Uncovering Insecure Designs of Cellular Emergency Services (911)

Yiwen Hu[†], Min-Yue Chen[†], Guan-Hua Tu[†], Chi-Yu Li[‡], Sihan Wang[†],
Jingwen Shi[†], Tian Xie[†], Li Xiao[†], Chunyi Peng[△], Zhaowei Tan[◇], Songwu Lu[◇*]
[†]Michigan State University, [‡]National Yang Ming Chiao Tung University,
[△]Purdue University,[◇] University of California, Los Angeles

ABSTRACT

Cellular networks that offer ubiquitous connectivity have been the major medium for delivering emergency services. In the U.S., mobile users can dial an emergency call with 911 for emergency uses in cellular networks, and the call can be forwarded to public safety answer points (PSAPs), which deal with emergency service requests. According to regulatory authority requirements for the cellular emergency services, anonymous user equipment (UE), which does not have a SIM (Subscriber Identity Module) card or a valid mobile subscription, is allowed to access them. Such support of emergency services for anonymous UEs requires different operations from conventional cellular services, and can therefore increase the attack surface of the cellular infrastructure. In this work, we are thus motivated to study the insecurity of the cellular emergency services and then discover four security vulnerabilities from them. Threateningly, they can be exploited to launch not only free data service attacks against cellular carriers, but also data DoS/overcharge and denial of cellular emergency service (DoCES) attacks against mobile users. All vulnerabilities and attacks have been validated experimentally as practical security issues in the networks of three major U.S. carriers. We finally propose and prototype standard-compliant remedies to mitigate the vulnerabilities.

CCS CONCEPTS

• **Security and privacy → Mobile and wireless security**.

KEYWORDS

Cellular Networks, Emergency Services, 911 (9-1-1), Security

1 INTRODUCTION

Emergency services are a vital lifeline to people in emergency conditions. The globally-deployed cellular networks with ubiquitous coverage have been the most accessible channel to emergency users. To ensure the availability for emergency uses, cellular standards and regulatory authorities have stipulated requirements for the offering of cellular emergency services. Specifically, from the GSM Association (GSMA) standard [23], emergency services must be supported by mobile phones without SIM (Subscriber Identity Module) cards, which are indicated as anonymous user equipments (UEs), and be free of charge for mobile users. The 3GPP standard [3] requires emergency services to be provided with higher priority than other services. In the U.S., Federal Communications Commission (FCC) [30] stipulates that cellular carriers have to deliver all wireless 911 calls to the public safety answering point (PSAP), which deals with emergency service requests, without respect to call validation results. Thus, cellular emergency services have become highly available and reliable for emergency uses.

The security research of emergency services has attracted much attention recently. Several attacks have been proposed to threaten emergency services, but they mainly focus on distributed denial-of-service (DDoS) attacks [17, 29, 38] against PSAPs (e.g., 911 call centers) rather than the cellular emergency services. Many solutions [19, 27, 31, 36, 37] have been thus introduced to address them. For the cellular emergency services, there have been also some proposed attacks [25, 26, 28] from the literature. Specifically, Lee *et al.* [28] and Hussain *et al.* [26] uncover that fabricated emergency alerts can be sent to victim UEs based on the abuse of cellular alert protocols and the hijacking of paging channels, respectively. Hou *et al.* [25] allow the adversary to not only bypass the victim UE's screen lock to dial any numbers on the emergency panel, but also block phone calls made to a set of numbers in a specific area, by providing the victim UE with a list of fake local emergency numbers via control-plane signaling messages.

The above attacks corresponding to the cellular emergency services mainly target the vulnerabilities on the UE side, but the security of the cellular infrastructure supporting emergency services still remains unexplored. Moreover, the cellular emergency services operate differently from conventional cellular services. Once any conventional designs are applied to the emergency services without careful reviews from a security perspective, security vulnerabilities may arise. Furthermore, allowing anonymous UEs to access the emergency services can increase attack surface of the cellular infrastructure. We are thus motivated to study whether the emergency services in the cellular infrastructure introduce any new security threats to mobile ecosystem or not.

Surprisingly, we discover four security vulnerabilities from the cellular emergency services in the cellular networks of three major U.S. carriers: unverifiable emergency IP-CAN session requests (V1), improper cross-layer security binding (V2), non-atomic cellular emergency service initialization (V3), and improper access control on emergency IP-CAN (IP Connectivity Access Network) sessions (V4). We then develop two proof-of-concept attacks based on them. The first attack is the denial of cellular emergency service (DoCES) developed based on V1 and V2; it allows the adversary to prevent mobile users from accessing cellular emergency services, and only two SDR (Software-defined Radio) platforms servicing as an attack UE and a sniffer are needed. This attack includes three

*The first two authors contribute equally to this work.

ACM MobiCom '22, October 17–21, 2022, Sydney, NSW, Australia

ACM ISBN 978-1-4503-9181-8/22/10…$15.00
https://doi.org/10.1145/3495243.3560534

Category	Attack	Victim	Description and Threat	Vulnerability	Attack Procedure	Attack Cost	Applicability	
							System	Carrier
Denial of Cellular Emergency Service (§4)	UE detaching	Individual	Adversary detaches the victim's emergency IP-CAN session, thereby preventing them from accessing all emergency services.	V1: Unverifiable emergency IP-CAN session requests (§4.1)	Using an SDR-based attack UE to send fabricated Attach Request/SIP Cancel/SIP Bye messages on behalf of victim UEs to the infrastructure while having an SDR-based sniffer to eavesdrop on nearby UEs' communication.	Two SDR cellular network platforms for serving as an attack UE and a sniffer.	4G	OP-I
	Call cancel	Individual	Adversary cancels the victim's emergency call attempt.	V2: Improper cross-layer security binding (§4.2)			4G, 5G†	OP-I*, OP-II*, OP-III*
	Call drop	Individual	Adversary terminates the victim's ongoing emergency call conversation with a PSAP.				4G, 5G†	OP-I*, OP-II*, OP-III*
Emergency IP-CAN Session Hijacking (§5)	Free Services	Operator	Adversary gains free data/voice/text services.	V3: Non-atomic cellular emergency service initialization (§5.1)	Using an SDR-based UE to serve as a Mobile-to-Internet gateway that provides UEs with free services via emergency IP-CAN session.	An SDR cellular network platform for serving as an attack UE.	4G, 5G†	OP-I, OP-II OP-III
	Data DoS/ overcharge	Individual	Adversary bypasses carriers' firewall protection and injects spams to impose denial of service or excessive data bill on the victim.	V4: Improper access control on emergency IP-CAN sessions (§5.2)	Using an SDR-based UE to communicate with other UEs via emergency IP-CAN session while bypassing carriers' internal firewall protection.		4G, 5G†	OP-II, OP-III
	Remote scanning	Individual	Adversary can remotely scan network services/applications available on the victim's device and launch remote attacks based on reported vulnerabilities.				4G, 5G†	OP-II, OP-III

†: Via empirical validation and/or 3GPP/GSMA standards study.
*: Validated via our testbed using emergency IP-CAN sessions established in tested carrier networks.

Table 1: A summary of the identified security threats of operational cellular emergency services.

variants, namely device detaching, call cancel, and call drop. The second attack developed based on V3 and V4 includes three variants, namely free data/voice/text service, data DoS/overcharge, and remote scanning. Table 1 summarizes the discovered vulnerabilities and attacks, which are experimentally confirmed in the three top-tier U.S. carriers. Notably, in this study, no emergency calls or texts are transmitted to real PSAPs due to ethical and illegal issues.

At the first glance, carriers should take the blame, since necessary security mechanisms are not deployed. However, after a careful analysis, we find that all identified vulnerabilities root in design defects of the cellular emergency standards, which span multiple protocols and network functions, so it is difficult for carriers to address them without significant effort. We further propose countermeasures including long-term security designs, which can address the vulnerabilities completely based on their root causes, and standard-compliant short-term remedies, which mitigate the vulnerabilities to reduce attack incentives.

This paper makes three key contributions: (1) we identify four vulnerabilities from cellular standard designs regarding emergency services, as well as validate them experimentally and analyze root causes; (2) we devise two proof-of-concept attacks with three variants each by exploiting the identified vulnerabilities and assess their real-world impact with three major U.S. cellular carriers; (3) we propose a suite of standard-compliant solutions and evaluate them based on a prototype. The lessons learned can secure both cellular network carriers and mobile users.

2 CELLULAR EMERGENCY SERVICE PRIMER

Network architecture. Figure 1 depicts a 4G/5G network architecture supporting cellular emergency services. The emergency service requests (calls or texts) are initiated by the UE with or without a valid SIM card and finally routed to PSAPs, which are connected to the cellular network through the Internet (IP) or the public switched telephone network (PSTN). Within the cellular network, an emergency service request from the UE in turn traverses radio access network (RAN), core network, and IP Multimedia Subsystem (IMS). Notably, 5G and 4G use distinct network entities for similar network functions; for example, the RAN uses base stations (BSs) to offer radio access; the BS is referred to as gNodeB in 5G

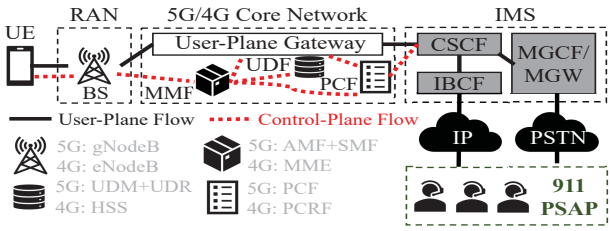


Figure 1: 5G/4G emergency service architecture.

and eNodeB in 4G. For simplicity, we intentionally avoid 5G/4G telecom jargons which are shown at the left bottom of Figure 1, but use generic names of network entities throughout this paper.

