vEPC-sec: Securing LTE Network Functions Virtualization on Public Cloud

Muhammad Taqi Raza, *Member, IEEE*, Songwu Lu, *Fellow, IEEE*, and Mario Gerla, *Fellow, IEEE*

***Abstract*—Public cloud offers economy of scale to adapt workload changes in an autonomic manner, maximizing the use of resources. Through network function virtualization (NFV), network operators can move LTE core to the cloud; hence removing their dependency on carrier-grade LTE network functions. Recent research efforts discuss performance, latency, and fault tolerance of LTE NFV, largely ignoring the security aspects. In this paper, we discover new vulnerabilities that LTE NFV face today with no standard solutions to address them. These vulnerabilities span at both LTE control and user planes. To address them, we propose `vEPC-sec` that cryptographically secures LTE control-plane signaling messages in the cloud. It provides distributed key management and key derivation schemes to derive shared-symmetric keys for securing the communication between any two network functions. Our approach provides encryption and integrity protection to the messages even during virtual machines scalability and failure recovery scenarios. `vEPC-sec` also prevents user-plane vulnerabilities by ensuring that LTE routing modules should faithfully forward the LTE subscriber packets.**

***Index Terms*—Security, network functions virtualization, 4G LTE, 5G networks, evolved packet core, fault tolerance, software defined networking.**

I. Introduction

LTE Network Function Virtualization (NFV) is a new trend that replaces carrier grade LTE core network functions with software running on commercial off-the-shelf servers in a cloud data center. On the one hand, NFV reduces operational and capital expenditure at traditional LTE network operators; on the other hand, it opens the cellular network business to small network operators. Network operators can take advantage of dynamic load balancing, the resource elasticity, and scalability that the cloud offers. This is a popular trend where a number of companies are offering multi-tenant LTE public cloud service following Amazon EC2 and Microsoft Azure style of business model [1]. In multi-tenant public cloud architecture, LTE network operators are cloud tenants that share compute, storage and network resources with each other.

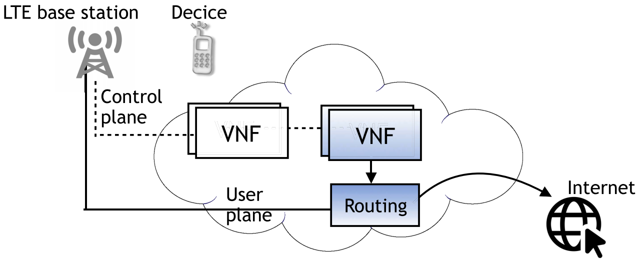


Fig. 1. LTE core NFs are moved to public cloud.

This fact has motivated us to study LTE core network security on the public cloud.

In our study, we find that available solutions provide detailed security guidelines to cryptographically secure both LTE signaling messages and data packets over the radio network [2]. They do not discuss, however, secure communication inside the LTE core network. Up till now, every network operator has privately operated its LTE packet core, shielding the backend packets processing and messages exchange from the outside world. In the age of multi-tenant LTE public cloud, LTE core traffic – not ciphered and transported as “clear text” – provides the adversary an opportunity to inspect subscriber traffic and to inject malicious network traffic.

Cloud service providers provision a number of virtual machines to host LTE Network Functions (NFs). These Virtualized Network Functions (VNFs) from different network operators share the same physical infrastructure. They communicate with LTE radio network via two different channels of control and user planes, as shown in Figure 1. Although cloud service providers logically isolate traffic from different tenants, they cannot guarantee that LTE VNF selection procedure always chains a VNF to the same tenant. This motivates an attacker to hijack VNF selection procedure to get associated with victim tenant’s network. After that, he gets control over the behavior of victim tenant’s VNFs.

We outline that an adversary can bring three different types of vulnerabilities and can launch a number of attacks in LTE NFV. First, the attacker can lie about the status of one-hop away neighbor and tricks the victim VNF to delete all associated subscribers records. Second, the adversary can put memory pressure by simply sending one false LTE paging notification message. This tricks LTE VNF to reserve memory space for tens of hundreds of devices and disrupts the memory resource allocation scheduling at victim VNF. Third, an adversary can inject fake IP packets into neighboring NF’s user-plane module. This renders victim NF’s data forwarding

Manuscript received August 7, 2018; revised January 22, 2019 and March 8, 2019; accepted March 9, 2019. Date of publication April 1, 2019; date of current version August 22, 2019. The work of M. T. Raza and S. Lu was supported by the NSF under Grant 1423576 and Grant 1526985. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Sherman S.-M. Chow. *(Corresponding author: Muhammad Taqi Raza.)*

M. T. Raza and S. Lu are with the Department of Computer Science, University of California at Los Angeles, Los Angeles, CA 90095 USA (e-mail: taqi@cs.ucla.edu; slu@cs.ucla.edu).

