Detecting Problematic Control-Plane Protocol Interactions in Mobile Networks

Guan-Hua Tu, Yuanjie Li, Chunyi Peng, *Member, IEEE*, Chi-Yu Li, and Songwu Lu, *Senior Member, IEEE*

***Abstract*—The control-plane protocols in 3G/4G mobile networks communicate with each other, and provide a rich set of control functions, such as radio resource control, mobility support, connectivity management, to name a few. Despite their significance, the problem of verifying protocol correctness remains largely unaddressed. In this paper, we examine control-plane protocol interactions in mobile networks. We propose CNETVERIFIER, a two-phase signaling diagnosis tool to detect problematic interactions in both design and practice. CNETVERIFIER first performs protocol screening based on 3GPP standards via domain-specific model checking, and then conducts phone-based empirical validation in operational 3G/4G networks. With CNETVERIFIER, we have uncovered seven types of troublesome interactions, along three dimensions of cross (protocol) layers, cross (circuit-switched and packet-switched) domains, and cross (3G and 4G) systems. Some are caused by *necessary yet problematic cooperation* (i.e., protocol interactions are needed but they misbehave), whereas others are due to *independent yet unnecessary coupled operations* (i.e., protocols interactions are not required but actually coupled). These instances span both design defects in 3GPP standards and operational slips by carriers and vendors. They all result in performance penalties or functional incorrectness. We deduce root causes, present empirical results, propose solutions, and summarize learned lessons.**

***Index Terms*—Control-plane, mobile networks, protocol verification.**

I. INTRODUCTION

THE 3G/4G mobile network is the largest wireless infrastructure deployed today, serving billions of mobile users with ubiquitous data and carrier-grade voice services. A salient feature of its design is its control-plane protocols. Compared with the Internet, these components provide more complex signaling functions. They follow the layered protocol architecture (see Fig. 1 for an illustration), and run at both network infrastructure and end devices. These protocols work together to offer control utilities vital to mobile networks, including radio resource control, mobility support, session management for data and voice, *etc.*.

In this work, we study control-plane protocol interactions in mobile networks. We focus on a set of critical functions (see Table I for the list), and seek to uncover problems during inter-protocol communications. Our research is motivated by three factors. First, problematic inter-protocol signaling each leads to functional incorrectness or performance penalty. For example, mobility management may make a wrong decision upon receiving duplicate signaling messages from the underlying radio resource control layer, thereby leading to network failure in that the user device unnecessarily loses its network access (out of service). Second, although each signaling protocol may be well designed individually, proper interactions among them in the networked environment are not guaranteed. Despite prior empirical assessment effort (e.g., conformance testing, field testing), verification for correctness on multiple protocol interactions through formal methods is still missing. Third, patterns of inter-protocol communication on the control plane are much richer and more complex than their Internet counterparts. They call for domain-specific verification. In addition to the cross-layer[1] (between layers of the protocol stack) case, protocol interactions exhibit in both cross-domain (between packet switching (PS) and circuit switching (CS) domains) and cross-system (between 3G and 4G systems) scenarios in mobile networks. Since both data and carrier-grade voice are indispensable services, signaling protocols thus regulate both PS and CS domains. Moreover, inter-system switching between 3G and 4G is also common due to incremental deployment, hybrid operation, user mobility, or even voice calls for 4G LTE users. Signaling protocols consequently need to work cross 4G and 3G systems.

To this end, we devise CNETVERIFIER, a two-phase signaling diagnosis tool. We first adopt publicly available 3GPP standard specifications as the reference design, and perform protocol screening using domain-specific model-checking methods. It helps us to determine a candidate set of potential design defects based on design documents only. Given this candidate set, we further instrument the device for empirical validation over operational 3G/4G networks. Through the validation phase, we not only identify real design defects, but also discover operational slips that may not show up during the screening phase. The use of 3GPP standards addresses the challenge due to a relatively closed system. Compared with the Internet, mobile networks remain rather closed: signaling exchanges are

Manuscript received August 08, 2014; revised February 11, 2015; accepted February 11, 2015; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor K. C. Almeroth. Date of publication March 13, 2015; date of current version April 14, 2016. This work was supported in part by NSF CSR-1017127, CNS-1421440 and CNS-1423576. G.-H. Tu and Y. Li contributed equally to this work. An earlier version of this paper appeared in ACM SIGCOMM 2014.

G.-H. Tu, Y. Li, C.-Y. Li, and S. Lu are with the Department of Computer Science, University of California, Los Angeles, CA 90095 USA (e-mail: ghtu@cs.ucla.edu; yuanjie.li@cs.ucla.edu; lichiyu@cs.ucla.edu; slu@cs.ucla.edu).

C. Peng is with the Department of Computer Science and Engineering, The Ohio State University, Columbus, OH 43210 USA (e-mail: chunyi@cse.ohio-state.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNET.2015.2404336

[1]We use inter-layer and cross-layer interchangeably in this paper.

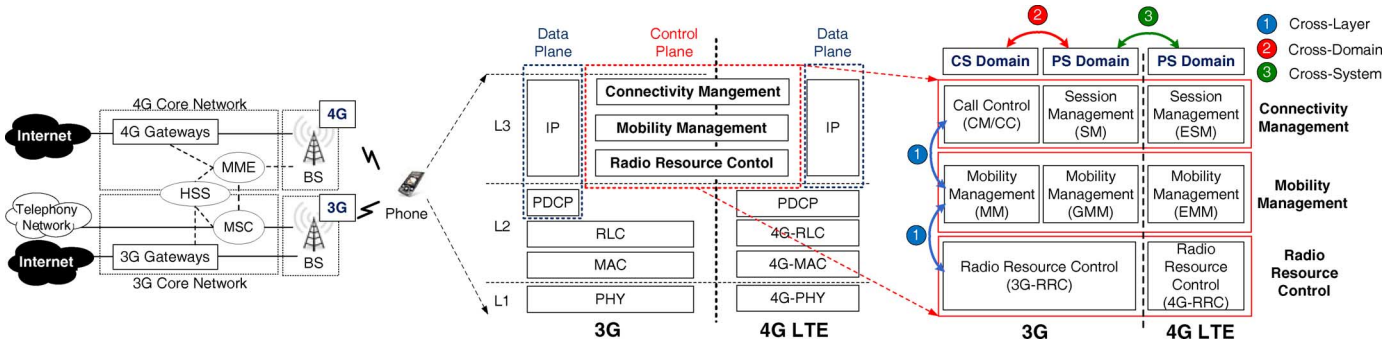


Fig. 1. 4G/3G network architecture and control-protocol interactions in three dimensions.

TABLE I
STUDIED PROTOCOLS ON NETWORK ELEMENTS AND DEVICES, AS WELL AS CURRENT COVERAGE OF CNETVERIFIER (NOTE THAT THE UNSUPPORTED SIGNALING MESSAGES AND STATES ARE FOR OPTIONAL, ADVANCED FEATURES, E.G., MULITICAST SERVICE OR GROUP CALL) OR INTERWORKING WITH NON-3GPP NETWORKS (E.G., CDMA2000 [10])

Function	Name	Net. Element	Standard	Description	#Signal	#Covered	Ratio	#State	#Covered	Ratio
PS/CS	CM/CC	3G MSC	TS24.008	CS Connectivity Management	35	14	40%	41	37	90.2%
	SM	3G Gateways	TS24.008	PS Session Management	24	17	70.8%	9	6	66.7%
	ESM	4G MME	TS24.301	4G Session Management	23	19	82.6%	7	7	100%
Mobility	MM	3G MSC	TS24.008	CS Mobility Management	22	15	68.2%	44	35	79.6%
	GMM	3G Gateways	TS24.008	PS Mobility Management	23	17	73.9%	25	25	100%
	EMM	4G MME	TS24.301	4G Mobility Management	32	21	65.6%	21	21	100%
Radio	3G-RRC	3G BS	TS25.331	Radio Resource Control	53	31	58.5%	5	5	100%
	4G-RRC	4G BS	TS36.331	Radio Resource Control	40	18	45%	2	2	100%
Total					252	152	60.3%	154	138	89.6%

not readily accessible from carriers, nor from devices during normal operations. It is thus difficult to both detect potential issues and validate them. To address state explosion issues in model checking, we exploit domain knowledge to model protocol behaviors and usage scenarios, and perform property checking with aggregation.

We apply our tool and delve into all above three dimensions. The study yields several interesting findings. We show two classes of problematic interactions among signaling protocols. They are exemplified using seven concrete instances (see Table III). In the first class, we show that some inter-protocol communications are necessary yet troublesome. The necessity of signaling synergy is partly driven by the requirement for carrier-grade voice support, partly by inter-system switching in hybrid 3G/4G deployments, and partly by mobility management. However, interactions among signaling protocols are not always designed and operated right: (S1) a user device is temporarily out of service because its vital context in 4G is shared but not well protected (being deleted after inter-system switching); (S2) Users are denied of network access right after being accepted because higher-layer protocols make unrealistic assumptions on lower layers; (S3) 4G users get stuck in 3G because inconsistent policies are used for CS and PS domains in 3G and 4G; (S4) 4G users lose their data services because the handoff policy to support voice calls between 2G/3G and 4G considers the CS domain only, thus imposing (unexpected) negative impact on the PS domain. The second class concerns independent operations by protocols. We discover that, some are unnecessarily coupled and have unexpected consequence: (S5) outgoing calls are delayed for unjustified location updates because cross-layer actions are "improperly" correlated and prioritized; (S6) User devices become out of service when the failure is propagated from one to another system; (S7) PS data sessions suffer from rate reduction (51%–96% drop observed) when traffic in both domains shares the same channel. We validate most instances with traces collected from our tool when running tests over two US carriers. We further conduct a four-week user study to assess their real-world impact. We further propose solutions that help to resolve above issues.