In the core network, the user-plane gateway (UPG) in the user plane is to route user traffic packets from the UE to the IMS network and eventually to the external network (e.g., PSAPs); it provides the emergency IP connectivity for emergency services with the functionality of UE IP address assignment and IMS server selection. In the control plane, there are three main control functions: (1) Mobility Management Function (MMF) manages radio access, user mobility, authentication, resource reservation, and emergency IP connectivity establishment; (2) User Data Function (UDF) is responsible for storing user and service subscription information; (3) Policy Control Function (PCF) is in charge of generating billing policies, QoS parameters, routing control rules and so on. The PCF also creates policies for the emergency IP connectivity and provisions them to the UPG or the MMF to assist in the control for voice and text emergency services.

The IMS provides emergency voice and text services over IP for UEs. It consists of three key network entities: Call Session Control Function (CSCF, referred to as IMS server hereafter), Media Gateway Control Function/Media Gateway (MGCF/MGW), and Interconnect Border Control Function (IBCF). The IMS server is responsible for IMS service signaling, which runs Session Initiation Protocol (SIP) [33]. The MGCF/MGW is connected to the traditional PSTN, whereas the IBCF is a session border controller which is interconnected to other IP/IMS networks.

IMS emergency service flow. Figure 2 illustrates a service flow for the cellular emergency voice/text service. To establish an emergency session with the PSAP, the emergency UE needs to perform

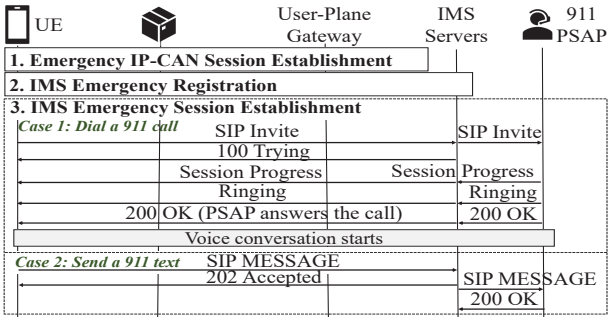


Figure 2: IMS emergency service flow.

the following three actions. First, *Emergency IP-CAN Session Establishment* allows the UE to obtain the emergency IP connectivity to communicate with the IMS server; an IP-CAN session is identified by the UE's IP address and identity information. Second, *IMS Emergency Registration* [3, 4] has the IMS server and the UE authenticate with each other and enables the UE to register the emergency service. Third, *IMS Emergency Session Establishment* allows an emergency UE to establish an IMS emergency call/text session with the PSAP [3, 4, 21, 22] through the IMS server. The UE sends SIP INVITE and SIP MESSAGE messages to the IMS server for establishing emergency call and text sessions, respectively. Notably, anonymous UEs may be still allowed to access the IMS emergency service without being registered in accordance with local regulatory requirements [8].

3 THREAT MODEL, METHODOLOGY, AND ETHICAL CONSIDERATION

Threat model. In this work, the adversary uses an SDR-based UE to attack operational cellular networks and cellular UEs in the two presented attacks; in particular, the attack SDR-based UE does not need to have any SIM card installed, but can successfully connect to operational cellular networks. In the first attack presented in Section 4.3, the victims are the cellular users who connect to operational emergency services using anonymous UEs. For the second attack in Section 5.3, the victims are cellular operators and non-emergency cellular users. In all the attacks, neither the operational cellular networks nor the victim UEs are compromised; the adversary is assumed to adhere to all cryptographic assumptions (e.g., a ciphered message cannot be decrypted without the ciphering key).

Experimental methodology. We validate the presented vulnerabilities and attacks in the operational cellular networks of three U.S. carriers, which are denoted as OP-I, OP-II, and OP-III. Two kinds of emergency UEs, which connect to operational cellular emergency services, are tested in the experiment: (1) commercial off-the-shelf (COTS) UEs, including Samsung Galaxy S8/S10, Google Pixel 3/5, and Apple iPhone 13; and (2) SDR-based UEs developed based on the srsRAN [35], which is an open-source 4G/5G software radio suite. Notably, all the vulnerabilities and attacks are validated in only 4G networks due to two major reasons. First, no COTS UEs which can be locked in the 5G network are found. Current COTS 5G UEs may switch to the legacy 3G network and make circuit-switched (CS) emergency calls; this fallback may cause emergency calls to reach PSAPs accidentally. Second, there are no SDR-based 5G UE platforms that can stably connect to operational 5G networks. However, it does not mean that our findings are limited to 4G networks only; more discussions about the applicability of the vulnerabilities and attacks in 5G networks are given in Section 7.

Ethical consideration. We understand that some feasibility tests and attack evaluations may be detrimental to cellular network carriers and users. We thus proceed with this preliminary study in a responsible manner. Specifically, two approaches are adopted. First, we use our own devices as the victims and purchase unlimited service plans in all the tested carrier networks. Second, all the vulnerability validation and attack experiments are conducted with small-scale tests based on the principle that aims to identify security issues of the cellular emergency services instead of aggravating damages. Notably, *in all the experiments, no emergency calls/text messages are sent to operational IMS servers or PSAPs.*

Responsible disclosure. We have reported all discovered vulnerabilities to tested carriers and provided them with standard-compliant remedies. Since those vulnerabilities may not be completely addressed at the publication of this paper, the names of those carriers are not disclosed.

4 DENIAL OF CELLULAR EMERGENCY SERVICE

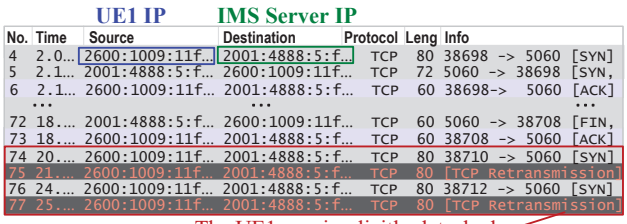
For emergency use, UEs shall be always allowed to make emergency calls/texts through a cellular network no matter whether they have valid subscriptions to the network, according to the FCC 911 requirements [30]. That is, for any U.S. cellular networks, anonymous UEs can access their cellular emergency services. The goal of this anonymous access is to maximize the availability of emergency services through cellular networks in emergency conditions. It can be also enabled for the UEs with valid subscriptions at the time when they are unable to access the emergency services from their home carrier networks; they are thus allowed to connect to other carrier networks and have the emergency services. However, we discover that such anonymous emergency service access is not well protected, thereby leading to a potential security threat, DoCES. It has two vulnerabilities: unverifiable emergency IP-CAN session requests (V1) and improper cross-layer security binding (V2).

In the following, we first introduce each vulnerability and then present the DoCES attack with several variants.

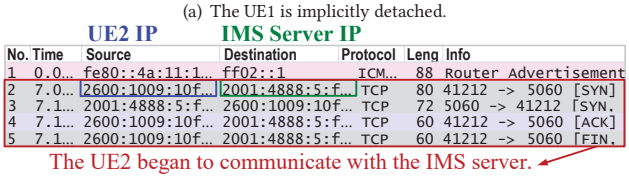
4.1 V1: Unverifiable Emergency IP-CAN Session Requests

Since an anonymous UE that attempts to consume the emergency service of a cellular network does not have any security association with the network infrastructure, the establishment procedure of the emergency IP-CAN session cannot be protected and its initial request is naturally unverifiable. When a duplicate establishment request is maliciously presented to the network, the network cannot differentiate it from the initial request. The impact of that malicious duplicate request depends on how the network deals with multiple emergency IP-CAN session requests from the same anonymous UE.

Surprisingly, the 4G and 5G standards take different approaches to handle the duplicate request. The 4G standard (i.e., TS24.301 [9]) stipulates that the MMF shall either reject it with a reason that multiple PDN connections for a given APN are not allowed, or accept it while implicitly detaching the existing established emergency



(a) The UE1 is implicitly detached.



(b) The UE2 establishes an emergency IP-CAN session successfully.

Figure 3: UE2's duplicate request makes UE1's ongoing emergency IP-CAN session be detached from the OP-I network.

IP-CAN session (i.e., the infrastructure detaches the session without providing any notification to its owner UE.). On the other hand, the 5G standard (i.e., TS23.501 [5]) specifies that the duplicate request shall be always rejected.

As a result, the adversary may have a chance to prevent anonymous UEs from accessing the emergency services by sending fabricated emergency requests to the network before or after valid requests. Since the requests are not ciphered or integrity-protected, they can be easily fabricated based on the same device ID.

4.1.1 Experimental validation. We validate this vulnerability using two SDR-based UEs: UE1 and UE2; neither of them has a SIM card installed. At the beginning, UE1 performs the establishment procedure of an emergency IP-CAN session with a tested 4G cellular network. Afterwards, UE2 sends the same cellular network a duplicate establishment request with the UE1's device identity, i.e., International Mobile Equipment Identity (IMEI). Once the UE1's emergency IP-CAN session is interrupted by the duplicate request, UE1 can be implicitly detached and then lose the IP connectivity. To detect whether this implicit detachment indeed happens, we make UE1 keep attempting to establish a new TCP connection with the assigned IMS server; the failure of any TCP connection establishment can indicate the loss of the IP connectivity.

We conduct this experiment with all the three carriers. The result shows that the UE2's duplicate request can successfully interrupt the ongoing emergency IP-CAN session of the UE1 in the OP-I network, but it does not work in the networks of OP-II and OP-III. As shown in Figure 3(a), the TCP connections cannot be established over the emergency IP-CAN session of the UE1 due to the implicit detachment caused by the UE2's duplicate session request; afterwards, the UE2 can communicate with the IMS server over the newly established IP-CAN session, as shown in Figure 3(b).