M. Gerla, deceased, was with the Computer Science Department, University of California at Los Angeles, Los Angeles, CA 90095 USA. M. Gerla’s memorial is at the end of the paper.

Digital Object Identifier 10.1109/TIFS.2019.2908800

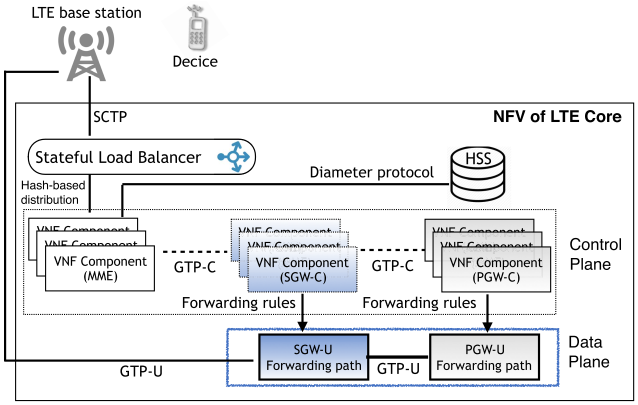


Fig. 2. LTE architecture over NFV: an overview.

module to process fake IP packets that impact the performance of other IP data packets flows.

To address these vulnerabilities, we put forward `vEPC-sec`. It is a solution that provides ciphering and integrity protection to LTE control-plane messages and prevents fake IP packets injection into user-plane. It provides shared-symmetric keys to VNFs to cryptographically protect their control-plane messages in a multi-tenant public cloud environment. `vEPC-sec` is designed to meet the cloud requirements of scalability and fault tolerance. The VNF might have lost the shared symmetric-keys during failure recovery or scaling to a new instance. `vEPC-sec` detects such a scenario through messages exchange with the peer VNF. It assigns a fresh shared-symmetric key to the recovered or scaled VNF. It also performs *key change on the fly* re-keying procedure with the VNFs whose keys need to be updated as well.

Our solution ensures that any attempt to inject fake IP packets should be detected and blocked. To achieve this, we add default packets forwarding policy as to ‘drop’ the packet. Further, through `vEPC-sec`, we can detect replaying of IP packets by malicious user-plane module that results in subscriber overbilling issue [3]. Our solution also identifies if data packets are illegally throttled at malicious forwarding module by delaying their delivery. To achieve these, our idea is to map radio data packets sent at LTE base station with the IP data packet received at LTE core. Because the LTE core network forwards the same packet that it has received from the base station, any missing/duplicate number of packets can be detected.

II. LTE–NFV in a Nutshell

LTE network consists of three main components: LTE device, LTE base station and LTE core, as shown in Figure 2. LTE NFV architecture virtualizes LTE core network functions over the cloud and eliminates reliance on vendor-specific proprietary hardware. Softwarization of LTE NFs accelerates the innovation by lowering operational and capital expenditures [4], [5]. LTE core (also known as Evolved Packet Core (EPC)) is composed of a number of Network Functions (NFs): the Serving Gateway (SGW), the PDN Gateway (PGW), the Mobility Management Entity (MME), the Home Subscriber Server (HSS), and a few others. These LTE EPC NFs (implemented as virtualized NF (VNFs) over cloud) handle control-plane and data-plane traffic through separate network interfaces and protocols. Cloud providers host EPC NFs on separate virtual machines (VMs) for scalability and flexibility purpose [6]–[8].

As shown in Figure 2, LTE control-plane traffic from radio network is sent to MME VNF. MME acts as a central management entity that authenticates and authorizes the device, handles device procedures (such as device registration, handover, location update, and service provisioning). It is also responsible for setting-up device data channel (i.e., data bearers) with SGW and PGW VNFs. In a virtualized environment, both SGW and PGW are divided into control-plane and user-plane modules. The control-plane modules are responsible of assigning IP address(es) for device and creating packet forwarding rules. These packet forwarding rules are sent to corresponding SGW and PGW user-plane modules that enforce the data packets forwarding policy for that device. Such decoupling of SGW and PGW into control and user planes is important for LTE data service performance that allows data packets to be forwarded without going through the virtualization layer.