TABLE II
SUMMARY OF VIOLATIONS DURING THE SCREENING PHRASE

Property	#. violations	Last Visited State		
		CM	MM	RRC
DATA_OK	32	22	8	2
CALL_OK	8	0	8	0
MM_OK	5	0	0	5

The rest of the paper is organized as follows. Section II introduces the background on control-plane protocols and the problem to address. Section III describes CNETVERIFIER, our tool for protocol analysis. Section IV presents a case study to illustrate how CNETVERIFIER works. Section V discloses our findings on problematic interactions, and Section VI presents the proposed solution. Section VII compares with the related work and Section VIII concludes the paper.

II. CONTROL-PLANE PROTOCOLS IN 4G/3G NETWORKS

We first introduce necessary background on 4G/3G network architecture and essential control utilities, including connectivity management for data and voice, mobility management and radio resource control. We then describe the problem to verify control-plane protocol interactions and its challenges as well.

TABLE III
FINDING SUMMARY

Category	Problematic Interaction	Consequence	Protocols	Dimension	Type
Necessary but problematic cooperation	**S1:** States are shared but unprotected (deleted during inter-system switch) between 3G and 4G.	Users are temporarily *out-of-service* just after 3G→4G switch.	SM/ESM, GMM/EMM	cross-system	design
	S2: MME assumes and depends on reliable transfer of signals which is offered by RRC.	Users are temporarily *out-of-service* during the attach procedure.	EMM, 4G-RRC	cross-layer	design
	S3: RRC state change policy is inconsistent for inter-system switch.	User devices get stuck in 3G.	3G-RRC, CM, SM	cross-domain, cross-system	design
	S4: Voice service continuity is ensure at the cost of disrupted data service.	All data services are unable to use.	4G-RRC, EMM, ESM	cross-domain, cross-system	operation
Independent but coupled operation	**S5:** Location update is unnecessarily served at higher priority than outgoing call/data requests.	Outgoing call/Internet access is delayed.	CM/MM, SM/GMM	cross-layer	design
	S6: Information and action on location update failures in 3G are exposed (propagated) to 4G.	Users are temporarily *out-of-service* in 4G due to failures in 3G.	MM, EMM	cross-system	design, operation
	S7: 3G-RRC configures a shared channel with a single modulation scheme for both data and voice.	PS rate declines (*e.g.*, 96.1% in OP-II) during ongoing CS service.	3G-RRC, CM, SM	cross-domain	operation

A. Control-Plane Protocol Primer

Fig. 1 depicts the 4G/3G network architecture and the main protocols on the control plane. The network architecture consists of base stations (BSes) and the core network. The BSes offer radio access to user devices (e.g., phones), whereas the core network connects them to external networks such as the wired Internet or the telephony network. The 4G LTE network supports PS only. It has three core elements: 4G gateways, MME (*Mobility Management Entity*) and HSS (*Home Subscriber Server*). 4G gateways are mainly responsible for packet transfer on the data plane, forwarding packets between the Internet and the 4G BSes, akin to edge routers over the Internet. In addition, 4G gateways also perform several vital control functions for connectivity management, including IP address allocation, data connectivity setup/release, packet filtering, and policy enforcement. MME is the key control component for vital signaling functions for radio access control and mobility management, such as location update or paging. HSS is a centralized database that stores user information to facilitate control functions. In contrast, 3G supports both CS and PS services. Accordingly, its core network consists of 3G gateways in the PS domain and MSC (*Mobile Switching Center*) in the CS domain. 3G gateways relay data packets, similar to 4G gateways, whereas MSC pages and establishes CS services (i.e., voice calls) for the mobile devices.

Similar to the Internet, mobile network protocols have adopted a layered structure. The protocol family spans both data and control planes. The data plane is responsible for actual data and voice transfer, whereas the control plane provides a variety of signaling functions to facilitate the data plane. Specifically, there are three major control functions provisioned at three sub-layers: 1) Connectivity Management (CM), which is responsible for creating and mandating voice calls and data sessions; 2) Mobility Management (MM), which provides location update and mobility support for call/data sessions; 3) Radio Resource Control (RRC), which controls radio resources and helps to route signaling messages. We next introduce major procedures at each sublayer.

1) Connectivity Management (CM): CM regulates data and voice services within mobile networks, through Call Control (voice) and Session Management (data) in CS and PS domains. Specifically, to enable *data* service, the mobile device has to first establish a bearer with the core network in advance. This bearer offers a virtual pipe between the device and the 4G/3G gateway, which carries the subsequent IP data packets. This is realized through *Evolved Packet System (EPS) Bearer Setup Activation* procedure [2] in 4G, or *Packet Data Protocol (PDP) Context activation* procedure [3] in 3G, which is mandated by Evolved Session Management (ESM in 4G) or Session Management (SM in 3G). Once it succeeds, the core network assigns an IP address, reserves resources to meet QoS requirements and establishes the routing path for the device. The essential configuration for data sessions (e.g., IP address and QoS parameters) is stored and maintained in the 4G EPS bearer (or 3G PDP context) at both the device and the 4G/3G gateways.

In 3G, voice calls are supported in the CS domain and handled by the Call Control (CC) protocol at the phone and MSC. In 4G, they are designed to run over PS via Voice-over-LTE (VoLTE) technique [4]. However, due to high deployment cost and complexity of VoLTE, most operators adopt another voice solution, Circuit-Switched Fallback (CSFB), which switches 4G users to 3G and uses CS voice services in 3G [5].

2) Mobility Management (MM): Mobility management is to offer wide-area coverage and ubiquitous services for user devices. In terms of involved control protocols, mobility support is realized through MM, GMM, and EMM in 3G CS, 3G PS and 4G PS (see Fig. 1), respectively. In essence, it provides two core functions: location update (knowing where the mobile device is) and handoff/switch (changing its serving base station if needed). Location update is done through one of the following procedures: *location area update* via MSC (3G CS), *routing area update* via 3G Gateways (3G PS) or *tracking area update* via MME (4G). Mobile networks use two types of handoff: intra-system and inter-system. In an intra-system handoff, the user roams within 3G or 4G only, whereas in an inter-system switch, the user migrates between 3G and 4G. Once the migration succeeds, the device still updates its location to the new serving network via the above procedure.

In addition to mobility support, the attach/detach procedure is mandated by Mobility Management control protocols (i.e., MM, GMM and EMM) running on mobile devices, 3G MSC, 3G Gateways and 4G MME, respectively. The mobile device must *attach* to the mobile network before using any network service[2] (e.g., data or voice). It happens when the device powers on.

[2]The only exception is to make emergency calls.

Once completed, the device is "*registered*" and allowed to use network services until being detached. The *detach* procedure can be triggered either by the device (e.g., the phone powers off) or the network (e.g., under resource constraints). Once detached, the device enters the "*deregistered*" (i.e., "out-of-service") state and cannot access any service.

3) Radio Resource Control (RRC): RRC controls radio resources between the device and the BS. An established RRC connection is the prerequisite for any communication (data, voice or signaling) with the mobile network. RRC defines two states of IDLE and CONNECTED to represent whether the RRC connection has been established or not. To improve energy efficiency, RRC adopts multiple connected sub-states. Specifically, 3G possesses three sub-states of DCH, FACH and PCH, while 4G uses three modes of Continuous Reception, Short and Long Discontinuous Reception. Both DCH and Continuous Reception modes consume more power but send packets faster, whereas others sustain low-rate communication with less radio resource and power consumption.

B. Problematic Control-Plane Protocol Interactions

An individual protocol cannot work alone to offer any service in mobile networks. Multiple control-plane protocols have to work in concert through proper inter-protocol communication. For instance, making a voice call in 3G involves three signaling protocols of CM/CC, MM and RRC. A call request is captured by the call control module (CM/CC) on the caller phone. It invokes a call setup procedure through the signaling exchange between the caller and the callee. This further triggers the underlying MM and RRC to get ready for delivering the CM/CC signaling messages. Assume RRC is idle at the start. CM thus has to ask MM, which further triggers RRC to establish an RRC connection first. Once RRC activates a radio bearer between the phone and the BS, an MM session with the MSC is further established on top of this RRC connection, in order to transfer the subsequent call control signaling messages to the destination.

In this work, we verify the correctness of protocol interactions on the control plane. Although each signaling protocol may be well designed and tested individually, appropriate interactions among them in the networked environment are not guaranteed. We focus on a set of critical functions (connectivity management, mobility management and radio resource control) and seek to uncover troubling inter-protocol communications.

There are three main challenges. First and foremost, we must tackle interaction complexity. Compared with their Internet counterparts, control-plane protocols fulfill more functions and are more sophisticated. Their interactions not only cross different layers, but also exhibit in both cross-domain and cross-system scenarios. For example, the phone conducts PS data transfer for a mobile application upgrade, while making a call in the CS domain; In another common scenario, the phone runs Google Maps for navigation on a highway and has to migrate from 4G to 3G and vice versa due to insufficient 4G coverage. Such rich interaction patterns stem from diversified usage scenarios such as concurrent voice and data sessions, 3G/4G switch due to mobility and hybrid deployment. On the other side, incremental deployment and hybrid operations are natural with the evolution of mobile networks. The legacy 2G network is designed for CS voice, but gradually migrates to the 3G/4G technology that supports both data and voice. As a result, hybrid 2G/3G/4G deployment is the norm rather than an exception. During the evolution process, similar functions are required but realized by different signaling protocols. For example, mobility management is realized by MM in 3G CS, but by GMM in 3G PS, whereas by EMM in 4G. To support both data and voice, signaling protocols must regulate both PS and CS domains. Inter-system switch between 3G and 4G is also common due to hybrid deployment and user mobility. All lead to much richer interactions among signaling protocols.