4.1.2 Root cause and lessons. The emergency IP-CAN session requests from anonymous UEs are unverifiable, since they do not have any security context shared with the cellular networks. However, allowing anonymous UEs to have the emergency services cannot be simply prohibited, since it is critical for emergency conditions. Moreover, duplicate emergency session requests cannot be simply

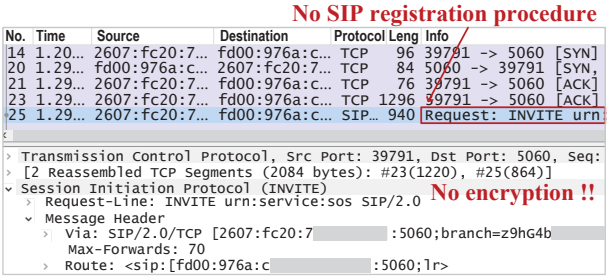


Figure 4: An unencrypted emergent call message is observed for a COTS phone without any SIMs in the OP-III network.

forbidden either, because they may be sent by benign anonymous UEs after a system or software crash. It thus calls for a new security mechanism that cannot only secure the cellular network with offered emergency services but also keep the high availability of the emergency services to anonymous UEs.

4.2 V2: Improper Cross-layer Security Binding

The UE with a valid mobile subscription cannot establish IPSec security associations with the IMS server for the emergency services until it completes the IMS emergency registration [14], since the IPSec ciphering and integrity keys are derived from the registration procedure. It appears that the network-layer security (i.e., IPSec) is bound to the application-layer security (i.e., SIP registration). Therefore, when anonymous UEs are allowed to skip the IMS registration due to no security context shared with the core network, the IPSec security associations with the IMS server cannot be built. It can leave the IMS emergency sessions of anonymous UEs to be unprotected; thus, the sessions may suffer from attacks.

4.2.1 Experimental validation. We validate this vulnerability by observing whether anonymous UEs indeed have unprotected IMS emergency call sessions. In the experiment, COTS UEs and operational cellular networks are considered. In order to prevent any emergency call signaling messages from being routed to PSAPs, we develop a smartphone application, namely 911-CallBlocker, which discards all the SIP INVITE messages sent from the smartphone to the network infrastructure. After activating the 911-CallBlocker at the tested smartphone without any SIM card (i.e., anonymous UE), we dial 911 while using TCPDump to record all the packets.

This experiment is conducted for all the three carriers. Figure 4 shows a representative trace from an anonymous UE connecting to the emergency services of the OP-III network. For all the carriers, we make two observations. First, the IMS emergency registration procedure is not performed. Second, the SIP INVITE messages are all sent in plain-text without ciphering protection. Thus, the critical session information (e.g., call-ID and call tag) can be leaked to the adversary; it can thus allow the adversary to manipulate ongoing emergency call sessions.

4.2.2 Root cause and lessons. The current cross-layer security design that binds the IPSec security association establishment to the IMS registration does not come without any reasons. It is necessary for non-emergency UEs to do the IMS registration; when the registration fails, no IMS services are provided to the UEs. That is, the

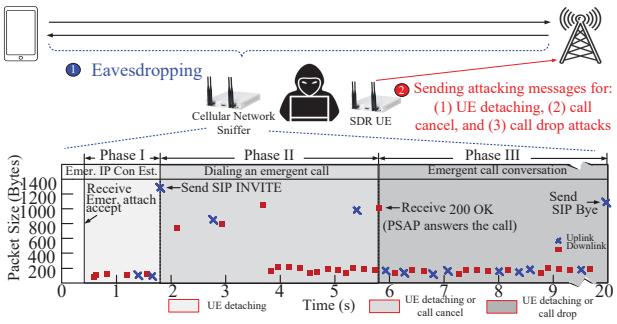


Figure 5: DoCES attack.

IPSec is needed only when the registration succeeds; the cross-layer security binding is thus reasonable and can work properly.

However, this security binding should not be directly applied to the cellular emergency services without any modifications. Anonymous emergency UEs can skip the IMS registration but are still allowed to establish IMS emergency sessions. Without the registration, the improper security binding causes the IPSec security association establishment to be skipped. Such design is explicitly stipulated in the 3GPP/GSMA emergency service standards [3, 23], so it can happen in all standard-compliant mobile devices. As a result, it calls for a security mechanism that is decoupled from the IMS registration and can protect the emergency service sessions.

4.3 Proof-of-concept Attacks

We exploit the above two vulnerabilities to launch the DoCES attack against anonymous UEs. This attack contains three attack variants that together can almost cover the entire life cycle of an emergency call, as illustrated in Figure 5; specifically, they are *UE detaching*, *call cancel*, and *call drop* attacks. Launching this attack requires two device components: (1) a cellular network sniffer, which eavesdrops on the communication of nearby UEs and identifies attackable UEs (i.e., anonymous UEs initiating cellular emergency services), and (2) an SDR-based UE, which sends attack messages to the cellular networks where victim UEs are. The cost of this attack is to have two SDR platforms compatible with 4G/5G networks for serving as the sniffer and the attack UE. Notably, this attack does not require the adversary to deploy rouge cellular infrastructure or install cellular signal jammers near victims. Moreover, the adversary does not need to be at the scene of victims; instead, the sniffer, together with the attack UE, can be deployed at any location where the victims' communication can be eavesdropped on.

We next present the experimental setting and then elaborate on each attack variant. Note that the following evaluation results demonstrate that adversaries could prevent mobile users from accessing emergency services in certain settings, but these should not be interpreted as common failures of operational cellular systems.

4.3.1 Experimental setting. We evaluate the DoCES attack with three variants on an emulation testbed deployed over the networks of the three carriers. Using the emulation testbed is to prevent any emergency calls from being sent to PSAPs. Figure 6 shows the testbed with three major parts, namely the emergency service system, the attack system, and the victim UE. The emergency service system includes an IMS server developed based on the open-source LinPhone VoIP SIP server [18] and an emulated IP-based PSAP;

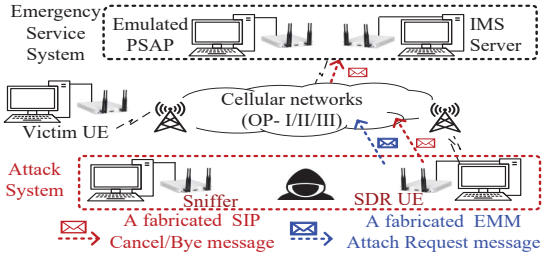
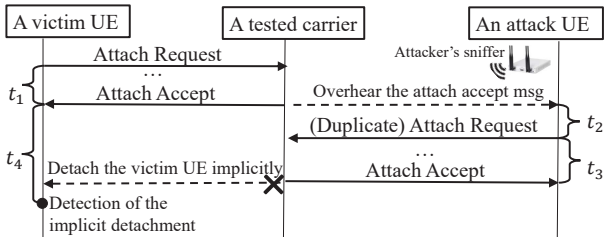
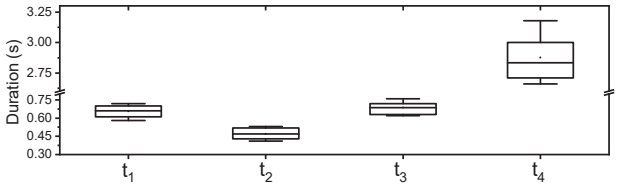


Figure 6: An emulation testbed for DoCES attack evaluation.



(a) Attack procedure.



(b) The time durations shown in the above figure (0/25/50/75/100th percentiles).

Figure 7: UE detaching attack.

both of these two components are emulated using SDR-based UEs connecting to the tested cellular network via emergency IP-CAN sessions. Thus, all the SIP messages generated by the victim UE are sent to the emulated PSAP rather than actual PSAPs. The attack system consists of a cellular network sniffer and an SDR-based UE with the LinPhone VoIP SIP client installed; the UE also connects to the tested cellular network with an emergency IP-CAN session. The victim UE is built based on the same SDR-based UE as the one in the attack system. Notably, only the IMS-related activities are emulated, but the underlying communications are still based on the emergency IP-CAN sessions established between the SDR-based UEs and the operational cellular networks.

4.3.2 UE detaching attack. We first devise the UE detaching attack that implicitly detaches emergency UEs based on the vulnerability V1. To exploit the V1, the attacker needs to identify potential victim UEs which are establishing emergency IP-CAN sessions, and obtain their device identities. To this end, a cellular network sniffer can be deployed to monitor particular control-plane signaling messages including EMM Attach Accept and EMM Attach Request [9] from nearby cellular UEs.

Figure 7(a) illustrates the procedure of this attack. While a victim UE nearby the sniffer performs the EMM Attach procedure [9] to establish an emergency IP-CAN session with a cellular network, the sniffer in the attack system can overhear the EMM Attach Accept message, which indicates the finish of the session establishment, from the cellular network. Afterwards, the SDR-based attack UE can fabricate a duplicate Attach Request message using the UE's

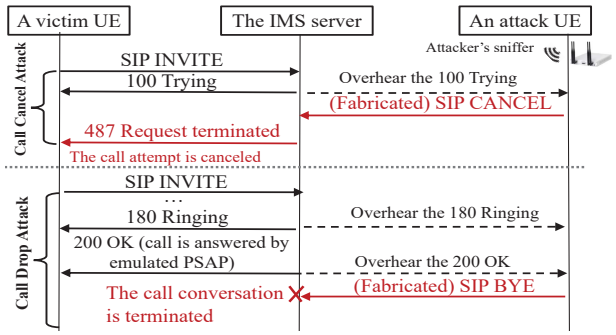


Figure 8: Message flows of call cancel and call drop attacks.