VNFs selection in LTE: LTE network operators require that the appropriate EPC VNFs are selected to serve their subscribers according to device geographical area (known as tracking area in LTE), and type of radio network (macro/micro base station) it uses. To achieve this, network operators configure a number of EPC VNFs and create a pool of these VNFs. The best available VNFs – closer to the device and not heavily loaded – are selected to serve the subscriber device during its registration procedure with LTE network. This VNFs selection can be achieved either through stateful load balancer or through LTE standardized procedure [9]. In the first approach, the stateful load balancer sends a query to VNF pool database and gets the IP addresses of MME, SGW and PGW VNFs to serve the subscriber. This is a standard cloud-based approach implemented in today’s public clouds. Examples include, Microsoft Azure’s backend pool [10] and Amazon EC2 spot fleet [11]. In the second approach, configured LTE VNFs are registered at Domain Name Server (DNS). During device registration procedure MME VNF makes a DNS query to select best possible SGW and PGW VNFs instances to serve the subscriber. These DNS queries are made using UDP as transport protocol, as the standard states “DNS resolvers in EPC core network nodes shall support recursive queries and responses over UDP transport as specified in IETF RFC 1035” [9]. This selection of VNFs is vulnerable especially when query request/response are not cryptographically protected.

III. LTE Security Over Public Cloud

LTE standard secures device communication with LTE base station, and EPC through symmetric keys. On receiving device registration request, MME contacts HSS and retrieves the device symmetric session key (known as $\mathrm{K}_{\mathrm{ASME}}$ key). MME further derives separate ciphering and integrity keys to secure the device connection with radio network and LTE core [2]. MME secures its communication with LTE base station and HSS through secure SCTP and diameter protocols, respectively [12], [13]. However, SGW communication with MME, PGW and LTE base station is carried through unsecured IP/UDP based GPRS Tunneling protocol (GTPs). LTE system security and LTE network domain security standards do not discuss securing GTP control and user plane protocols [2]. In this paper, we first show new LTE vulnerabilities that unsecured GTP protocols bring in public cloud and then provide a framework to secure GTP protocols communication.

Threat model: Our threat model is similar to [14] and [15] where we consider a cloud service provider that hosts multiple LTE network operators (i.e. tenants). These tenants serve in a competitive LTE market where multiple tenants compete by providing LTE service in the similar geographical areas. To gain a competitive edge, a malicious tenant has a benefit to attack other tenants' LTE VNFs. The first step the malicious tenant takes is to trick victim tenant's VNF to get associated with one of the malicious tenant's VNF. By doing so, it gets control over the behavior of victim tenant. This is challenging, especially, when cloud service provider isolates traffic from different tenants through virtual LANs and/or source/destination addresses hash based forwarding. To solve this challenge, the malicious tenant exploits the fact that a VNFs selection query is made to select the best available VNFs during device registration procedure (§II). The malicious tenant can hijack the response of that query by replacing victim tenant's SGW IP address to one of its SGW IP address. He does not need to hijack every VNFs selection response; rather, hijacking one out of few thousands of responses is sufficient. Further, the malicious tenant can also control the number of VNFs selection requests. It can do so by first becoming the customer of victim tenant (by purchasing LTE service plan under victim tenant), and then sending a number of device registration requests to trigger SGW selection procedure at the cloud.

The malicious SGW that associates itself with the victim tenant network strictly follows LTE standard operations to avoid being detected through cloud intrusion detection box. The threats it can bring include: (a) sending wrong status information to MME regarding PGW, (b) sniffing unprotected GTP messages exchanged between source and destination MMEs, (c) putting memory pressure through false paging notification message(s), and (d) injecting fake IP packets to impact the performance of other IP flows. We assume that victim tenant VNFs are not compromised and function according to LTE standard protocols.

IV. LTE–NFV Vulnerabilities

A. Purging Subscribers' Context From MME

The malicious SGW can remove all subscribers' context from MME by sending PGW restart notification message. LTE EPC NFs employ a mechanism, known as *path management* [16], in which the availability of directly connected peer NFs can be determined for reliability purpose. The NF sends *echo request* message to its peer NF, and on receiving the *echo response* message it determines the reachability of peer NF. These periodic heartbeat messages are also used to adjust the retry timer value for lost signaling messages. Once an NF is detected to be non-responding (i.e., no *echo response* message is received for a certain number of tries), it is marked as failed. The failure indication is also sent to next hop NFs which are not directly connected with non-responding NF. We take an example of PGW failure. The PGW is directly connected to SGW, and its connection with MME goes through SGW. When SGW determines that the PGW has failed, it sends failure notification signaling message to MME (refer to section 7.9.5 PGW Restart Notification in TS 29.274 [17], and 16.1A.2 PGW Failure in TS 23.007 [18] for detailed procedure). On receiving the PGW failure notification, MME clears all those subscribers records which are served by failed PGW. MME then sends *Implicit Detach Request* message to all these subscriber devices. On receiving *Implicit Detach Request*, devices first locally deregister from LTE network and then re-initiate the registration procedure (i.e., Attach Request procedure). As new registration requests (i.e., *Attach Request* messages) from these subscriber devices arrive at EPC, a different PGW is selected (either by stateful load balancer or through DNS resolution).