Second, we must exploit *limited* information to reveal possible problems on both network and end-device sides. Different from the Internet, the mobile network infrastructure remains opaque. No full-stack protocol implementation or even the reference codes are released. Signaling exchanges within mobile networks are not readily accessible. There are no open-source diagnosis tools (akin to ICMP-based tools over the Internet) for us to probe the control-plane operations. Without open access to the infrastructure, black-box testing is deemed less effective to verify the operation correctness. Moreover, even when some suspicious problems are identified, it is harder to validate them and quantify their negative impact. The end device can expose very limited information during its normal operation mode. All these factors make the verification of signaling protocols more challenging than the Internet case.

Third, we need to explore as many scenarios as possible using a formal method. The industry has adopted a large number of testings, including conformance testing, KPI (Key Performance Index) testing and field testing, to ensure satisfactory performance of mobile networks [6]. These empirical approaches, which rely on test cases pre-defined by experts, have been largely successful for decades. However, they cannot formally specify exact test conditions to ensure protocol interaction verification. Existing tests measure the phone/network performance (e.g., conformance test [7], KPI test), while tolerating certain failure percentage (e.g., $<5\%$). Field tests run in the operational network environment, but their cases focus on common scenarios or stress tests. Since problematic interactions are not expected to be the norm, they might reside in these failure cases or less common cases that have not fully been tested. Moreover, these empirical runs are not good for troubleshooting; extra efforts are needed to identify their root causes. Therefore, we adopt a model-checking based method for formal verification. To this end, the concrete challenge is how to model protocol interactions without their real implementations. Some problems might be carrier-independent due to design defects in 3GPP standards, while others are caused by imprudent implementation and inappropriate configurations. We need to differentiate them whenever possible.

III. CNetVerifier: Mobile-Specific Verification of Control-Plane Protocol Interactions

We propose CNetVerifier, a **C**ellular **Net**work **Verifier** to examine control-plane protocol interactions. It helps to uncover both *design flaws* originated from the 3GPP standards and *operational slips* in the carrier's practice.

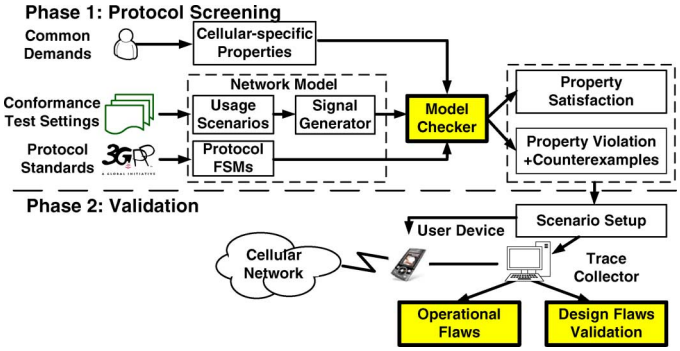


Fig. 2. CNETVERIFIER overview.

```
while (true) do
  msg = mm→receive(); //Blocked if no msgs received
  switch (mm_state) do
    case MM_IDLE   //MM state is idle
      switch msg do
        case LU request      // Receiving a location update request
          mm→send(RRC, RRC_CONN_REQUEST)    // activate RRC
          mm_state = MM_WAIT_RR_LU          // update mm_state
        case CM request      // Receiving a CM service request
          mm→send(RRC, RRC_CONN_REQUEST)
          mm_state = MM_WAIT_RR_MM
        case ...
        case Error
    case MM_WAIT_RR_MM   // another MM state
    case ...
end while
```

Fig. 3. Partial pseudo codes of MM in CNETVERIFIER.

Our tool carries out two-phase protocol diagnosis, as shown in Fig. 2. During the *screening* phase, it determines *possible* design defects in control-plane protocol interactions defined by 3GPP standards via model-checking techniques, and produces a super-set of counterexamples and their property violations. Afterwards, CNETVERIFIER moves to the *validation* phase. Given each counterexample, we set up its experimental scenario and perform tests over real networks. In addition to singling out real design flaws, these experiments further make possible to identify operational slips that are not discovered during the *screening* phase.

Our two-phase diagnosis approach seeks to address all three challenges. The screening phase is based on 3GPP design documents. This enables us to investigate the signaling exchange without accessing the carrier infrastructure or empirical traces. The publicly available 3GPP specifications serve as the reference design, which should regulate core functions of almost all mobile network implementations. Consequently, this phase might discover possible logical defects which are likely applicable to most mobile networks. To filter out those false design flaws from screening and assess their impacts, we further employ the validation phase over multiple carriers. With the assistance of the hints from the first phase, we perform empirical testing in the target scenarios. Moreover, the second phase helps to identify implementation- and measurement-dependent issues (operational slips and implementation bugs). In a nutshell, our approach helps to uncover as many problematic interactions as possible. Compared with the approach of empirical testing, our scheme not only avoids enumerating all test scenarios in practice, but also helps to deduce root causes to problematic behaviors (the exact design/operation flaws). We next elaborate on techniques for each phase.

A. Protocol Screening via Mobile-Specific Model Checking

The core of protocol screening is a mobile-specific model-checking tool, which is written in Spin[8]. It has two steps. First, we define mobile-oriented properties and model signaling protocol interactions and use scenarios. Second, given these inputs, CNETVERIFIER checks whether a set of desired properties are violated. If so, it thus outputs a counterexample for each concrete instance of property violation, which indicates a possible design defect. Although model checking has been widely used in protocol verification and diagnosis [9], three domain-specific issues need to be addressed in the mobile context: 1) How to model cellular networks? 2) How to define the desirable properties? and 3) How to check each property given the model?

1) Modeling Protocol Stacks: We start with modeling all three control functions in mobile networks. This is derived from the 3GPP standards, which specify operations for each protocol [11]–[14]. Table I lists the studied protocols, including three core functionalities of PS/CS services, mobility management and radio resource control, as well as current coverage of CNETVERIFIER.

Each protocol is modeled as two finite state machines (FSMs): one running at the user device and the other operating in the network. We use a virtual network element to denote multiple physical network components (for instance, CM/MM, SM/GMM, ESM/EMM are operated at MSC, 3G Gateways and MME, respectively). We follow the specifications and implement the abstract FSM for each protocol. Specifically, we first extract its high-level structure directly from the standards and define the corresponding states in FSM. In the three connectivity management protocols (CM/CC, SM and ESM), each state denotes the activation status of the session contexts (e.g., idle, pending, active). In mobility management, each state indicates the registration status of the user device (attached or detached), and/or the ongoing signaling exchange status between the user device and the core network (e.g., waiting for certain message, being connected or released); In RRC, the states represent the status of RRC connections between the user device and the BS.

We further model state transitions by leveraging their characteristics in the mobile network context. All state transitions are triggered or accompanied by the delivery and/or receipt of signaling messages. These messages include voice/data connectivity context activation/deactivation/modification, intra-/inter-system handoff commands, radio connectivity setup/release request/response, and user registration/de-registration to the network, etc. Some signals are initiated by usage scenarios while others follow the dependency of messages, regulated by the standard. More details are described below. Fig. 3 gives an example on how MM is partially represented in the FSM. In CNETVERIFIER, each FSM is event driven using signals from other protocols or input scenarios. FSMs of all protocols run concurrently.

To reduce model complexity, we simplify the representation of each signaling message, and safely neglects some fields in the message (e.g., checksum, user/network identity, physical layer

configurations). This is based on the premise that the concrete setting is used to complete functions at each individual protocol, which has been well designed and operated. Moreover, we simplify protocols not covered in Table I. We assume the protocols below RRC have two options: normal radio transmission and radio link failure. The one above CM is directly driven by user requests. We do not consider data-plane protocols and simply assume that all the transfer succeeds. Such simplifications do not compromise the completeness of the control-plane protocol interaction model, because we still enumerate all possible signaling exchange sequences in the screening phase.

In Table I, we further specify the number of signaling messages and states of each control-plane protocols defined by 3GPP standards and the current coverage by CNETVERIFIER. CNETVERIFIER covers around 89.6% (i.e., 138/154) of states, and supports 60.3% (i.e., 152/252) of signaling messages. Both are deemed sufficient for our study. We currently focus on the essential control utilities including (de)registration, authentication, call setup/release, data service session setup/release, location update, radio connection establish/release, handoff, etc. We do not cover the optional or advanced functions. For example, the advanced call service (e.g., hold, redirect, group, voice broadcast, send DTMF code), short message service and multicast service are not implemented. Second, those unsupported features do not require many new states (i.e., 16 here). Our CNETVERIFIER thus captures the major interaction cases.

2) Modeling Usage Scenarios: A usage scenario shows how the device accesses the mobile network and drives the transitions of all FSMs in CNETVERIFIER. Usage scenarios largely depend on user demands and operation policies; they are not defined by the 3GPP standards. They can be represented by [*traffic-type, mobility-type, network-type*], implying that the device uses data/voice/no service under a given mobility pattern, in a specific mobile network system. Specifically,

- *Traffic-type*. This can be either idle, data or voice or both. For data, we allow for various data rates.
- *Mobility-type*. This specifies mobility speed and the occurrence indicator of an intra-/inter-system handoff.
- *Network-type*. Each phone device uses at most one mobile network at a time, and cannot concurrently access both 3G and 4G. This is the default setting for most phones in reality. We test both scenario with the initial attach to 3G or 4G.

Ideally, we should test all combinations of usage scenarios for complete verification. However, certain usage scenarios may have unlimited choices (e.g., various speeds for user mobility, traffic arrival patterns of PS services). Enumeration is thus deemed unrealistic. Moreover, considering choices for parameters and event orders, the number of combinations explodes. In our previous work [1], we addressed this problem with sampling and randomization, but at the cost of certain missed cases.