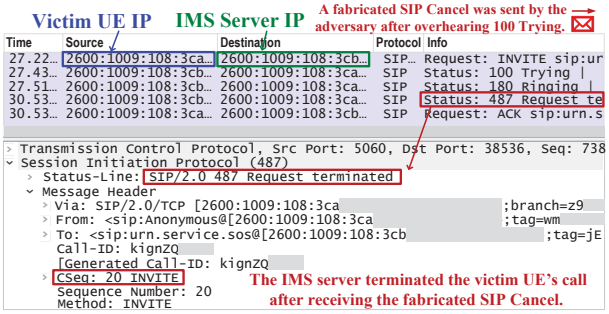


Figure 9: An emergency call is terminated by a fabricated SIP CANCEL message sent by the adversary.

IMEI. Once the attack succeeds, the network implicitly detaches the victim UE while replying Attach Accept to the attack UE.

We evaluate this attack by conducting the attack procedure for 10 runs in the OP-I network, where V1 exists. The evaluation result shows that the victim UE can be implicitly detached in all the experiment runs; that is, it does not receive any notification from the network after being detached. Figure 7(b) shows the measured values of the time durations in the attack procedure. It is observed that the attacker can successfully detach the victim UE within 2.66~3.18 s (i.e., t4) right after the emergency session is established. Note that getting the IMEI requires capturing the EMM Attach Request message from the uplink traffic. Although our open-source cellular radio sniffer can be used for only downlink traffic, it can be done by other commercial sniffers (e.g., WaveJudge 5000 LTE Analyzer). The victim UE's IMEI is thus given in the experiment.

4.3.3 Call cancel attack. We then devise the call cancel attack that cancels the victim UE's emergency call attempt based on the vulnerability V2, which allows the attacker to overhear and fabricate SIP messages. As shown in the upper part of Figure 8, the attacker can send a fabricated SIP Cancel message to the IMS server on behalf of the victim UE after overhearing the SIP 100 Trying message. Once the IMS server accepts the fabricated message, it will cancel the victim UE's call attempt by replying the Request Terminated. Notably, to fabricate a valid SIP Cancel message, the adversary can obtain required session information including Call-ID, tag@From, and branch@Via [33], from the SIP 100 Trying message.

The experimental setting of the attack evaluation is built as follows. Each of the victim UE, the emulated PSAP, the emulated IMS server and the adversary's SDR-based UE obtains an emergency IP-CAN session from the tested cellular network. Both the victim UE and the emulated PSAP are registered to the emulated IMS.

In the evaluation, the victim UE initiates a SIP call to the emulated PSAP; meanwhile, the attack UE launches the call cancel attack. The result shows that the victim UE receives a 487 Request terminated message from the IMS server, which indicates the victim UE's emergency call is successfully canceled. Figure 9 shows a representative trace of this successful attack result in the OP-I network; the same results are observed in all the three carriers.

4.3.4 Call drop attack. Similar to the call cancel attack based on V2, this attack is launched by sending a fabricated SIP message to the IMS server, but it has two major differences. First, it can terminate an ongoing emergency call conversation between the victim and the PSAP. Second, the fabricated SIP message is the SIP Bye, which requires an additional piece of SIP session information, tag@To, compared with the SIP Cancel; it can be collected from the SIP 180 Ringing. As shown in the lower part of Figure 8, after overhearing the SIP 200 OK message, the attacker can send a fabricated SIP Bye message to the IMS server on behalf of the victim UE. Once the IMS server accepts the fabricated message, the victim UE's ongoing emergency call will be terminated. The experiment setting of this attack evaluation is the same as the previous one, besides that the emulated PSAP answers the victim's call. The result shows that the the victim UE does not receive any messages from the IMS server but the voice conversation is terminated.

5 EMERGENCY IP-CAN SESSION HIJACKING

The emergency IP-CAN session is established whenever a cellular emergency service is requested. Particularly, the emergency service request can be issued from anonymous UEs and be free of charge for cellular users due to its emergency purpose [3, 9, 10, 23]. It can be thus more vulnerable than other non-emergency services. However, we discover that no additional security mechanisms are introduced to protect the emergency IP-CAN session; thus, it could be arbitrarily established and then hijacked to launch a variety of attacks, e.g., free data/voice/text service and DoS attacks.

In the following, we first identify two vulnerabilities, non-atomic cellular emergency service initialization (V3) and improper access control on emergency IP-CAN sessions (V4), and then propose three proof-of-concept attacks.

5.1 V3: Non-atomic Cellular Emergency Service Initialization

The cellular emergency service initialization is triggered right after a user submits an emergency call/text request on the UE. It consists of three actions, as described in Section 2. For the timely delivery of an emergency service request, the initialization is expected to have the atomic property where those three steps are executed continuously without being decoupled or being interleaved with other UE actions. Specifically, the UE can only do IMS emergency registration or/and establish an emergency session with a PSAP whenever an emergency IP-CAN session, which is built for the exclusive use, is established. After the initialization, the emergency service request can reach the PSAP.

However, the cellular network infrastructure may not fulfill this property, since no related security mechanisms are stipulated in the 3GPP/GSMA standards [3, 9, 10, 23]. It may allow an adversary

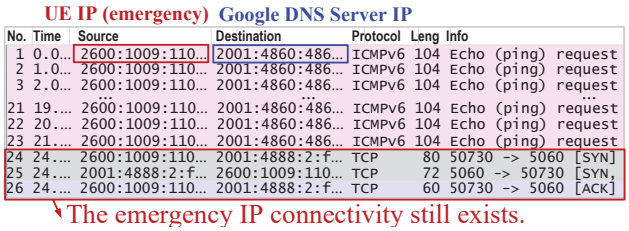


Figure 10: The UE can keep the emergency IP-CAN session active by periodically sending packets out.

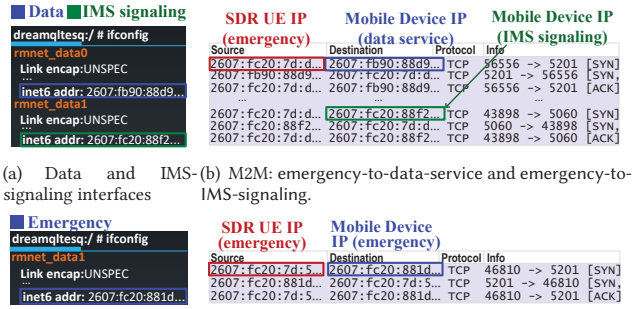
to establish an emergency IP-CAN session to abuse while skipping the last two initialization actions. Without the IMS emergency registration or/and session establishment, the IMS server and the PSAP cannot be aware of the abuse. More threateningly, the emergency IP connectivity can be requested by anonymous UEs, so it is challenging to trace back to the adversary.

5.1.1 Experimental validation. We validate this vulnerability by developing an SDR-based UE using the srsRAN [35]. The UE without any SIM card installed is made to perform the emergency IP-CAN session establishment with three 4G carriers, but skip the last two initialization actions and transmit no packets to the infrastructure.

We have two findings. First, the anonymous UE can successfully obtain an IP address for the established emergency IP connectivity from each carrier. Second, the emergency IP connectivity can be interrupted by the infrastructure (i.e., the UE is implicitly detached), after a period of time, which is 10 s, 5 s, and 3 s for OP-I, OP-II, and OP-III, respectively. It can be thus inferred that an inactivity timer is deployed to protect the emergency IP connectivity from being abused. Nevertheless, we discover that the UE can prevent the interruption by sending packets out periodically; moreover, the destination is not necessarily to be the IMS server. As shown in Figure 10, the UE can keep the emergency IP connectivity active by sending ICMP packets to the Google DNS server; notably, no ICMP response packets are received by the UE, but the major purpose that the emergency IP connectivity appears to be in use with those outgoing packets has been achieved. In sum, *an adversary can obtain the emergency IP connectivity and keep it active for a long time.*

5.1.2 Root cause and lessons. This vulnerability can be attributed to a design defect that the cellular infrastructure does not enforce the atomicity of the cellular emergency service initialization. This design defect appears when the emergency service migrates from the 2G/3G circuit-switched (CS) system to the 4G/5G packet-switched (PS) one without a careful security review. In the CS system, the emergency service initialization is completely taken charge of by a single network entity, MSC (Mobile Switch Center [1]), so the atomicity can be easily ensured by the MSC.

However, the emergency service becomes to be IMS-based in the PS system and the initialization is decomposed into two parts, the emergency IP-CAN session establishment and the IMS emergency registration/session establishment, which are managed by the MMF and the IMS server, respectively. Without an additional security mechanism stipulated to protect the emergency service initialization among them, they do not cooperate to ensure the atomicity. Specifically, the MMF can know which UEs obtain the emergency IP connectivity, but have no information about whether those UEs continue to proceed with the IMS emergency service operation;



(a) Data and IMS-signaling interfaces

(b) M2M: emergency-to-data-service and emergency-to-IMS-signaling.

(c) Emergency-service interface.

(d) M2M: emergency-to-emergency

Figure 11: An SDR-based UE uses the emergency IP-CAN session to communicate with another UE in OP-III.

on the other hand, the IMS server does not know which UEs have gained the emergency IP connectivity. Thus, it calls for a concerted solution to ensure the atomicity.

5.2 V4: Improper Access Control on Emergency IP-CAN Sessions

The access control on emergency IP-CAN sessions is fulfilled by the PCF to provision PCC (Policy and Charging Control) rules for MMFs or UPGs [7, 12]. For an IP-CAN session, each PCC rule identifies a set of service flows based on the 5-tuple information (i.e., source/destination IP addresses, source/data port numbers, and transport protocol ID) and the corresponding service flows are managed based on an associated policy control setting, including precedence, QoS parameters (e.g., maximum uplink/downlink throughput), gate status (allowed or disallowed), etc. Thus, for the exclusive use of the emergency service, the emergency IP-CAN session should be restricted to deliver traffic to the IMS server based on given PCC rules. However, the cellular network standards [7, 12] do not stipulate such a regulation or give the PCF the information of the IMS server assigned to emergency UEs during their emergency IP-CAN session establishment, so the restriction may be ignored. Without the access control, adversaries may abuse emergency IP-CAN sessions to access the Internet or other cellular devices.