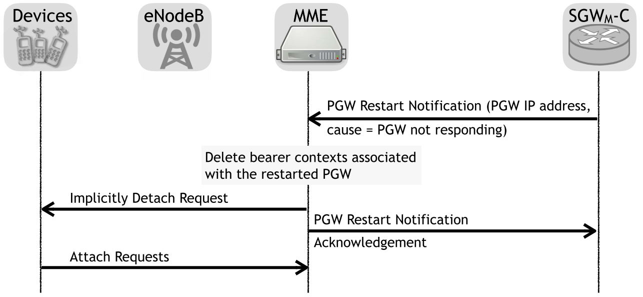

Devices
eNodeB
MME
SGW$_M$-C
PGW Restart Notification (PGW IP address,
cause = PGW not responding)
Delete bearer contexts associated
with the restarted PGW
Implicitly Detach Request
PGW Restart Notification
Acknowledgement
Attach Requests

Fig. 3. Malicious SGW (SGW$_M$) tricks MME to delete all subscribers records by sending false PGW restart notification message.

Malicious SGW adopts LTE failure recovery procedure in its advantage. It sends *PGW Restart Notification* message to MME, as shown in Figure 3. On receiving *PGW Restart Notification* message, MME sends *Implicit Detach Request* message to all those subscribers which are connected to the reported PGW. Thereafter, all these subscribers will send *Attach Request* message to MME. MME will authenticate these devices and will assign them new PGW that will assign IP addresses to these subscribers.

This vulnerability is quite powerful in two aspects. First, when malicious SGW reports PGW failure to MME, then this failure is cascaded to other SGWs too, as MME clears all subscribers' contexts related to the reported PGW. LTE design choice of associating multiple SGWs with one MME and a PGW is due to avoid IP address[1] change during device mobility. When the device moves around, the SGW is relocated by keeping the PGW unchanged, hence the device keeps the same IP address. Second, this vulnerability brings incast micro-burst (signaling spikes) at the stateful load balancer and MME VNF that may render them non-response for a short period of time, as shown in Figure 4. This is because on receiving the *Implicit*

[1] PGW assigns IP address(s) to every subscriber device.

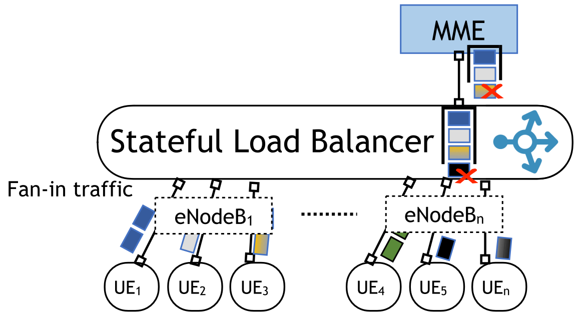


Fig. 4. Incast micro-burst problem: simultaneous initialization of attach request procedure from a quite number of devices (i.e., UEs) bring signaling spikes at cloud.

Detach Request message, all devices (associated with different LTE base stations but one MME VNF[2]) initiate the LTE Attach procedure at roughly the same time. These signaling messages, arriving from distributed LTE base stations (i.e., eNodeBs), are received by a stateful load balancer that forwards to MME VNF (approximately at same messages arrival rate).

Vulnerability 1: Malicious SGW can disrupt LTE service provided by victim tenant. The exploit of this vulnerability can be cascaded to other victim's SGWs, as MME clears device sessions spanned over multiple SGWs. Also, this vulnerability inadvertently creates micro-burst at the cloud when distributed LTE devices try to re-attach with MME.

B. Memory Pressure Through a False Downlink Notification Message

A false downlink notification message from malicious SGW VNF renders MME VNF to reserve the memory for hundreds of devices. The downlink notification message enables the device to re-establish the session with LTE network that it has been torn down while entering into low power idle state. The device enters into idle state when it has no data to send or receive. In the idle state, the device releases its radio connection with eNodeB to conserve the battery. The device periodically listens to the broadcast paging message (which is a downlink notification message for device) to check if there is any incoming data waiting to be transmitted at EPC. The paging message is initiated by MME when it receives a downlink data notification signal for a particular device from SGW. On receiving the paging message, the device establishes the radio connection with eNodeB followed by *Service Request* initial NAS message. On receiving the *Service Request* message from the device, the MME authenticates the subscriber and modifies the data bearer at SGW and PGW.