In this paper, we take a different approach using guided sampling and randomization. We observe that although usage scenarios may have unlimited choices, they can be mapped to limited events. Such mappings are determined by discrete protocol event handling. For example, as traffic data rates vary within a large dynamic range from several bps to tens of Mbps, there is no difference on CM and MM, and only limited effect on RRC. In 3G, various data rates lead to four possible RRC states: DCH (high-rate), FACH (low-rate), PCH (only downlink broadcast, extremely-low), IDLE (almost no data, or with an extremely large interval). This is because RRC uses thresholds to determine its modes. The signaling protocols are designed to handle limited events, instead of handling infinite feasible inputs. We thus use predefined thresholds and a binary search algorithm to determine sampling parameters. Recall the data-rate example. Initially, we consider the minimum and maximum rates (say, 1 bps and 30 Mbps). If all network states and transitions are identical, we stop. Otherwise, its middle value (say, 15 Mbps) is considered next time unless the states are the same, given two different inputs.

3) Defining Desirable Properties: We focus on detecting those troublesome protocol interactions that lead to user-perceived problems. The properties to be checked denote the requirements for the services offered to users. Consequently, we define three mobile network-oriented properties: 1) DATA_OK: Packet data services should be always available once attached to 3G/4G, unless being explicitly deactivated. 2) CALL_OK: Call services should also be always available. In particular, each call request should not be rejected or delayed without any explicit user intervention (e.g., hang-up at the originating device). 3) MM_OK: Inter-system mobility support should be offered upon request as long as the coverage is allowed. For example, a 3G $\leftrightarrow$ 4G switch request should be served if both 3G and 4G are available. We only consider inter-system mobility, because intra-system mobility is always supported in practice. Note that, DATA_OK and CALL_OK represent the expected behaviors for network services, while MM_OK is for mobility support. In CNETVERIFIER, these properties act as logical constraints on the PS/CS/mobility states.

4) Property Checking: We perform formal model checking procedures to examine whether any property is violated. The model checker first creates the entire state space by interleaving all FSMs for individual protocols. For each scenario, the signal generator creates a sequence of initial signaling messages, which decides the initial state of the model. The depth-first algorithm is then applied to explore state transitions from the initial state (i.e., the device attempting to attach to 3G/4G networks) under different usage scenarios. In particular, for each state, if there can be multiple output signaling messages, we create a new brach from this state for each message. For example, given an RRC connection setup request, both accept and reject messages are considered. This way, we can test all possible cases for the responses. Moreover, for each message delivery, our implementation considers two possibilities of success and loss. This helps us to understand how the signaling protocols behave in response to the signaling loss/corruption. As a result, we enumerate all possible message delivery cases in a dynamic network environment.

To detect a violation, we derive constraints and mark certain states as an "error" for three properties.

- DATA_OK: except the normal deactivation state, all states whose connectivity context status are "inactive" are marked as erroneous states;

- CALL_OK: for each state, if the transition to this state involves a call rejection message, it is marked as an error;
- MM_OK: for each state, if it receives a 3G $\leftrightarrow$ 4G switch request but it does not have a direct transition to the other system, or it does not have available transitions to other states that can further move to the alternative system, it is marked as an error.

Once an error is hit, a counterexample, including all internal states, triggering events and case settings, is generated for the property violation. The protocol screening proceeds until all cases have been exhausted. Finally, it outputs all counterexamples and their violated properties.

To address the state explosion problem, we reduce the number of states to be searched through two measures. First, we leverage the domain knowledge of mobile networks. We do not intend to capture problems under those random and meaningless inputs. Instead, we are interested in identifying what issues occur in cases of our concerns. We refer to the conformance testing settings which are clearly stipulated by 3GPP standards [15]. As a result, we do not encounter severe state explosion since the benchmark test scenarios are not randomly created, but from 3GPP conformance testing designed for user devices. Second, given diverse usage scenarios, we avoid repetitive search of the same problematic state transitions. This is achieved by reusing the property-checking results from previous usage scenarios. For each scenario, once its property checking is finished, we aggregate all the found violations if they share the same error state. We check with the last visited state and trace back for longest suffix of all violations. To this end, we trace back each violation and find out the longest suffix of all violations. This suffix would end up in a state shared by all violations, which finally leads to the error state. With this state, we can analyze the root cause of the problem. Once all possible violations are covered, we mark it as "common error state". For followup usage scenarios to be checked, if the "common error state" is hit with the cause, we stop the search and directly report the property violations, together with the suffixes for it generated from previous scenarios.

B. Phone-Based Experimental Validation

The main task in the validation phase is to conduct experiments, collect protocol traces from real networks and compare them with the anticipated operations. There are two issues. The first one is how to reconstruct counterexample scenarios. Though the screening phase provides detailed settings, not all are feasible on the phone side. Some input events on the network side are still beyond our control. Constrained by this, we develop automatic tools on the phone to execute as many runs as possible. Specifically, we build `AutoCaller` (used in the case study) to automatically dial out, answer and terminate voice calls. For PS data services, we implement `AnyOnOFF` to keep switching on and off data services.

The second issue is trace collection. The current mobile network is operated as a black box. It is thus hard to obtain protocol traces from mobile operators. Instead, we retrieve protocol traces from user devices. Fortunately, most vendors allow for developers to retrieve signaling protocol traces using their debugging tools. We use QxDM [16], a Qualcomm extensible diagnosis monitor, for this purpose. We consequently collect five types of information: 1) timestamp of the trace item using the format of hh:mm:ss.ms(millisecond); 2) trace type (e.g., STATE); 3) network system (e.g., 3G or 4G); 4) the module generating the traces (e.g., MM or CM/CC); and 5) the basic trace description (e.g., a call is established).

We run experiments over two Tier-1 US operators. They together serve more than 140M subscribers. For privacy concerns, they are denoted as OP-I and OP-II. We use five smartphone models that support dual 3G and 4G LTE operations: HTC One, LG Optimus G, Samsung Galaxy S4 and Note 2, and Apple iPhone5S. They cover both Android and iOS. All phones are used in all validation experiments. In addition to the counterexamples from the screening phase, we also test with common usage scenarios to study whether any operational slip is observed to break three properties in practice. To measure the uplink/downlink speed of the Internet access, we use Speedtest (http://www.speedtest.net) on the phone. Each experiment has 10 runs unless explicitly stated.

1) Limitations: Before elaborating on our findings, we want to point out several downsides of CNETVERIFIER. First, it focuses on the control-plane protocol interaction, thus simplifying data-plane operations (e.g., ignoring data communication latency and call durations). Second, the defined properties are from the user's perspective. They may not uncover some issues concerned by operators. Third, though some heuristics are adopted in our model to reduce state explosion, there can be other effective techniques to completely eliminate the issue. Advances in model checking are also orthogonal to our work. For the scope of our study, we target the completeness of the model from the control-plane perspective. Fourth, due to limited access to mobile networks, not all findings can be validated by experiments. Finally, we mainly conduct experiments according to those counterexamples reported during the *screening* phase. Not all operation slips will be identified.

IV. CASE STUDY: CNETVERIFIER DURING A VOICE CALL

To illustrate how CNETVERIFIER works, we use "*dial-a-voice-call*" as the case study. We first demonstrate how troubling behaviors are identified in the screening phase and then present its root cause and empirical evidence. We further extend a single violation to multiple ones.

A. Screening: When a Call Request Meets a Location Update

We consider a usage scenario of making an outgoing call when an intra-system handoff occurs in 3G networks. That is, our input setting is as follows: $traffic$ = outgoing-voice-only, $mobility$ = intra-system handoff, and $network$ = 3G. Given this scenario, depending on the order of voice request and mobility, the input message generator produces two sequences of initial messages: (a) attach request $\rightarrow$ call request $\rightarrow$ location update; and (b) attach request $\rightarrow$ location update $\rightarrow$ call request. Protocol screening runs as specified in Section III-A. The messages are fed to the FSMs, triggering the corresponding state transitions. In subcase (a), both the location update and the call requests are satisfied. No violations are reported. However, in subcase (b), a call rejection message is invoked, leading to an error against CALL_OK. As a result,

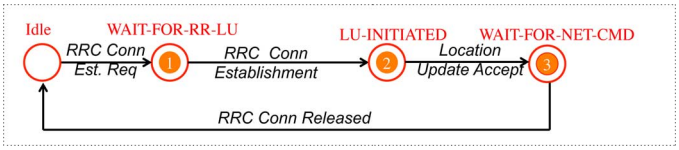


Fig. 4. FSM of the location update procedure at MM.

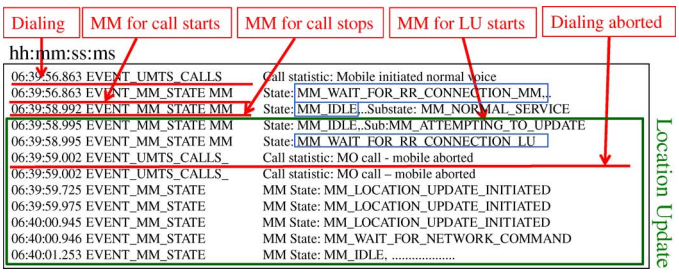


Fig. 5. A signaling trace of an aborted-call example. A call dialing when driving aborts due to the location update at MM.

CNETVERIFIER reports a CALL_OK violation, together with the complete message sequences and state transitions.

The problem occurs when the call request message arrives while the MM stays at a special waiting state (called as WAIT-FOR-NET-CMD). Fig. 4 shows a simplified FSM for location update at the MM sublayer. Following the location update procedure, the MM on the phone side first establishes an RRC connection and then updates its location to MSC (State 2). Upon receiving the accept message, it enters into the WAIT-FOR-NET-CMD state, which is an interim state before releasing its MM session. However, according to the 3GPP standard [11], this state is unable to serve user call requests. Any new call request is either immediately rejected or delayed at this state, until MM returns to the IDLE state.

B. Validation: Observations in Real Networks

We validate the above finding using experiments over a major US carrier. Since we are unable to control or predict when the location update is performed in reality, we run a driving test on the freeway, while making successive call dialings. We use `AutoCaller` to repeat dialing-hangup on our test phone. To examine how these protocols react, we use QxDM [16], a testing tool provided by Qualcomm, to record signaling traces on the phone.