5.2.1 Experimental validation. We conduct an experiment to examine whether the emergency IP-CAN session is restricted to only service flows between the UE and the IMS server. Two types of service flows which do not reach the IMS server are tested for those three U.S. 4G carriers: mobile-to-Internet (M2I) and mobile-to-mobile (M2M), which represent the communication between the UE using the emergency IP-CAN session and Internet hosts, and the communication between that emergency UE and another tested UE, respectively. For the M2M case, we further test three kinds of IP-CAN sessions that may be used by the tested UE: (1) the data-service IP-CAN for Internet access, (2) the IP-CAN of the IMS call signaling, and (3) the emergency IP-CAN. Notably, the UE creates a network interface for each IP-CAN session; as shown in Figure 11(a), the interfaces of the data-service and IMS-signaling IP-CAN sessions can be observed, whereas Figure 11(c) shows the interface of the emergency IP-CAN session.

Carriers	Mobile-to-Internet	Mobile-to-Mobile		
		E2E	E2IMS	E2D
OP-I	X	O	X	X
OP-II	X	O	X	O
OP-III	X	O	O	O

Table 2: The available communication cases based on the emergency IP-CAN session vary with carriers.

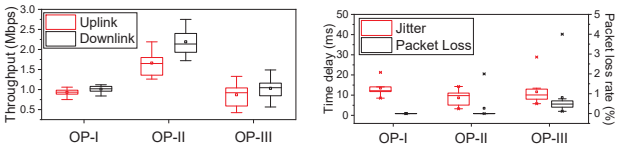
In this experiment, we still use the SDR-based UE without SIM card to obtain an emergency IP-CAN session from each tested carrier network. For the M2I case, the UE is tested to communicate with the Google DNS server using the emergency IP-CAN. In the M2M case, two phones are connected to the tested carrier network; one phone with a valid SIM card can obtain two IP-CAN sessions for *data service* and *IMS signaling*, respectively, whereas the other phone without SIM card can obtain an emergency IP-CAN session. Four phone models, including Samsung Galaxy S8/S10 and Google Pixel 3/5, are tested. The SDR-based UE is tested to communicate with those two phones through each of those three different IP-CAN sessions based on their corresponding IP addresses. The tested communication is based on the ICMP echo request/reply and the TCP three-way handshake.

Table 2 summarizes the result for all the three tested carriers. We have two observations. First, the M2I communication based on the emergency IP-CAN is forbidden for all the tested carriers. Second, all the carriers allow the emergency IP-CAN to have the M2M communication, but the allowable cases vary with the carriers. Specifically, the OP-III allows the communication for all the three different cases, as shown in Figure 11, whereas OP-I permits only the emergency-to-emergency (E2E) communication, and two communication types, E2E and emergency-to-data-service (E2D), are available for OP-II. In sum, all the tested carriers have improper access control on the emergency IP-CAN session.

5.2.2 Root cause and lessons. The root cause of this vulnerability is a lack of an access control mechanism on the emergency IP-CAN session in the standards, so it can be attributed to a design defect. At the first glance, designing the access control mechanism is straightforward, since the only requirement is to install the PCC rules that can restrict the emergency IP-CAN to the IMS server only. Specifically, during the emergency IP-CAN session establishment, the MMF or the UPG should provide the PCF with the IMS server information and then the PCF produces the corresponding PCC rules for the installation.

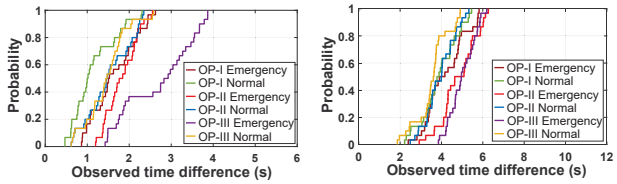
However, the real situation is much more complex; the IMS server may not be always determined during the emergency IP-CAN session establishment. The IMS server can be also assigned based on the DNS (Domain Name Service) or DHCP (Dynamic Host Configuration Protocol) services after the UE obtains the emergency IP-CAN [4]. In this case, the PCC rules cannot be produced and installed until the IMS emergency registration proceeds; during the registration, the IMS server needs to notify the PCF after receiving the UE's SIP Register message [2]. But, the adversary is allowed to skip the registration and bypass this notification. Thus, installing the PCC rules for the access control should be designed to be independent of the emergency registration.

Figure 12: Exploiting the E2E communication to enable free data service using a Mobile-to-Internet gateway.



(a) Throughput. (b) Packet loss rate and jitter.

Figure 13: The min/med/max and 25th/75th percentiles of throughput, jitter, and packet loss rate for the data service over the emergency communication channel.



(a) Call setup time (b) Text delivery time

Figure 14: The CDF results of the call setup and text delivery times observed in the emergency attacks and normal cases.

5.3 Proof-of-concept Attacks

We devise three proof-of-concept attacks, namely free data/voice/text services, data DoS/overcharge, and remote scanning, using the vulnerabilities V3 and V4. The cost of these attacks is to have an SDR platform compatible with 4G/5G networks; it serves as a M2I gateway that provides the free services over an emergency IP-CAN session, and an attack UE for the first and last two attacks, respectively. We next elaborate on the details of each attack.

5.3.1 Free data/voice/text service attack. The adversary can exploit the E2E communication, the delivered data of which are free of charge, to obtain free data/voice/text service. To achieve it, an M2I gateway needs to be deployed to forward data between the UE with an emergency IP-CAN session and the Internet, as shown in Figure 12. At the gateway, the SDR UE connects to the cellular infrastructure using an emergency IP-CAN session and receives/transmits all data to/from the other UEs through the free E2E communication, the Wi-Fi router connects to the Internet, and the computer forwards data between the SDR UE and the router.

We next evaluate the data service over that free-of-charge communication channel in all the three carrier networks. We use IPerf to assess its throughput, jitter, and packet loss rate with 20 runs each. As shown in Figure 13, the median values of the uplink and downlink throughput range from 0.83 Mbps to 2.17 Mbps, all the jitter values are smaller than 30 ms, and all the packet loss rates are smaller than 1%. Note that the measured throughput is constrained by the SDR-based UE, which supports only a single antenna [15] with the current srsRAN version (20.10), so the adversary may increase the throughput using more advanced UEs in this attack.

We further use Google Voice over the free-of-charge channel to have voice and text services at no cost [39]. We assess the voice and

text services by considering the call setup time and the text delivery time, respectively. Figure 14 plots the CDF results by comparing the attack with normal cases, where the UE with a valid mobile service subscription uses the Google voice. It is seen that this attack can offer comparable performance to normal cases. Specifically, they have the ranges of the call setup time, 0.86s~3.87s and 0.47s~2.58s, respectively, whereas those of the text delivery time are 2.39s~6.27s and 1.87s~5.46s, respectively.

5.3.2 Data DoS/overcharge attack. The adversary can further use the E2D communication to launch a data DoS/overcharge attack against cellular users. The spamming data can be generated from the attack UE's emergency interface at no cost and sent to a victim UE's data interface, thereby consuming the data quota of the victim's data service plan. It can cause the victim UE to suffer from an overcharged bill or the data DoS, where its subscribed data quota is exhausted. In particular, massive cellular IoT devices (e.g., water and electricity meters) are more vulnerable to this attack, since they usually have only a small amount of data quota with high unit rates (e.g., $0.99 per MB) in common IoT service plans. The prerequisite of this attack is to obtain the IP addresses of potential victim UEs. To target cellular IoT devices in this attack, the adversary can remotely identify their IP addresses by probing them based on the operation of the cellular IoT power saving mechanism (PSM) [40]. The adversary can also attack specific UEs and steal the information of their IP addresses by installing the malware or launching phishing attacks.

We validate the feasibility of this DoS/overcharge attack for both OP-II and OP-III using four different victim UEs, including Samsung Galaxy S8 and S10, Google Pixel 3 and 5. Each validation test consists of the following three steps. First, we obtain the latest data usage amount three days after powering off the victim UE. Second, after powering on the victim UE, we use the attack UE to send spamming data from its emergency interface to the victim UE's data interface. The spamming packets are the UDP datagrams created by the attack UE using a randomly selected UDP destination port number and the victim UE's IP address. The victim UE may reply ICMP Port Unreachable error message to the attack UE. Third, we power off the victim UE and keep it for three days; afterwards, we query the latest data usage amount again.

We show the evaluation result of OP-III only, since the attack becomes unavailable for OP-II during the evaluation experiment[1]. In the experiment, we vary spamming rates from 50 Kbps to 400 Kbps and for each test, the spamming attack lasts for 30 s. Figure 15 shows the volume of spamming data which are sent, received, and charged in the OP-III network. It can be seen that the victim is charged for all the spamming data.

5.3.3 Remote scanning attack. The E2D communication also allows the adversary to scan victim UEs remotely for vulnerability discovery while bypassing cellular network firewalls. Specifically, the adversary can send probing packets (e.g., TCP SYN) to various

[1]This attack was successfully validated for OP-II in August 2021, but it became unavailable later in December 2021 when a comprehensive attack experiment was conducted. The observed difference between these two experiment times is that the IP addresses assigned to non-emergency IP-CAN sessions change from IPv6-based to IPv4-based, whereas those of emergency IP-CAN sessions are still IPv6-based. Such changes in the network configuration/infrastructure could be the reason why the E2D communication over V4 becomes unavailable in the OP-II network.