An attacker can exploit this LTE feature to put memory pressure at MME VNF. In his approach, the malicious SGW VNF sends a false downlink paging notification message to MME VNF, as shown in Figure 5. In that message, it puts the message cause as paging message and provides bearer identities for up to one thousand devices.[3] On receiving the paging message notification from SGW VNF, MME VNF reserves the memory for every device addressed in the downlink notification message. It then initiates the paging procedure through eNodeBs[4] that send broadcast paging messages addressing multiple devices. On receiving the paging message, the subscriber devices first initiate the radio connection with eNodeB and transition into connected state. They send *Service Request* messages to MME VNF to establish their data channel. MME VNF authenticates these devices and establishes their bearer resources for uplink/downlink transfer of data packets. Because these *Service Request* messages were initiated due to false downlink notification from malicious SGW VNF, there exists no data activity from/to devices. The MME awaits for device inactivity timer to expire (usually set as 11-12 seconds [19]) before releasing the connections for devices, and hence clearing the memory.

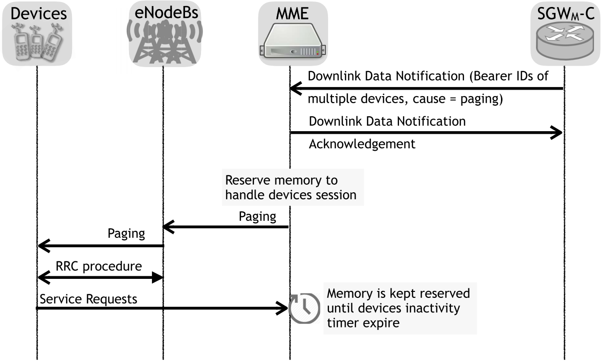

Devices
eNodeBs
MME
SGW$_M$-C
Downlink Data Notification (Bearer IDs of multiple devices, cause = paging)
Downlink Data Notification Acknowledgement
Reserve memory to handle devices session
Paging
Paging
RRC procedure
Service Requests
Memory is kept reserved until devices inactivity timer expire

Fig. 5. Malicious SGW VNF generates a paging downlink notification message addressing a number of devices towards victim's MME VNF. The MME VNF reserves the memory and initiates the paging procedure towards target devices. Through this procedure, malicious SGW VNF can increase memory usage at victim's MME VNF.

Through this vulnerability, malicious SGW VNF can keep MME VNF memory occupied by periodically (at an interval of 12 seconds) sending a downlink notification message. As a result, the attacker can slow down messages processing at victim VNF and incur control-plane latencies [20].

Vulnerability 2: False downlink data notification signaling message puts memory pressure on MME VNF that lets victim VNF reserve memory space for a number of devices. It also impacts victim tenant's subscribers' devices that end up consuming significantly higher battery power.

C. Slowing GTP Forwarding Plane by Injecting Fake IP Packets

We find that a malicious SGW VNF can throttle the victim tenant's user-plane traffic by simply sending fake IP packets to the victim tenant's forwarding plane. On device registration, once the user is authenticated and authorized by MME, PGW control plane assigns the IP addresses and packet forwarding precedence priority.[5] It also disseminates these policies to SGW control plane VNF. Both SGW and PGW apply IP

[2] During device mobility, the device changes LTE base station by performing X2 handover, but keeps same MME [8].

[3] One GTP-C payload message size is 64KB, whereas device identifier length is 64 bytes. LTE standard allows SGW to include multiple devices bearer identities in single downlink notification message [17]

[4] Paging message is sent in registered tracking area of the device. This tracking area spans over multiple eNodeBs.

[5] For example, IP address assigned for voice traffic has higher packet forwarding priority than default IP address assigned to access the Internet.

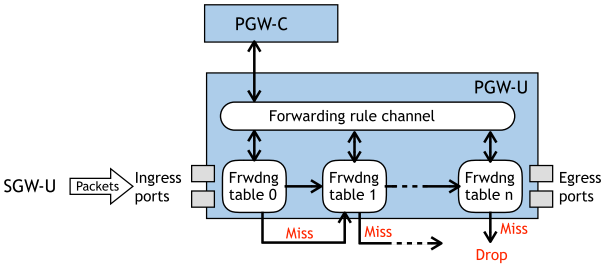


Fig. 6. Packet forwarding tables at SGW-U and PGW-U are arranged from highest priority rules to lowest priority ones. The malicious SGW-U can inject fake packets towards PGW-U that force PGW-U to search the rule at all of its forwarding tables. This procedure slows down the packet forwarding of legitimate IP packets.

packet forwarding rules at their user-plane forwarding engine, as *<rule, action>* tuples. The rule represents the matching of the different packets according to policy, and the action refers to the basic operation to be carried out over the incoming packets. Much like OpenFlow switching tables [21] and Service Data Flow Templates in LTE (Figure 6.5 in LTE policy and charging control architecture specification [22]), these rules are installed in forwarding tables of SGW-U and PGW-U, as shown in Figure 6. The tables closer to ingress ports store high precedence rules compared to the tables which are closer to egress ports. If the incoming packet does not match any rule at all the tables, it is dropped.