Fig. 5 gives an example trace, where the dial request gets aborted due to location update. When moving on the freeway, a call is dialed from the car at the 56th second and MM switches to the WAIT-FOR-RR state as expected. However, at the 58th second, MM changes back to IDLE and starts the location update procedure in 0.003 s, followed by the dialing abort at the 59th second. The call dialing is interrupted and rejected due to location update running in MM. This shows how the CS voice call is blocked due to independent location update (Instance S5), which has been described in our previous work [1]. Moreover, location update can impede PS data services, similar to its blocking in the CS domain.

C. Generalization to Other Instances

We find more violation instances, where call dialing can be denied or delayed due to other events. In fact, we identify

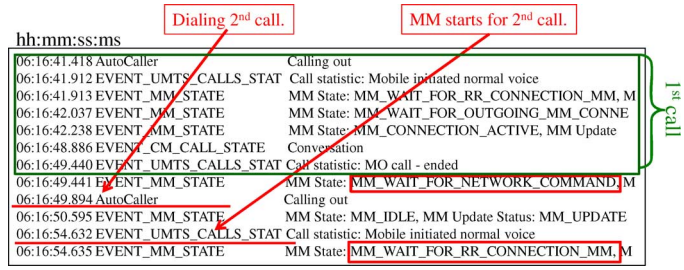


Fig. 6. A signaling trace of a delayed-call example. The dialing is delayed for about 5 seconds by MM.

eight violations against CALL_OK in the screening phase (see Table II). Similar to the above example, all these violations share one common erroneous state WAIT-FOR-NET-CMD, as well as the event trigger (dialing arrives when MM is in this state). The eight instances represent all possible reject causes, which are stipulated in [11], including network failure, congestion, no authorization, service option not supported, request service option not subscribed, temporarily out of order, etc. In case of a found violation, we further enumerate all the possible causes and mark this erroneous state complete. This way, for other scenarios to be checked, WAIT-FOR-NET-CMD would be marked as "common error state". Any scenario that hits this state with the same reason would incur property violation. This ensures the comprehensive coverage of all erroneous cases, and alleviates the issue of state explosion.

Taking this approach, we can extend the result from this case study to other usage scenarios. We report another example setting where the same violation occurs but under a different condition. It occurs when the user re-dials right after a call ends. There is no location update. This is a reasonable user request. However, it is not well supported. Once the call ends, MM enters into the WAIT-FOR-NET-CMD state. For the same reason, the user dialing request cannot be served. However, different from the location update example, we discover the call is just delayed but not aborted. Fig. 6 shows an example trace. It is easy to see that "*dial-a-voice-call*" takes about 5 seconds to respond (i.e., the new call is dialed at the 49th second but being delayed until the 54th second). In contrast, the first call takes merely 0.8 seconds to go through after dialing (the wait time for processing is about 0.5 seconds). The delay difference is about 4.5 seconds. This extra delay is consistently observed in all experiments.

V. FINDINGS ON PROBLEMATIC INTERACTIONS

Through CNETVERIFIER, we uncover protocol interaction problems in both design specifications and operational practices. In this section, we first give an overview of our major findings. We then briefly describe the instances disclosed in our previous work [1], and extend the findings to a new feature in mobile networks. We uncover another problematic interaction which has not been reported yet. Finally, we present results from a 20-volunteer user study.

A. Overview and Classification

Table II summarizes 45 identified violations during the screening phase. All those violations against CALL_OK and MM_OK occur in an erroneous state at MM and RRC. We

classify all the violations based on their root causes and summarize them as seven instances in Table III. There are two categories. The first class, *necessary yet problematic cooperations*, denotes those protocol interactions that are required but misbehave. The second class, *independent yet unnecessarily coupled operations*, refers to those interactions that are not needed but indeed occur. Both categories are found in all three dimensions: cross-layer, cross-domain and cross-system.

1) Cross-Layer: Upper-layer and low-layer protocols directly interact with each other via the interface between them. Two representative instances are found in this category. In both cases, the principle of protocol layering is not properly honored. In the first instance (S2), the low-layer RRC protocol fails to offer reliable and in-sequence signal delivery required by the upper-layer EMM protocol. Subsequently, the signaling exchange between the device and the network can be lost or delayed, triggering wrong reactions from EMM. It denies user's network access right after accepting the access request. In the second instance (S5), CM/SM and MM/GMM protocols, running on different layers in 3G, should act on outgoing call/data requests and location updates independently and concurrently. However, they prioritize location updates over call/data requests. This incurs head-of-line blocking, and outgoing calls and data are unnecessarily delayed (or even aborted).

2) Cross-Domain: The CS-domain voice and the PS-domain data have distinctive requirements. Data values throughput and lossless delivery, whereas voice prefers timely (even lossy) delivery. They thus demand differential treatment. As a result, standards stipulate different protocols for each domain (CS/PS). However, protocols in both domains are not always distinguished. Three instances are discovered, where identical operations are performed on traffic from both domains. In the first case (S3), RRC keeps its state for the aggregated CS and PS data traffic. Due to improper coordination on RRC states, when the CS traffic terminates, the PS data may get stuck in 3G without going back to 4G. In the second case (S4), MM and RRC select a 2G network to support voice call continuity when a 4G LTE phone moves into an area covered by 2G and 3G only. Such handoff selection only takes CS into account but neglects the ongoing PS service. The PS service thus often gets aborted since data service is not well supported in 2G. In the third case (S7), carriers ask RRC to assign PS and CS sessions over a shared channel, using a single modulation scheme. The PS data rate thus drops significantly.

3) Cross-System: Cross-system interactions occur upon 2G/3G $\leftrightarrow$ 4G switch. In this scenario, both systems may be motivated to share or even act on certain state information. However, their interactions might improperly protect and utilize the shared information. Four instances are also uncovered. Information vital to data services is contained in the PDP context in 3G, which is equivalent to the EPS bearer in 4G. Such contextual information should be correctly protected while being shared during the cross-system operations. However, in the first finding (S1), 3G is allowed to delete the PDP context. Consequently, the 4G network cannot recover its EPS bearer state after 3G $\leftrightarrow$ 4G switch. The user device is thus temporarily out of service in 4G. In the second case (S6), both 3G and 4G share information on location update failures. However, actions on location update failures from one system are unnecessarily propagated to the other, while they should be confined within either 3G or 4G. 4G acts on the user device to handle failure signals from 3G. The user consequently loses its network access. Both S3 and S4 suffer from the problematic cooperation in CS and PS domains, as well as that in 4G and 3G/2G systems.

B. Previous Findings on Problematic Interactions

We summarize each problematic instance (S1–S3 and S5–S7) of our previous work. More details can be found in the conference version [1].

1) S1: Unprotected Shared Context in 3G and 4G: The first instance arises during *cross-system* signaling exchanges between 3G and 4G. When the user device switches from 4G to 3G or vice versa, data services are indeed migrated accordingly. However, under certain conditions, when the user switches back to the 4G network, the device might be temporarily *out of service*. The involved protocols are SM/GMM in 3G and ESM/EMM in 4G. These protocols should interact, because they need to support seamless PS data sessions during inter-system switches between 3G and 4G. They thus share contexts in 3G and 4G. However, such shared states, which are mandatory in 4G, might be deleted in 3G, thus causing state recovery failure after a successful inter-system switch. Specifically, the protocol screening phase reveals how it occurs. The context contains critical information for the PS connectivity (e.g., IP address remains the same before and after the switch). A violation against DATA_OK is reported when the device moves back to 4G after the 3G PS connectivity context is deactivated in 3G (for various reasons listed in [1], including insufficient resource, unaccepted QoS, regular deactivation, etc.). We note that the inter-system switch between 3G and 4G is commonly observed in practice, for example, when roaming in an area with hybrid 3G/4G deployment, or due to load balancing of user traffic or for 4G voice call support. Our experiments validate its existence and show that this out-of-service status may last from several to tens of seconds in operational networks.

2) Lesson: Regarding the shared context, the actions and policies should be coherent between different systems. Otherwise, cross-system inconsistency may arise.

3) S2: Out-of-Sequenced Signaling: S2 is induced by improper cross-layer protocol interactions in 4G networks. Two involved protocols are EMM and RRC. EMM takes wrong actions when communicating with RRC. It assumes that RRC offers reliable, in-sequence signaling messages, but this is not guaranteed by the underlying RRC protocol. Even worse, the EMM design does not anticipate any lost or delayed signaling exchange. This leads to unexpected consequence. The user device is detached from 4G right after a successful attach. It thus is temporarily *out-of-service* and loses 4G access. In particular, the problem occurs during the attach procedure which is used to register the user device before using any mobile network service. It is realized through a series of signaling exchange between the user device and the MME: attach request $\rightarrow$ attach accept $\rightarrow$ attach complete $\rightarrow$ location update. CNETVERIFIER identifies two error cases with lost or duplicate signals. The first

case happens when the *attach complete* message from the device is lost. MME does not receive the attach complete message and thus has an inconsistent EMM state from the device. The device believes it is registered and thus triggers a location update, while the network believes the device is deregistered and rejects this request. This rejection leads to the network detachment on the device side after the prior attach success. The second case is observed when duplicate *attach request* messages are received at MME. The device fails to receive the attach accept response from the MME (possibly due to heavy traffic load or wireless loss), and retransmits the request. Upon receiving this duplicate signal, EMM at MME deregisters the device and remove its PS connectivity context according to the standard [13].

4) Lesson: During cross-layer protocol interactions, the core operation of upper-layer protocols should not rely on the not-always-guaranteed features from lower-layer protocols. Otherwise, they operate at the risk of failures.