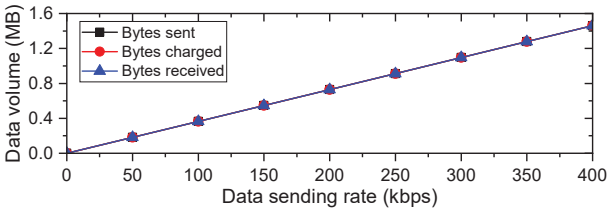


Figure 15: The volume of spamming data which are sent, received, and charged from the data DoS/overcharge attack against a victim device in the OP-III network.

No.	Service ID	Service Application	Protocol	Port	Reported CVE
1	saphostctrl	AirDrop - File Share For Android	TCP	1128	CVE-2019-9832
2	lm-x	Opera Mini Browser		6200	CVE-2021-23253
3	ultraseek-http	AirDroid - File Transfer&Share		8765	CVE-2019-9599
4	amcs	Sand Studio -Screen mirroring		8766	CVE-2015-5661
5	http	ES - File Explorer, File Manager		59777	CVE-2019-6447
6	upnp	UPnP Simple Service Discovery	UDP	1900	CVE-2021-27239
7	bfd-control	Bidirectional Forwarding Detection		3784	CVE-2021-28496
8	zeroconf	Multicast DNS (mDNS)		5353	CVE-2017-6519
9	oma-ulp	OMA User Plane Location		7275	CVE-2016-10416
10	unknown	Eques Smart Door Control		27341	CVE-2019-15745

Table 3: The result of the remote scanning attack against a Samsung S8 in the OP-III network; only the services and ports with reported CVE are listed.

port numbers of the victim UEs, and then determine which ports are open and which services are running at each victim UE based on the responses (e.g., TCP SYN+ACK or ICMP Port Unreachable) corresponding to the probing packets. The collected information of each UE is then used to query the CVE (Common Vulnerabilities and Exposures) database to examine whether the UE has any potential security vulnerabilities.

We validate this attack by using Nmap, which is an open-source utility for network discovery and security auditing, to send the probing packets from the attack UE's emergency interface to the victim UE's data interface. This validation test is conducted in OP-II and OP-III, both of which allow the E2D communication, with three victim UEs, including Samsung S8, Pixel 5, and iPhone 13. We discover that to scan 5,000 ports, the attack UE needs to send and receive around 322.8 KB and 306.1 KB, respectively, and it takes around 13 s. Table 3 summarizes the scanning result obtained from S8 in OP-III with a list of services and ports associated with reported CVE vulnerabilities.

6 COUNTERMEASURES

All the discovered vulnerabilities root in design defects of the cellular emergency services stipulated in the 3GPP/GSMA standards. However, addressing them based on their root causes to have a secure design may not be practical in the short term, since the required design changes lie in some core network functions and even security functions of billions of UEs. It cannot be achieved without significant effort or a long time. In the following, we first present long-term secure designs that can address the vulnerabilities, together with their expected overhead, and then introduce three short-term, yet low-overhead, remedies that can mitigate those vulnerabilities.

6.1 Long-term Security Designs

We present the design change required for each vulnerability below.
V1 (unverifiable emergency IP-CAN session requests). It calls for a device-level authentication mechanism, which can make differences on emergency IP-CAN session requests from different UEs,

even when the UEs do not have SIM cards. It requires each UE to have device credentials (e.g., certificates), but it is not easy to upgrade each UE to get and install a carrier-certified certificate since the process requires the device owner to be involved but not an automatic upgrade with a software patch due to security concerns.
V2 (improper cross-layer security binding). The cross-layer security binding between the establishment of IPSec security association and the IMS registration shall be decoupled. However, such design change could incur a large overhead, since the general IMS operation for both emergency and non-emergency services needs to be modified; specifically, the derivation of the IPSec security context needs to be removed from the IMS registration procedure.
V3 (non-atomic cellular emergency service initialization). The three steps in the cellular emergency service initialization need to be combined into an atomic operation. Specifically, the request of the emergency IP-CAN session establishment piggybacks the requests of both IMS emergency registration and session establishment procedures. Once this combined request arrives at the core network, the corresponding emergency call attempt can reach the IMS server so that the emergency IP-CAN session cannot be hijacked without raising awareness from the IMS. However, handling that combined request requires modifications on the MMF, the UPG, and the IMS server, which cannot be done in a short time.
V4 (improper access control on emergency IP-CAN sessions). The MMF or the UPG shall provide the PCF with the IP address of the IMS server assigned to each emergency UE so that the PCF can install a proper access control rule that can restrict the emergency IP-CAN session to the IMS server only. However, the assignment of the IMS server can be done through the DHCP or DNS service, after the establishment of the emergency IP-CAN session [4]; there could still exist a window period when the emergency IP-CAN session is not restricted and may be abused. Thus, the IMS server assignment shall be executed during the emergency IP-CAN session establishment. However, this proposed design can incur a large overhead due to the required support of multiple core network functions, e.g., MMF, UPG, PCF, and IMS server.

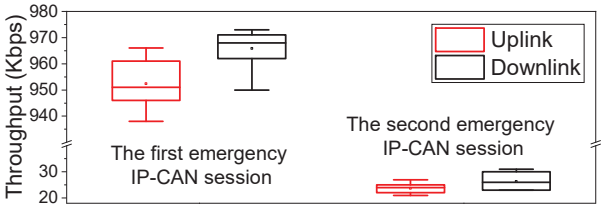
6.2 Short-term Remedies

In this section, we propose a suite of standard-compliant remedies, which can reduce attack incentives or mitigate attack damage, instead of fully addressing the vulnerabilities.
Restricted resource on duplicate emergency IP-CAN session. Simply rejecting each duplicate emergency session request is seemingly an effective solution to address V1, but the duplicate ones may be sent by benign UEs in some scenarios. For example, while a user is having an emergency call, the smartphone may be accidentally rebooted due to some unexpected software/hardware errors [16, 24]; this accidental event does not allow the smartphone to perform the detach procedure of the emergency IP-CAN session and the session is not released, so when the user dials an emergency call again after the smartphone reboots, a duplicate emergency session request can be generated. As a result, this simple-rejection method may hurt the availability of the emergency service for benign UEs. In order to not only defend against the DoCES attack but also keep the service availability, we propose to accept duplicate emergency session requests but restrict their session capability while keeping the existing emergency sessions that are duplicated.

Specifically, the duplicate emergency IP-CAN sessions are restricted to only the access of basic IMS emergency services (e.g., 31 Kbps for voice calls with the basic audio codec [11]), but not allowed to access video calls or voice calls with high audio quality codecs. Even though duplicate emergency sessions are established by the adversary, the resources available to be abused are limited, since these duplicate emergency sessions are granted only the minimum resource supporting the basic IMS emergency service; the attack incentive can be thus greatly reduced. On the other hand, when the duplicate ones are created by benign UEs, they are still available to offer the emergency services.
Enabling TLS protection over IMS emergency session. V2 can be addressed by enabling the ciphering and integrity protection over IMS emergency sessions. However, emergency UEs may not have credentials to do IMS emergency service registration and then establish IPSec security associations with their IMS servers. We then propose a standard-compliant method that an emergency UE establishes a TLS session with its IMS server using only the server's certificate prior to the IMS emergency service registration [14]. The TLS session can protect the IMS signaling messages with ciphering and integrity, thereby preventing fabricated SIP messages. Notably, this approach does not require significant support from carriers, since it was originally stipulated by the cellular network standards [14] to be used as an optional security mechanism to improve the security of IMS service access.
Delay authorization of emergency IP-CAN session. To address V3 and V4, we propose to delay authorization of each emergency IP-CAN session. The initial IP-CAN session obtained from the emergency IP-CAN session establishment for a UE is deemed as a temporarily-authorized session, the availability of which is only authorized for a short time period (e.g., 3 s); moreover, the bandwidth of this temporarily-authorized session is also limited to a small value (e.g., 31 Kbps). Its permanent authorization is delayed until the IMS server assigned to the UE receives SIP messages from the UE, and then determined by the IMS server. If no anomaly happens, the IMS server authorizes the session permanently by instructing the PCF to remove the session's time constraint and install proper PCC rules to restrict the IP-CAN session to the IMS server only. With this mechanism, even though the adversary may abuse the IP-CAN session during the initial, temporarily-authorized time period, their incentive can be largely decreased by that short abuse time. Notably, not all UPGs understand the IMS-related messages, so the permanent authorization of the emergency IP-CAN session cannot be done at the UPG during its establishment procedure.

6.3 Prototype and Evaluation

We prototype and evaluate the above three standard-compliant remedies. To emulate the cellular emergency service architecture, we use srsRAN (v20.1) [35], Open IMS Core [32], and LinPhone Voice client [18] to serve as the 4G LTE infrastructure, the IMS core with an IMS server, and the Voice over IMS app, respectively.
Restricted resource on duplicate emergency IP-CAN session. We upgrade srsRAN to support the emergency IP-CAN session establishment and modify the PCF to limit the maximum throughput of duplicate emergency IP-CAN sessions to 31 Kbps. In the experiment, we make the UE establish two emergency IP-CAN



(a) Restricted resource on duplicate emergency IP-CAN session.

LinPhone client OpenIMS server

No.	Source	Destination	Protocol	Length	Info
4	192.168.200.130	192.168.200.131	TLS...	585	Client Hello
6	192.168.200.131	192.168.200.130	TLS...	814	Server Hello, Certi
8	192.168.200.130	192.168.200.131	TLS...	194	Client Key Exchange
9	192.168.200.131	192.168.200.130	TLS...	310	New Session Ticket,
...	...	...	...	...	SIP Invite 100 Trying
61	192.168.200.130	192.168.200.131	TLS...	1447	Application Data
66	192.168.200.131	192.168.200.130	TLS...	396	Application Data

(b) TLS-protected IMS emergency session.