An attacker (SGW-C VNF) exploits the fact that incoming packet rule is searched at all forwarding tables before taking the action of dropping the packet. It first installs few fake IP packets forwarding rules as the highest precedence at its user-plane and then starts injecting these packets. When the fake IP packets arrive SGW-U from SGW-C, they are matched at 1st forwarding table and are sent to PGW-U. The PGW-U does not contain any entry of these fake IP packets as these IP addresses were never assigned by PGW-C. However, PGW-U needs to find the match of all fake IP packets at all of its forwarding table before discarding these packets. This process of matching of IP packets at all forwarding tables introduces extra packet-processing overhead that slows down other legitimate IP packet flows sharing the common hardware resource.

Vulnerability 3: Injecting fake IP packets slow down the forwarding plane performance of victim tenant.

V. Solution Goals and Overview

Goals: We want to achieve the following two goals in our solution.

1) GTP-C ciphering and authenticity: We want to cryptographically secure GTP-C communication.
2) GTP-U faithful packets forwarding: We want to ensure that the filtered packets reach PGW-U from SGW-U. Moreover, SGW-U should not be able to replay or delay packets.

Architecture and `vEPC-sec` *component:* Figure 7 provides an overview of our architecture. We propose a distributed architecture in which an LTE–NFV over the cloud is decomposed into several LTE–NFV subnets. Dividing vEPC into subnets ensure fault-tolerant and scalable network design.

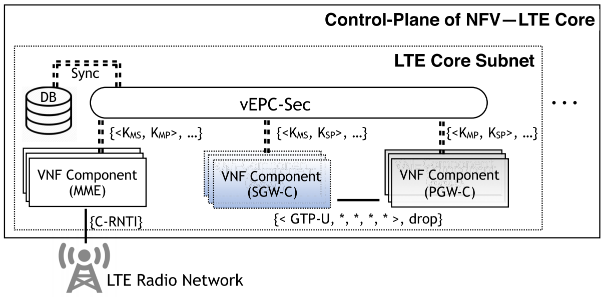

Control-Plane of NFV—LTE Core
LTE Core Subnet
Sync
DB
vEPC-Sec
• • •
{<K_{MS}, K_{MP}>, ...}
{<K_{MS}, K_{SP}>, ...}
{<K_{MP}, K_{SP}>, ...}
VNF Component (MME)
VNF Component (SGW-C)
VNF Component (PGW-C)
{C-RNTI}
{< GTP-U, *, *, *, * >, drop}
LTE Radio Network

Fig. 7. Our solution architecture to cryptographically secure GTP-C. Interfaces shown in double dotted lines are secured through TLS.

Our solution introduces `vEPC-sec` component, a central entity for providing key management to GTP-C traffic. It also ensures that only legitimate data packets are forwarded from SGW-U to PGW-U. We assume that `vEPC-sec` component is highly reliable with 1:1 redundancy [23], and communicates with EPC VNFs over secure channels only.

Solution overview: We propose `vEPC-sec` that (1) cryptographically secures communication over GTP-C, and (2) prevents illegitimate packets injection at GTP-U.

At GTP-C, our idea is to provide a distributed key management scheme from which LTE EPC VNFs derive ciphering and integrity keys to encrypt and integrity protect their messages. When EPC VNF is selected to serve the subscriber, it connects with `vEPC-sec` over a secure interface (shown as double-dotted lines in Figure 7) and requests the shared symmetric keys to communicate with other EPC VNFs. In the request message, it includes the VNF identities with which it wants to communicate, as well as its own identity. `vEPC-sec` first checks whether all these virtualized EPC (vEPC) instances are part of the same tenant by contacting local database. If the answer is positive, then it runs Key Derivation Function (KDF) and generates 3 pairs of keys so that MME, SGW and PGW VNFs can independently communicate with each other. `vEPC-sec` then sends these keys to corresponding EPC VNFs. Every VNF then locally derives integrity and ciphering keys against both keys it has received from `vEPC-sec`. Thereafter, the signaling messages between a pair of VNFs are ciphered and integrity protected. Our solution addresses **vulnerability 1** (§IV-A) and **vulnerability 2** (§IV-B) when MME only accepts integrity protected and ciphered messages from PGW-U (sent via SGW-U) using derived shared keys between MME and PGW VNFs. Therefore, SGW-U cannot lie that the message is originated from PGW-U.