5) S3: Stuck in 3G Due to Inconsistent Cross-Domain/System State Transition: S3 is induced by inconsistent state transitions during the CSFB (CS Fallback) procedure [5]. CSFB is a voice solution to the LTE network. It leverages the CS domain in 2G/3G networks since LTE supports PS only. In the worst case, a 4G device gets stuck in 3G, thus losing its 4G connectivity and high-speed access. Such a consequence is not anticipated by the design of CSFB, which intends to move the device back to 4G after the call. In particular, the problem occurs when certain PS service is still ongoing after the CSFB call. It is related to 3G RRC. The standards offer three options to migrate 3G to 4G: 1) *RRC connection release with redirect* if the RRC state is not idle; 2) *inter-system handoff*, if the RRC state is in a specific connected state; and 3) *inter-system cell selection*, with RRC being idle. However, for those carriers with Option 3) only, the device gets stuck in 3G as long as the PS service is ongoing (RRC is not idle) after CSFB voice call is released. We validate it in our experiments over two U.S. carriers. This is also observed in another recent study [17]. Note that the inter-system switch works well without CSFB. CSFB is a new procedure introduced to the existing infrastructure. It induces a 3G $\rightarrow$ 4G switch when the call ends in the CS domain. However, its RRC state is also affected by the PS domain, not only the CS domain. This unfortunately disables the 3G $\rightarrow$ 4G switch in the carrier network adopting the option of *inter-system cell selection*.

6) Lesson: The original, well-crafted functions may become error prone as new features (e.g., CSFB) are introduced. All options should be prudently examined and regulated. Otherwise, the desirable functions may be compromised by certain overlooked options.

7) S5: HOL Blocking for Independent Updates: S5 is on unnecessary coupling between cross-layer protocols in 3G. Both voice and data may suffer from head-of-line (HOL) blocking, due to independent, yet unnecessarily prioritized location update at underlaying layers. The involved protocols are CM/MM and SM/GMM for the CS domain and the PS domain, respectively. The details are elaborated in the case study (Section IV). Location update is one major function in mobility management. Without it, the network cannot route *incoming* calls or packets to the user. It is not only performed for roaming users, but also used for periodic refresh without mobility or after inter-system switching. In 3G CS domain, the *location update* is initiated by MM protocol on user device, and sent to MSC. However, the location update will make MM migrate to the `WAIT-FOR-NET-CMD` state, which rejects to serve the CS/PS service request from CM/SM. The root cause is that location updates are processed with higher priority. However, this reasoning has a catch. Note that the call/data request is *outbound*. The network does *not* need to know the device location. Moreover, If this call request is served first, MSC also *implicitly* updates the location for the device as a byproduct of call serving. This implicit update is realized without any extra resource usage. We conduct validation experiments of driving with (a) call requests or (b) data services. We indeed observe the extra delay (mainly 2 –5 seconds). In the worst case, we even observe that call requests might abort due to location update on popular smartphones (e.g., Samsung Galaxy S4).

8) Lesson: Some procedures at upper and lower layers are independent, yet coupled in their execution order. Without prudent design, HOL blocking may occur.

9) S6: 3G Failures Propagated to 4G System: S6 is a cross-system coupling case. It is an operational slip found in our experiments. The involved protocols are MM in 3G and EMM in 4G. In both OP-I and OP-II, the failure of location area updates in 3G is propagated to 4G, and processed by 4G. This may force 4G users out of service temporarily. Two location updates are performed in 3G when using CSFB for voice calls. The first update is needed after the 4G $\rightarrow$ 3G switching once the call starts. It is initiated by the device. When the call completes, the second update in 3G is activated after the device switches back to 4G. It is done by the network. The update is first processed by MME in 4G, which relays the update request to MSC in 3G. Among both location updates, one is deemed redundant. It yields no benefit, but incurs penalty. In OP-I, the first update hurts. This update is delayed until the call terminates. In OP-II, the second update causes damage. The first update completes first, since it takes more time for the carrier to switch from 3G back to 4G. The success of the first update may trigger MSC in 3G to refuse the second update request. It leads to a detach request sent by 4G to the device, and the user becomes *out-of-service*.

10) Lesson: For similar functions in different systems, their failure-handling actions should be coordinated to resolve conflicts. Naively exposing them to user devices is not a good practice.

11) S7: Fate Sharing for Voice and Data: S7 is another operational problem discovered during the validation experiments on S3 (CS+PS services). We observe that, when both PS and CS access the 3G network from the phone, the PS data rate decreases significantly, compared with the case of 3G PS only. For example, the downlink throughput decreases up to 3.5–5.8 Mbps, about 73.9% in OP-I and 74.8% in OP-II. This is due to improper cross-domain coupling between PS and CS in 3G.

Our collected traces show that, both carriers use RRC to configure radio channels for CS or PS or both. For concurrent CS and PS traffic, RRC uses a shared channel and applies a single modulation scheme. Before the call is made, RRC uses a high-rate modulation scheme . Once the voice call starts, both OP-I and OP-II downgrade the highest-rate modulation to the lower one. Consequently, the user thus suffers from large rate drop in

its data service. The root cause is that CS voice and PS data share the fate (using the same RRC configuration) while they indeed have distinctive requirements (low-throughput and high-reliability for voice but high-throughput for data).

12) Lesson: When two domains have different goals and properties, their services should be decoupled whenever possible. Otherwise, one domain's services may be sacrificed.

C. New Problematic Interactions in VoLTE

We next apply CNETVERIFIER to assess a new feature in 4G networks, VoLTE (voice over LTE) [4]. VoLTE has been designated as the ultimate voice solution in 4G LTE, though its actual deployment in the US has just started (since late 2014) [18], [19]. Instead of leveraging the CS domain in the legacy 2G/3G network, VoLTE directly supports voice calls in the PS domain. It carries voice calls in PS packets, akin to voice over IP (VoIP) in the Internet.

1) CNETVERIFIER *in VoLTE:* Two standard procedures are needed to support the VoLTE service in 4G networks. VoLTE [4] regulates basic call operations in the LTE PS domain. SRVCC (Single Radio Voice Call Continuity) [20] handles the case when the user leaves the 4G coverage. In SRVCC, an ongoing VoLTE call is migrated from the 4G network to a CS-based voice call in 2G/3G networks.

We extend CNETVERIFIER to support signaling protocols for both procedures in three aspects. First, VoLTE shares the PS domain with conventional PS data services, but uses high-priority radio resource control to ensure good voice quality. To this end, 4G ESM (Session Management) and 4G RRC use separate PS connectivity for VoLTE, which is assigned with higher priority (i.e., 1 for VoLTE control-plane messages and 2 for VoLTE data-plane voice packets [21]). We thus modify the 4G ESM and 4G RRC modules to allow for the establishment and release of multiple connectivity, as well as different priority levels. Second, SRVCC requires a cross-system, cross-domain handoff (from 4G PS to 2G/3G CS) when a VoLTE call user moves into the non-4G area. We thus update the FSMs for CM/CC, SM, MM, GMM, 3G-RRC, ESM, EMM, 4G-RRC based on the 3GPP standard [20]. For example, we update the 4G RRC signaling message to specify its capability to support SRVCC. Third, we add VoLTE test scenarios accordingly. All conventional CS calls are replaced by VoLTE calls if applicable (in 4G LTE). Mobility settings with various coverage combinations of 4G/3G/2G, 3G/2G, or 3G only, etc. have been considered.

2) Issues and Root Causes: We identify three instances which violate DATA_OK and MM_OK. The first two are reported in the screening phase, and have the same causes as S1 and S3. The only difference is that the inter-system switch is incurred by SRVCC (i.e., the VoLTE user roams from 4G to 3G). Specifically, in S1, the PS connectivity context shared in 3G and 4G is still deleted in 3G, and it cannot be recovered when it moves back to 4G. For S3, the user device still gets stuck in 3G without going back to 4G even when the voice call ends with an ongoing PS data session. The problem is identical to the case when CSFB is used. In fact, it indicates that S1 and S3 are caused by common design defect of cross-system/cross-domain protocol interactions. The inter-system switch suffers from these inappropriate interactions, no matter how this switch is triggered (by user mobility, CSFB or VoLTE-SRVCC). We further conduct experiments to validate and assess the above issues on OP-I, since OP-II currently does not support SRVCC. We first make a VoLTE call and then move out of the 4G coverage. We see that, SRVCC is triggered and the user device moves to 3G. We then deactivate all PS connectivity contexts in 3G (from Setting $\rightarrow$ Mobile Network $\rightarrow$ Mobile data) and move back to 4G. We observe that, user device would first become "out-of-service" and then return to the 4G LTE network. For S3, we run the same test of initiating a 60-minute UDP download session at 32 kbps rate when moving back to 3G. We then hang up the voice call. The user device gets stuck in 3G for an hour. Both experimental results match our previous findings in S1 and S3.

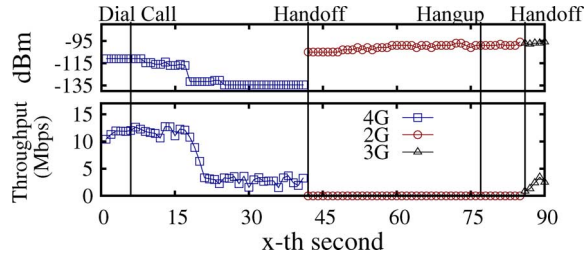


Fig. 7. Throughput of user's UDP downlink session while a VoLTE call is handoffed to 2G CS call by SRVCC.

We further identify a new operational problem S4 in the validation. We find that, data services are disrupted due to an inappropriate handoff performed by SRVCC. When the user leaves 4G and moves to an area with 2G and 3G coverage only, EMM/GMM/MM (driven by SRVCC) does not always move to 3G. Instead, 2G may be chosen. Once the user is moved to the 2G network, the transmission rate of data service quickly shrinks to zero. Even after voice call ends, it still gets stuck in 2G. Fig. 7 plots the data throughput and the radio signal strength of the serving base station where 2G is selected by SRVCC. The ongoing data services (e.g., android FTP) might get aborted during SRVCC.