Figure 16: Short-term remedy evaluation.

Time	Protocol	Info
18.287...	S1AP/NAS...	InitialContextSetupRequest, Attach accept,
18.327...	S1AP	UECapabilityInfoIndication, UECapabilityIn
18.532...	S1AP	InitialContextSetupResponse
18.533...	S1AP/NAS...	UplinkNASTransport, Attach complete, Activ
18.533...	S1AP/NAS...	DownlinkNASTransport, EMM information
21.287...	S1AP	UEContextReleaseCommand
21.287...	S1AP	UEContextReleaseComplete

The UE was implicitly detached by the MME in about 3 seconds.

(a) MME implicitly detaches UE.

UE IP (emergency) IMS Server IP An invalid SIP Invite

Time	Source	Destination	Protocol	Info
5.4360...	172.16.0.2	172.16.0.1	ICMP	Echo (ping) request id=0x007a,
5.4722...	172.16.0.1	172.16.0.2	ICMP	Echo (ping) reply id=0x007a,
6.0713...	172.16.0.2	172.16.0.1	TCP	46483 -> 4070 [SYN] Seq=0 Win=
6.0921...	172.16.0.1	172.16.0.2	TCP	4070 -> 46483 [SYN, ACK] Seq=0
6.0921...	172.16.0.2	172.16.0.1	TCP	46483 -> 4070 [ACK] Seq=1 Ack=
6.6776...	172.16.0.2	172.16.0.1	SIP...	Request: INVITE tel:8881234567
8.4400...	172.16.0.2	172.16.0.1	ICMP	Echo (ping) request id=0x007a,
9.4719...	172.16.0.2	172.16.0.1	ICMP	Echo (ping) request id=0x007a,
10.491...	172.16.0.2	172.16.0.1	ICMP	Echo (ping) request id=0x007a,

UE was implicitly detached. There was no Echo reply being received.

(b) Dialing a non-emergency call.

Figure 17: UE is implicitly detached when no valid IMS emergency session is established within 3 seconds.

sessions, primary and secondary sessions, on the testbed and measure their throughput using Iperf. Figure 16(a) plots the throughput measured from 10 experiment runs. It is observed that the maximum throughput of the secondary emergency IP-CAN session is limited to 31 Kbps, whereas that of the primary one is as high as 973 Kbps. Together with the proposed delay authorization method, this remedy can largely decrease adversaries' incentives.

Enabling TLS protection over IMS emergency session. We enable the TLS support on the OpenIMS server and LinPhone Voice client. As illustrated in Figure 16(b), all the SIP messages of the emergency call establishment are protected by the established TLS session between the client and the server. It can thus prevent the DoCES attack, which relies on the SIP messages sent in plaintext.

Delay authorization of emergency IP-CAN session. We modify the PCF server to restrict the access of the emergency IP-CAN sessions with specified PCC rules at the UPG. For the delay authorization mechanism, a 3 s timer is set for each emergency IP-CAN session right after it is established. By default, after 3 s, it will be terminated by the UPG and its PCC rules will be removed; the Delete Bearer Request message [13] is sent to the MMF for the termination. For normal emergency service requests, the IMS server can receive a valid SIP INVITE message for the emergency IP-CAN session within that 3 s; then, it will authorize the emergency IP-CAN session by sending the AAR (Authentication Authorization Request) message [2] to the PCF through the standardized Rx interface [2].

We evaluate this remedy for the UE in three tested scenarios: (1) transmitting nothing to the infrastructure, (2) transmitting an invalid SIP INVITE message with a non-emergency phone number to the IMS server, and (3) transmitting a valid SIP INVITE message using urn:service:sos as the recipient's number to the IMS server. As shown in Figure 17, the UE will be implicitly detached by the infrastructure if no valid SIP INVITE message is received within 3 s after its emergency IP-CAN session is established. The result shows that the adversary cannot keep the emergency IP-CAN session being alive for a long time without a valid IMS emergency session.

7 DISCUSSION

Launching attacks from COTS UEs? Some attacks (e.g., data DoS/spamming/free attacks) can be launched from COTS UEs, but they need to be finished within a short time period, because the UEs can be switched to the legacy 3G network, where the attacks are not allowed, after they fail to communicate with the IMS emergency service server. We have developed a tool on the COTS UEs to intercept all the SIP messages and reply to some critical messages so that any emergency calls will not be sent to PSAPs after an emergency IP-CAN session is established. However, the tool can only delay the fallback switch without avoiding it. Notably, completely preventing the fallback requires to compromise the UEs' modems or finds COTS UEs supporting 4G/5G network services only.

Potential DoS attacks. The emergency IP-CAN sessions have higher priority than the non-emergency ones, so the adversary can exploit the vulnerability V1 (unverifiable requests) to establish multiple concurrent emergency sessions and generate as many packets as possible to exhaust a cell's limited radio resource, thereby causing a cell DoS attack where non-emergency UEs have no available bandwidth for network services. Due to the ethical reason, this DoS attack cannot be assessed in operational networks. We then evaluate it using our srsRAN testbed with SDR-based UEs and 4G LTE infrastructure. On the testbed, we set the maximum bandwidth value to 3 Mbps, which is the maximum uplink/downlink bandwidth observed in the tested carrier networks.

In the attack evaluation, we build multiple concurrent emergency IP-CAN sessions using ZeroMQ [41] and generate as much uplink traffic as possible from each session, while measuring the uplink/downlink throughput of a victim UE connecting to the same 4G network. We vary the number of concurrent emergency sessions from 0 to 4 and have 10 runs for each experimental setting. As shown in Figure 18, the uplink/downlink throughput values of the victim UE decrease with the increasing number of the emergency sessions; they reach 0 Mbps when there are 4 concurrent emergency sessions. This result confirms the feasibility of this attack.

However, the proposed short-term remedy, delay authorization of emergency IP-CAN session, can defend against this attack. It not only allows initial emergency sessions to be temporarily authorized for only a short time period (e.g., 3 s), but also limits the bandwidth of those temporarily-authorized sessions (e.g., 31 Kbps). Therefore, without thousands of emergency sessions being established within that short time period, a LTE cell with more than

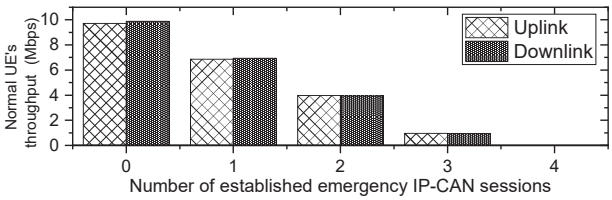


Figure 18: The median throughput values of the victim UE vary with number of emergency IP-CAN sessions in a cell.

100 Mbps bandwidth cannot be saturated; launching this DoS attack becomes almost impossible.

Applicability of vulnerabilities/attacks in 5G networks. We consider that the discovered vulnerabilities V2, V3, and V4, as well as their corresponding attacks, may still exist in 5G networks, according to an analysis of the related 3GPP/GSMA standards [3–7, 10]; however, V1 is not applicable to 5G networks, as described in Section 4.1. We elaborate on each vulnerability below. The V2 allows anonymous emergency UEs to establish IMS emergency sessions without doing IMS registration; it is a design issue of the IMS system. As the 4G network, the 5G cellular emergency service is supported by the IMS [4], so the V2 can also happen in 5G networks. The V3 stems from a design defect that the 4G MME does not know whether the 4G UE with an emergency IP-CAN session indeed establishes an IMS emergency session with PSAPs. We discover that this defective operation still exists in 5G networks. Specifically, the 5G AMF (Access and Mobility Management Function) [5] serving the similar role as the 4G MME is responsible for the emergency IP-CAN session establishment, but no interfaces are introduced for the communication between the AMF and the IMS emergency server; the AMF has no way of knowing any IMS emergency session establishment, so the V3 can be still applicable to 5G networks. For the V4, the PCF in 4G networks does not have information about the IMS emergency service server assigned to each emergency UE, so it cannot restrict the access of the UE's emergency IP-CAN session to the server only. According to the 5G standard [7], the PCF is still not given any information about the IMS emergency server. Thus, the V4 can be applied to 5G networks.

8 RELATED WORK

We classify the related work of the emergency service security into non-cellular and cellular categories.

Non-cellular Emergency Service Security. Several studies have been proposed to examine the security of non-cellular emergency services. Specifically, Goebel *et al.* [20] presented the vulnerabilities of the 9-1-1 call system from the perspectives of confidentiality, integrity, and availability. Fuchs *et al.* [19] developed an adapted intrusion detection architecture against the DoS attacks where a large number of faked VoIP-based emergency calls are generated. Seth *et al.* [34] designed a Wi-Fi based emergency service framework that enables mobile devices to contact the PSAP securely.

Cellular Emergency Service Security. The security issues of the cellular emergency service have attracted much attention in recent years. They can be classified into three categories. The first category of the studies is to launch or defend against the DDoS attack on the PSAP or the IMS emergency service server. Specifically, Mirsky *et al.* [29] showed that the adversary can jeopardize the statewide and nationwide PSAPs by generating random UE identities (e.g., IMEIs). Jung *et al.* [27] presented a CAPTCHA-based DDoS defense system that can protect the PSAP from DDoS attacks generated by compromised UEs (bots). Onofrei *et al.* [31] developed an adaptive firewall pinholing mechanism that can mitigate DDoS attacks against the server of the IMS emergency service. The second category is to examine the security issue that fabricated emergency/presidential alerts can be sent to UEs. Lee *et al.* [28] demonstrated that fabricated emergency alerts can be sent to UEs successfully. Hussain *et al.* [26] discovered that the adversary can hijack legitimate paging channels to send fabricated paging messages with emergency alerts to victim UEs successfully.