At GTP-U, we introduce the concepts of assigning SGW-U the role of firewall, and correlating data packets received at LTE radio network and PGW-U. SGW-U plays the role of a firewall when PGW-C assigns the default packet forwarding policy to drop the packet. As a result, SGW-U only allows those IP packets whose addresses are assigned by PGW-C (via SGW-C); hence addressing **vulnerability 3** (§IV-C). We further ensure that only those packets should reach PGW-U which are sent by the legitimate subscriber. That is, SGW-U should not be able to replay IP packets towards PGW-U. We achieve this by matching packets sent from LTE base station to vEPC-sec, and packets received at PGW-U from SGW-U. By correlating packets sequence numbers from both entities ensure that they were originated by the device and were not delayed/dropped by SGW-U. vEPC-sec also correlates IP address with device Cell Radio Network Temporary Identifier (C-RNTI).[6] C-RNTI and IP address mapping confirm that IP packets are originated by the legitimate device, hence avoid IP spoofing attack reported in LTE [3].

VI. Solution

A. LTE GTP-C Confidentiality and Integrity Protection

We propose distributed security keys derivation and management scheme for integrity protection and ciphering of GTP-C signaling messages.

Distributed security keys derivation and management for GTP–C: Our solution provides a security abstraction module vEPC-sec, responsible of providing symmetric keys to LTE VNFs. When a VNF is chosen to serve a subscriber (during device registration procedure), it first checks whether it has the symmetric keys to securely communicate with other selected VNFs or not. If the keys exist then the subscriber signaling messages exchanged between these VNFs are ciphered and integrity protected; otherwise, shared symmetric keys are retrieved from vEPC-sec over TLS connection.[7] LTE VNF retrieves the keys by sending *Keys Information Request* message, requesting security keys for GTP-C communication. This request includes its VNF identity (which is Universal Unique Identifier (UUID) assigned to VM [24]), as well as identities of other VNFs with which it will communicate. Upon the receipt of the *Keys Information Request* message from the LTE VNF, vEPC-sec contacts the database and determines whether all these VNFs belong to the same operator or not. If the response is negative, then an alarm message will be sent to NFV orchestrator to take further action. This is the first line of defense in which any attempt from malicious tenant to infiltrate into victim tenant network is thwarted. In case all VNFs included in *Keys Information Request* message belong to the same tenant, vEPC-sec computes K_{MS}, K_{MP}, and K_{SP} to secure the communication between MME – SGW, MME – PGW, and SGW – PGW VNFs, respectively. Each key is derived from the KDF by using inputs of 256 bits long vEPC-sec master key and RAND value, as well as identities of two VNFs with which the key will be shared. We have shown the keys derivation steps in Figure 8. vEPC-sec applies the H-MAC based KDF, as specified in TS33.220 3GPP specification [25].

After deriving the keys, vEPC-sec sends *Keys Information Response* message back to the VNF that has requested the keys. It also sends *Keys Allocation Request* message to other two VNFs for whom the keys were derived in the process. These messages contain two keys required to communicate with other two EPC VNFs as well as encryption and integrity algorithm identities[8] for further key derivation. On receiving the message from vEPC-sec, every VNF further derives ciphering and integrity keys. It inputs encryption algorithm identity and the received key value to derive the ciphering key. Similarly, it inputs integrity algorithm identity along with the received security key and derives the integrity key. We have shown keys derivation at VNFs in Figure 8. Once both ciphering and integrity keys have been derived, these are truncated and the 128 least significant bits are used. Thereafter, VNF can use these keys to send ciphered and integrity protected messages over GTP-C interface to other paired VNFs.

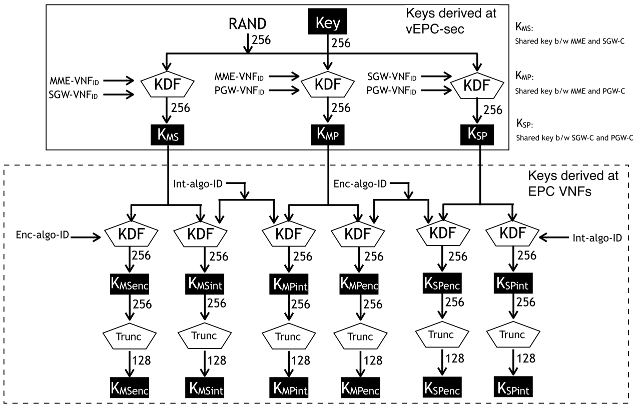


Fig. 8. Keys hierarchy and derivation for securing communication at LTE GTP-C interfaces.