This problem is rooted in SRVCC, which only considers whether voice call continuity can be ensured, regardless of the ongoing data service continuity. However, its handoff decision does affect both CS and PS domains. EMM is myopic to choose 2G when its signal strength is slightly stronger than 3G, but neglects its impact on the ongoing data service. Due to poor data support in 2G, mobile users cannot obtain any data service any more. Even worse, when the call ends, there is no extra mechanism in MM to trigger the device back to 3G. In fact, MM determines its target network simply based on the measured signal strength in 2G and 3G in practice.

We conduct more experiments to study how SRVCC selects the target cell (2G or 3G) for voice call continuity. Our test consists of four steps: 1) initiating a UDP downlink session in 4G LTE; 2) establishing a VoLTE call conversation; 3) walking along the test routes where SRVCC is observed and stoping walking after SRVCC is triggered; and 4) measuring the signal strength and data throughput of the serving cell for 30 seconds. We run experiments on two test routes (Route-I and Route-II) with 50 runs each. The results are summarized in Table IV. We make two observations. First, there are around 16–18% of

TABLE IV
PERFORMANCE OF SRVCC WITH ONGOING DATA SERVICES ON TWO ROUTES

	Route-1		Route-2	
Legacy voice	2G CS Call	3G CS Call	2G CS Call	3G CS Call
Signal (dBm)	-101 ∽ -105	-103 ∽ -110	-103 ∽ -105	-91 ∽ -95
Ratio[3]	38% (19/50)	46% (23/50)	0% (0/50)	82% (41/50)
Throughput	0 Mbps	1.6 Mbps	0 Mbps	5.1 Mbps

SRVCC inter-system switch failures on two routes. It implies that SRVCC is not well supported so far. Second, the network migrates the VoLTE call and data sessions to 3G networks on Route-II, where 46% to 3G networks and 38% to 2G networks on Route-I. This is because the signal strength of 2G networks on Route-I is comparable to 3G networks ($-101 \backsim -105$dBm v.s $-103 \backsim -110$ dBm), whereas the result is different on Route-II. SRVCC does not take into account of concurrent data services, and has no preference when selecting 3G or 2G.

3) Lesson: The inter-system switch decision should not only satisfy the demand in one domain but also consider the other.

D. User Study

To assess the real-world impact, we conduct a four-week user study with 20 volunteers, including students, faculty members, engineers and technology-unsavvy people. 12 people use 4G-capable phones, while others use 3G-only phones. Table V summarizes the results for seven instances S1-S7. Compared with our preliminary user study in [1], similar findings are observed but their occurrence ratios are lower.

S1: A user in 3G fails to switch to 4G if its PDP context is deactivated. We observe about 2.8% for S1 events in case of 4G $\rightarrow$ 3G switches with enabled mobile data access.

S2: 43 attaches are observed but none of them fails. It implies that S2 rarely occurs, possibly because all are performed in the area with good coverage (the weakest signal is -97 dBm).

S3: In S3, users do not immediately return to 4G when a CSFB call ends. Among 214 CSFB calls, 115 (39 in OP-I and 76 in OP-II) are made while mobile data is enabled. Our results show that OP-I users usually switch back to 4G within 3 seconds. It is because OP-I uses "*RRC Connection Release with redirect*," which can be triggered at RRC Non-IDLE state. However, OP-II users get stuck in 3G much longer because OP-II performs "*inter-system cell selection*," which occurs only at RRC IDLE state.

S4: Since the VoLTE service is only supported in a few new phone models in OP-I and OP-II to date, we do not observe any VoLTE call from our participants.

S5: We consider the HOL blocking issue for 3G CS calls. We check whether there is any location area update done in 1.2 s right after the outgoing call starts, because this update takes at least 1.2 s to complete. We observe 320 outgoing calls out of 506 CS calls in 3G. Eight (i.e., 2.5%) are affected. In case of longer location area updates (>1.2 s), the ratio is larger. For example, if we take the median time of update observed as criteria (i.e., 2 s and 1.8 s for US-1 and US-2), the occurrence ratio increases to 3.8%.

S6: We examine how often CS calls affect PS data traffic and how much data is affected during a call. It is observed that 34.2% 3G CS calls (173 out of 506) happen while data traffic is ongoing. For these calls, the average duration is 94 s, and the average affected data volume is 338 KB. Most calls (167 out of 173) affect the data volume less than 550 KB, whereas the remaining one call affects more than 1 MB data, and five calls affect more than 4 MB data (with the largest being 18.5 MB).

S7: In addition to S1, the failure of location update required by CSFB calls make the users fail to switch back to 4G after a CSFB call. It happens in 5 out of 214 calls (2.3%).

TABLE V
SUMMARY OF USER-BASED STUDY ON S1–S7

Problem	S1	S2	S3	S4	S5	S6	S7
Observed	√	×	√	×	√	√	√
Occurrence Prob.	2.8% (4/141)	0.0% (0/43)	66.1% (76/115)	0% (0/0)	34.2% 173/506	2.3% 5/214	2.5% 8/320

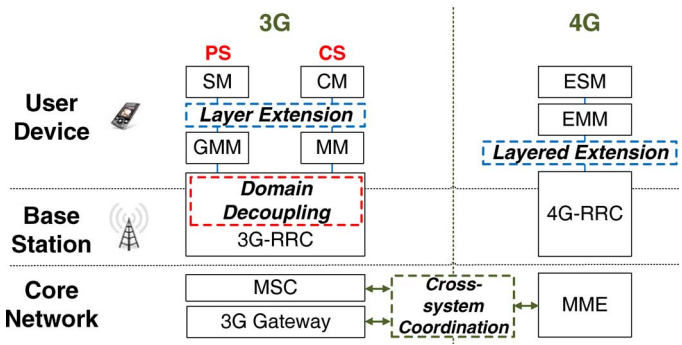


Fig. 8. Solution overview.

The above user study with small samples may not accurately quantify the real-world impact and can be further improved with more participants. The result partly confirms that current cellular networks are largely successful. However, it also shows that the found issues do occur in our daily life and affect our actual data usage.

VI. SOLUTION

We next present our solution, as well as its recommended implementation and prototype-based evaluation. It applies design guidelines along three dimensions. For cross-layer, protocol layering should be strictly honored. A lower-layer protocol should fully meet its upper layer's requirements without unnecessary blocking. For cross-domain, services from two domains should be decoupled. Actions for one domain should neither interrupt nor be constrained by the other. For cross-system, similar functions in different systems should collaborate to minimize conflicts.

The overall solution is shown in Fig. 8. It has three modules of layer extension, domain decoupling and cross-system coordination. We next elaborate on each component.

A. Layer Extension

We propose a slim layer with reliable transfer for the out-of-sequence signaling at EMM, and parallelize independent operations. In the former, the slim layer is inserted between EMM and RRC. Its reliable transfer ensures the end-to-end, in-order signaling exchange between the phone and MME (S2). To be compatible with the current system, it bridges the interface between EMM and RRC, and encapsulates the information on reliable transfer. For the latter, location update should be decoupled

from the CS or PS service request for MM and GMM, respectively (S4). Each MM/GMM maintains two instances. One is for the location update, whereas the other is for remaining functions such as outgoing CS/PS service requests.

B. Domain Decoupling

Two domains are coupled at the RRC layer. We hence propose a domain decoupling module in RRC. It aims to eliminate the unnecessary interference (e.g., triggered events in S3, disrupted data service in S4, and modulation downgrade in S5) between domains. For the triggering events, one domain should not be constrained by the other. That is, when CSFB is triggered in the CS domain, it should perform $3G \rightarrow 4G$ switch when the call ends. If the switch condition is satisfied (e.g., 4G is available), switching is executed, but not blocked by any operation in the PS domain. To this end, the base station adds a CSFB tag to assist the followup inter-system switching. On the other hand, if actions in one domain would disrupt those in the other (S4), such actions should be taken only if it is the only choice. For instance, when both 2G and 3G are available, the serving cell should move the device to 3G. Handover to 2G is only performed when 3G is not available. To avoid modulation downgrade, the 3G RRC can decouple PS and CS services by assigning different channels. Consequently, PS and CS services can be transmitted with different modulation schemes. To ensure decoupling, we differentiate CS/PS traffic and independently assign radio resources.

C. Cross-System Coordination

Similar functions in different systems should be coordinated. The key is to 1) share the information with each other and 2) collaborate to enforce proper decisions. Specifically, the 4G EPS bearer and the 3G PDP context are equivalent. Both are critical to data services. Two systems should regulate proper transitions when the user device switches across 3G and 4G. We recommend that user device should not detach itself without transferring its active PDP context from 3G to 4G. Instead, the device should immediately activate its EPS bearer after $3G \rightarrow 4G$ switch. Seamless system migration can thus be ensured. In case of failures in one system, the other system should help to recover from them if possible. For example, in the second issue, 4G MME should not detach the user device upon the location update failure in 3G. It should recover its location update with 3G MSC on behalf of the device.

We prototype our solution and emulate its control features at three key components (user device, base station, and cellular gateway). Our prototype is based on our proof-of-concept 3G/4G protocol stack, since the operational stack is not accessible. We use an Android phone as the user device, and two commodity PCs as the base station and the gateway. More implementation and evaluation details are in [1].

VII. Related Work

Mobile networking has been an active research area in recent years. New findings on mobile network performance are reported, including the interplay between applications and the infrastructure [22], [23], TCP over cellular channels [24], mutual interference between data and voice [17], and misbehaviors in cellular operations [25]–[27], to name a few. Our work differs from these studies in two aspects. First, these studies focus on the data transfer in the data plane, while we study the protocol interactions in the control plane. Second, they study protocol at the end hosts only, while our study spans on both the end device and the cellular infrastructure.

