F. Ricciato. Unwanted Traffic in 3G Networks. *SIGCOMM Comput. Commun. Rev.*, 36(2):53–56, Apr. 2006.

[26] P. Traynor, W. Enck, P. McDaniel, and T. La Porta. Mitigating Attacks on Open Functionality in SMS-capable Cellular Networks. In *MobiCom*, 2006.

[27] P. Traynor, P. McDaniel, and T. La Porta. On Attack Causality in Internet-Connected Cellular Networks. In *USENIX Security*, 2007.

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