The last category is to exploit the cellular emergency service or resources to attack UEs or carriers. Hou *et al.* [25] developed two attacks based on the emergency service: UE screen lock bypassing and call service DoS. The first attack allows the adversary to dial any number on the emergency panel of the victim's UE and the call can be routed to the number owner, whereas the second attack can block phone calls made to a set of any numbers in a specific area. The present study belongs to this category; however, it differs from the above study from two major aspects as follows. First, the explored vulnerabilities and attacks are different; this study mainly presents the free data service, data DoS/overcharge, and DoCES attacks. Second, the adversary in the above study requires deploying a malicious eNB and let victim UEs connect to the eNB, whereas only SDR-based UE without SIMs is needed in this work.

9 CONCLUSION

Cellular networks offer mobile users with ubiquitous emergency services. For emergency uses, anonymous UEs are usually allowed to access cellular emergency services, according to regulatory authority requirements. However, such emergency support increases the attack surface of cellular networks. It leads us to discover four security vulnerabilities and exploit them to develop several attacks including free data service, data DoS, and DoCES. All of the vulnerabilities root in cellular design defects, which happen because some conventional non-emergency functions and services are directly applied to the emergency service operation without being carefully reviewed from the security aspect. We have experimentally validated the vulnerabilities and attacks with three major U.S carriers, and shown that both carriers and mobile users may suffer from the attacks. We finally propose short-term remedies and evaluate their feasibility, but the ultimate solution still requires a concerted effort from the standard community, carriers, and device vendors.

10 ACKNOWLEDGMENTS

We appreciate the insightful and constructive comments from our shepherd and the anonymous reviewers. This work is supported in part by the National Science Foundation (NSF) under Grants No. CNS-1750953, CNS-1815636, CNS-1814551, CNS-2112471, CNS-2226888, and CCF-2007159, and by the National Science and Technology Council (NSTC) under Grants No. 109-2628-E-009-001-MY3, 110-2221-E-A49-031-MY3, 111-2218-E-A49-023, and 111-2218-E-A49-013-MBK. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors only and do not necessarily reflect those of the NSF and NSTC.

REFERENCES

[1] 3GPP. TS 23.002: Universal Mobile Telecommunications System (UMTS); Network architecture (Release 5), Sept. 2003. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=728.

[2] 3GPP. TS 29.211: Rx Interface and Rx/Gx signalling flows (Release 6), June 2007. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1671.

[3] 3GPP. TS 23.167: IP Multimedia Subsystem (IMS) emergency sessions (Release 17), Sept. 2021. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=799.

[4] 3GPP. TS 23.228: IP Multimedia Subsystem (IMS); Stage 2 (Release 17), Dec. 2021. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=821.

[5] 3GPP. TS 23.501: System architecture for the 5G System (5GS) (Release 17), Dec. 2021. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144.

[6] 3GPP. TS 23.502: 5G; Procedures for the 5G System (5GS)(Release 17), Dec. 2021. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3145.

[7] 3GPP. TS 23.503: Policy and charging control framework for the 5G System (5GS); Stage 2 (Release 17), Dec. 2021. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3334.

[8] 3GPP. TS 24.229: IP multimedia call control protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3 (Release 17), Sept. 2021. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1055.

[9] 3GPP. TS 24.301: Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3 (Release 17), Dec. 2021. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1072.

[10] 3GPP. TS 24.501: Non-Access-Stratum (NAS) protocol for 5G System (5GS); Stage 3 (Release 17), Dec. 2021. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3370.

[11] 3GPP. TS 26.114: IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction (Release 17), Dec. 2021. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1404.

[12] 3GPP. TS 29.212: Policy and Charging Control (PCC); Reference points (Release 17), Sept. 2021. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1672.

[13] 3GPP. TS 29.274: 3GPP Evolved Packet System (EPS); Evolved General Packet Radio Service (GPRS) Tunnelling Protocol for Control plane (GTPv2-C); Stage 3 (Release 17), Dec. 2021. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1692.

[14] 3GPP. TS 33.203: Access security for IP-based services (Release 17), Dec. 2021. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2277.

[15] 3GPP. TS 36.213: Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures (Release 17), Jan. 2022. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2427.

[16] Androidcentral. Shutting and rebooting down during calls, 2015. https://forums.androidcentral.com/moto-x/575325-shutting-rebooting-down-during-calls.html.

[17] Aschenbruck, N., Frank, M., and Martini, P. Present and future challenges concerning dos-attacks against psaps in voip networks. In *Fourth IEEE International Workshop on Information Assurance (IWIA'06)* (2006), IEEE, pp. 6–pp.

[18] Communications, B. Linphone - for smartphones, tablets and desktop platforms, 2020. https://www.linphone.org/.

[19] Fuchs, C., Aschenbruck, N., Leder, F., and Martini, P. Detecting voip based dos attacks at the public safety answering point. In *Proceedings of the 2008 ACM symposium on Information, computer and communications security* (2008), pp. 148–155.

[20] Goebel, M., Dameff, C., and Tully, J. Hacking 9-1-1: infrastructure vulnerabilities and attack vectors. *Journal of medical Internet research 21*, 7 (2019), e14383.

[21] GSMA. Official Document IR.92 -IMS Profile for Voice and SMS (Version 15.0), May 2020. https://www.gsma.com/newsroom/wp-content/uploads//IR.92-v15.0-4.pdf.

[22] GSMA. Official Document NG.111 -SMS Evolution (Version 2.0), Nov. 2020. https://www.gsma.com/newsroom/wp-content/uploads//NG.111-v2.0-1.pdf.

[23] GSMA. Official Document NG.119 -Emergency Communication (Version 1.0), July 2021. https://www.gsma.com/newsroom/wp-content/uploads//NG.119-v1.0-3.pdf.

[24] Harvey. How to fix a Galaxy S9 that reboots on its own during calls, 2022. https://thedroidguy.com/how-to-fix-a-galaxy-s9-that-reboots-on-its-own-during-calls-1091448.

[25] Hou, K., Li, Y., Yu, Y., Chen, Y., and Zhou, H. Discovering emergency call pitfalls for cellular networks with formal methods. In *Proceedings of the 19th Annual International Conference on Mobile Systems, Applications, and Services* (2021), pp. 296–309.

[26] Hussain, S., Chowdhury, O., Mehnaz, S., and Bertino, E. Lteinspector: A systematic approach for adversarial testing of 4g lte. In *Network and Distributed Systems Security (NDSS) Symposium 2018* (2018).

[27] Jung, S. W. Captcha-based ddos defense system of call centers against zombie smart-phone. *International Journal of Security and Its Applications 6*, 3 (2012), 29–36.

[28] Lee, G., Lee, J., Lee, J., Im, Y., Hollingsworth, M., Wustrow, E., Grunwald, D., and Ha, S. This is your president speaking: Spoofing alerts in 4g lte networks. In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services* (2019), pp. 404–416.

[29] Mirsky, Y., and Guri, M. Ddos attacks on 9-1-1 emergency services. *IEEE Transactions on Dependable and Secure Computing 18*, 6 (2021), 2767–2786.

[30] of Federal Regulations, C. FCC 911 Regulations: 47 CFR Part 9: 911 Requirements, 2021. https://www.ecfr.gov/current/title-47/chapter-I/subchapter-A/part-9.

[31] Onofrei, A. A., Rebahi, Y., Magedanz, T., Institute, F. F., et al. Preventing distributed denial-of-service attacks on the ims emergency services support through adaptive firewall pinholing. *International Journal of Next-Generation Networks* (2010).

[32] OpenIMSCore.org. Welcome to Open IMS Core's Homepage., 2008. http://openimscore.sourceforge.net/.

[33] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and Schooler, E. SIP: Session Initiation Protocol, 2002. https://datatracker.ietf.org/doc/html/rfc3261.

[34] Seth, M., Kasera, S. K., and Ricci, R. P. Emergency service in wi-fi networks without access point association. In *Proceedings of the 1st International Conference on Wireless Technologies for Humanitarian Relief* (2011), pp. 411–419.

[35] srsRAN. Get the srsRAN software and documentation., 2022. https://docs.srsran.com/en/latest/index.html.

[36] Subudhi, B. S. K., Catal, F., Tcholtchev, N., Chiu, K. T., Rebahi, Y., Boerger, M., and Lämmel, P. Performance testing for voip emergency services: a case study of the emynos platform and a reflection on potential blockchain utilisation for ng112 emergency communication. *J. Ubiquitous Syst. Pervasive Networks 12*, 1 (2020), 1–8.

[37] Tschofenig, H., Schulzrinne, H., Shanmugam, M., and Newton, A. Protecting first-level responder resources in an ip-based emergency services architecture. In *2007 IEEE International Performance, Computing, and Communications Conference* (2007), IEEE, pp. 626–631.

[38] Tsiatsikas, Z., Kambourakis, G., and Geneiatakis, D. At your service 24/7 or not? denial of service on esinet systems. In *International Conference on Trust and Privacy in Digital Business* (2021), Springer, pp. 35–49.

[39] Voice, G. Google Voice calling rates, 2022. https://voice.google.com/rates.

[40] Wang, S., Tu, G.-H., Lei, X., Xie, T., Li, C.-Y., Chou, P.-Y., Hsieh, F., Hu, Y., Xiao, L., and Peng, C. Insecurity of operational cellular iot service: new vulnerabilities, attacks, and countermeasures. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking* (2021), pp. 437–450.

[41] ZeroMQ. ZeroMQ: An open-source universal messaging library, 2022. https://zeromq.org/.