Protocol verification has been investigated on the Internet protocols [8], [28]–[30]. Recent efforts seek to validate the correctness of packet forwarding and processing, to eliminate loops, blackholes and/or crashes. Various techniques have been proposed, including controller program validation with symbolic execution[31], data-plane validation [32]–[35], header space analysis [36], etc.. Different from these studies, our verification is on the signaling protocol interactions part.

In mobile networks, most individual protocols/functions have been formally modeled and studied. For example, process calculus is applied to verify the functional correctness of mobility support[37], [38]. Formal models are also constructed for cellular mutual authentication protocol, and used to uncover the security loopholes [39], [40]. Our work differs in both the studied problem and the proposed solution. We study the interactions between protocols, and propose two-phase verification.

VIII. Conclusion

Control-plane protocols in cellular networks are more complex than their counterparts over the Internet. They have to work in more diversified usage settings, e.g., between CS and PS domains, and across 3G and 4G systems. They also support additional functions, including mobility, data and carrier-grade voice, fine control over radio resource. Consequently, inter-protocol signaling is widespread along all three dimensions of cross-layer, cross-domain, and cross-system scenarios.

Three cellular-specific lessons are learnt. First, in the cross-layer case, the well-tested layering rule from the Internet should be strictly honored. If the lower layer does not provide certain functions, the higher layer has to do so, or to be prepared to work without those functions. Coupling inter-layer actions is not a good practice unless properly justified. Second, in the cross-domain case, signaling design should recognize the inter-domain difference. Treating domains identically seems to reduce design and operation complexity, but makes it simplistic and error prone. Third, in the cross-system case, failure messages can be shared and even acted between systems. However, it is better not to expose such failure-handling operations outside the system unless absolutely needed.

In a broader scope, research on control-plane protocols in cellular networks warrants more efforts. 3G/4G is a large-scale infrastructure on a par with the wired Internet. There is no competing wireless technology for universal coverage and wide-area mobility support on the horizon. Given such a critical system indispensable to smartphones and tablets, more research is needed. The control-plane research in cellular networks also complements its counterpart on the Internet. While the Internet seeks to enhance its control plane, the cellular system needs to simplify its signaling design. Both can benefit from each other in the process.

REFERENCES

[1] G.-H. Tu, Y. Li, C. Peng, C.-Y. Li, H. Wang, and S. Lu, "Control-plane protocol interactions in cellular networks," in *Proc. ACM SIGCOMM*, Aug. 2014, pp. 223–234.

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

[3] TS23.060: GPRS; Service Description; Stage 2, 3GPP, 2006.

[4] *Voice over LTE*, GSMA VoLTE initiative [Online]. Available: http://www.gsma.com/technicalprojects/volte

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

[6] Authoritative Guide to Advanced LTE Testing, Ixia [Online]. Available: http://www.ixiacom.com/sites/default/files/resources/whitepaper/Authoritative-Guide-to-Advanced-LTETesting.pdf

[7] TS 36.521–1: Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) Conformance Specification; Radio Transmission and Reception; Part 1: Conformance Testing, 3GPP, 2014.

[8] G. J. Holzmann, *Design and Validation of Computer Protocols*. Holmdel, NJ, USA: Bell Laboratories, 1991.

[9] C. Baier *et al., Principles of Model Checking*. Cambridge, MA, USA: MIT Press, 2008, vol. 26202649.

[10] A. Damnjanovic, V. Vanghi, and B. Vojcic, *The CDMA2000 System for Mobile Communications: 3G Wireless Evolution*. New York, NY, USA: Pearson Education, 2004.

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

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

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

[14] TS36.331: Radio Resource Control (RRC), 3GPP, 2012.

[15] TS34.108: Common Test Environments for User Equipment (UE); Conformance Testing, 3GPP, 2014.

[16] Qxdm Professional™ Qualcomm Extensible Diagnostic Monitor, Qualcomm. [Online]. Available: http://www.qualcomm.com/media/documents/qxdm-professional-qualcomm-extensible-diagnostic-monitor

[17] G. H. Tu, C. Peng, H. Wang, C. Y. Li, and S. Lu, "How voice calls affect data in operational LTE networks," in *Proc. ACM MobiCom*, 2013, pp. 87–98.

[18] *AT&T: 2015 will be the Year of Mass Market VoLTE*, [Online]. Available: http://www.fiercewireless.com

[19] *Verizon Details VoLTE Rollout Plans*, Verizon, May 2014 [Online]. Available: http://newscenter.verizon.com/corporate/news-articles/2014/05-20-verizon-announces-volte/

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

[21] TS 23.203: Policy and Charging Control Architecture, 3GPP, 2013.

[22] N. Balasubramanian, A. Balasubramanian, and A. Venkataramani, "Energy consumption in mobile phones: A measurement study and implications for network applications," in *Proc. IMC*, 2009, pp. 280–293.

[23] U. Javed, D. Han, R. Caceres, J. Pang, S. Seshan, and A. Varshavsky, "Predicting handoffs in 3G networks," *SIGOPS Oper. Syst. Rev.*, vol. 45, no. 3, pp. 65–70, Jan. 2012.

[24] J. Huang, F. Qian, Y. Guo, Y. Zhou, Q. Xu, Z. M. Mao, S. Sen, and O. Spatscheck, "An in-depth study of LTE: Effect of network protocol and application behavior on performance," in *Proc. ACM SIGCOMM'13*.

[25] C. Peng, G. Tu, C. Li, and S. Lu, "Can we pay for what we get in 3G data access?," in *Proc. ACM MobiCom*, Aug. 2012.

[26] C. Peng, C. Y. Li, G. H. Tu, S. Lu, and L. Zhang, "Mobile data charging: New attacks and countermeasures," in *Proc. CCS*, Oct. 2012.

[27] G. H. Tu *et al.*, "Accounting for roaming users on mobile data access: Issues and root causes," in *Proc. MobiSys*, 2013.

[28] M. Musuvathi and D. R. Engler, "Model checking large network protocol implementations," in *Proc. NSDI*, 2004.

[29] M. A. S. Smith, "Formal verification of communication protocols," in *FORTE*, 1996, pp. 129–144.

[30] B. T. Loo *et al.*, "Declarative routing: Extensible routing with declarative queries," in *Proc. ACM SIGCOMM'05*.

[31] M. Canini, D. Venzano, P. Peresini, D. Kostic, and J. Rexford, "A NICE way to test openflow applications," in *Proc. NSDI*, 2012.

[32] H. Mai *et al.*, "Debugging the data plane with anteater," *ACM SIGCOMM Computer Commun. Rev.*, vol. 41, no. 4, pp. 290–301, Oct. 2011.

[33] H. Zeng, S. Zhang, F. Ye, V. Jeyakumar, M. Ju, J. Liu, N. McKeown, and A. Vahdat, "Libra: divide and conquer to verify forwarding tables in huge networks," in *Proc. NSDI*, 2014.

[34] A. Khurshid, W. Zhou, M. Caesar, and P. Godfrey, "Veriflow: Verifying network-wide invariants in real time," *ACM SIGCOMM Computer Commun. Rev.*, vol. 42, no. 4, pp. 467–472, Sep. 2012.

[35] M. Dobrescu and K. Argyraki, "Software dataplane verification," in *Proc. NSDI'14*.

[36] P. Kazemian, G. Varghese, and N. McKeown, "Header space analysis: Static checking for networks," in *Proc. NSDI*, 2012.

[37] F. Orava and J. Parrow, "An algebraic verification of a mobile network," *Formal Aspects of Computing*, vol. 4, no. 6, pp. 497–543, Nov. 1992.

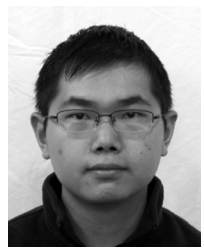
[38] G.-L. Ferrari, S. Gnesi, U. Montanari, and M. Pistore, "A model-checking verification environment for mobile processes," *ACM Trans. Software Eng. Methodol. (TOSEM)*, vol. 12, no. 4, pp. 440–473, 2003.

[39] C. Tang, "Modeling and analysis of mobile telephony protocols," Ph.D. dissertation, Stevens Inst. Technol., Hoboken, NJ, USA, 2013.

[40] TR33.902: Formal Analysis of the 3G Authentication Protocol, 3GPP, 2001.

Guan-Hua Tu received the Master degree in computer science from the University of California, Los Angeles (UCLA), CA, USA, in 2013. He is a 5th-year Ph.D. student in computer science at UCLA.

His research interests include wireless networking, network security, and mobile systems and applications with an emphasis on cellular data networks.



Yuanjie Li received the Bachelor degree in electronic engineering from Tsinghua University, China, in 2012. He is a 3rd-year Ph.D. student in computer science at the University of California, Los Angeles, CA, USA.

His research interests include mobile systems, cellular networks and data center networks.

Chunyi Peng (M'06) received the Ph.D. degree in computer science from the University of California, Los Angeles (UCLA), CA, USA, in 2013.

She now works as an Assistant Professor in the Department of Computer Science and Engineering at the Ohio State University, Columbus, OH, USA. Prior to UCLA, she worked as an Associate Researcher at Microsoft Research Asia. Her research interests focus on mobile networks, mobile sensing systems, wireless networking, and network security.

Chi-Yu Li received the Master degree in computer science from the University of California, Los Angeles, CA, USA, in 2013. He is currently a 6th-year Ph.D. student in the Department of Computer Science at UCLA.

His research interests include wireless networking, network security, and mobile systems and applications.

Songwu Lu is a Professor in the Computer Science Department at the University of California, Los Angeles, CA, USA. He has published 200+ technical papers, cited more than 20,000 times. His research interests cover wireless networking, mobile systems, sensor networks, and data center networking.

Prof. Lu is on the editorial board of IEEE/ACM TRANSACTIONS ON NETWORKING, and was on the boards of IEEE TRANSACTIONS ON MOBILE COMPUTING, *ACM Wireless Networks*, and *IEEE Wireless Communications Magazine* in the past.