Derivation of MAC, and ciphering the messages: After integrity and encryption keys derivation, we discuss how our solution encrypts messages and calculates their Message Authentication Code (MAC) for integrity check. Let's assume sender wants to send a message to receiver VNF over GTP-C. The sender VNF first calculates the MAC through EPS integrity algorithm [2]. The algorithm takes a number of input parameters: (1) 128 bit shared symmetric key between sender and receiver VNFs, (2) a 32-bit Nonce, (3) 1-bit direction of the transmission, and (4) the GTP-C signaling message itself. The Nonce value is a pseudo-random number to ensure that old messages cannot be replayed. The direction bit is 0 for uplink and 1 for downlink message. After calculating MAC, the sender then ciphers the message by using EPS encryption algorithm [2]. The input values to the algorithm are: (1) 128 bit shared symmetric key between sender and receiver VNFs, (2) a 32-bit Nonce, (3) 1-bit direction of the transmission, and (4) the length of the GTP-C signaling message to be sent. The encryption algorithm outputs keystream block equals to the length of the message. The message is then encrypted using a bit per bit binary addition of the plaintext GTP-C message and the keystream block. The sender sends the encrypted message, MAC and the Nonce value to the receiver. The receiver first ensures that the Nonce value is not the one it has received before. The receiver then calculates the MAC and matches it with received MAC value to ensure the integrity protection of the message. If the

[6] C-RNTI uniquely identifies the device over the air. This C-RNTI remains unchanged until the device releases its radio connection (i.e., RRC Connection Release).

[7] LTE VNF and vEPC-sec interface is protected using TLS. No message is exchanged until a secure tunnel is established

[8] vEPC-sec selects either 00010 or 0010 to represent AES or SNOW 3G algorithm identity, respectively [2], in its response message.

integrity check is passed, the receiver deciphers the encrypted message. The receiver recovers the message by generating the same keystream using the same input parameters by the sender and applying a bit per bit binary addition with the ciphertext.

Securing communication during device mobility: Up till this end, vEPC-sec secures the communication between MME, SGW and PGW VNFs. Due to device mobility, $\text{MME}_{\text{source}}$ needs to exchange handover signaling messages as well as providing K_{ASME} key to $\text{MME}_{\text{target}}$. To meet the security requirement in device mobility, we extend our key management technique for communication between two MME VNFs. On receiving the *Handover Required* message from LTE base station, $\text{MME}_{\text{source}}$ determines the address of $\text{MME}_{\text{target}}$ and asks vEPC-sec to provide the shared symmetric key to securely communicate with $\text{MME}_{\text{target}}$. vEPC-sec generates the K_{MM} and gives it to $\text{MME}_{\text{source}}$ along with integrity and ciphering algorithm identities. It also sends a message *Handover Key Establishment* to $\text{MME}_{\text{target}}$ that include K_{MM} and integrity and ciphering algorithm identities. The message from vEPC-sec explicitly informs $\text{MME}_{\text{target}}$ that it would receive a ciphered and integrity protected message from $\text{MME}_{\text{source}}$ which will be decoded using the provided key. On receiving the K_{MM}, $\text{MME}_{\text{source}}$ proceeds with handover procedure and sends ciphered and integrity protected handover signaling message to $\text{MME}_{\text{target}}$. The $\text{MME}_{\text{target}}$ receives the message from $\text{MME}_{\text{source}}$ VNF for which it has received the key, and deciphers the message after ensuring the message integrity check.

B. LTE GTP-U Faithful Packets Forwarding

Our solution ensures that SGW-U (1) does not inject any fake packets, (2) forwards the data packets without delaying, and (3) does not duplicate the packet forwarding. Moreover, vEPC-sec also prevents IP packets spoofing by attacker devices.

Making SGW-U the firewall for PGW-U: At the time of device registration, PGW-C VNF assigns the device IP address(es) and applies packet forwarding rules – that it receives from policy and charging LTE NF – at PGW-U. PGW-C also forwards the <*rule, action*> pair to SGW-C. SGW-C then installs these rules to its forwarding plane. This means both SGW-U and PGW-U apply identical packets forwarding rules. The SGW-U which receives the device data packets from radio network simply forwards these packets to PGW-U according to data forwarding policy. The vulnerability we discuss in §IV-C arises due to the fact that SGW-U forwards the fake IP packets and exhausts the forwarding table lookup at PGW-U. In principle, SGW-U should never forward a packet whose rule was not defined by PGW-C. It means there exists no practical use case scenario in which SGW-U and PGW-U forwarding policies ever mismatch. PGW-C uses this principle to address **vulnerability 3** (§IV-C). It simply explicitly provides packet drop policy to SGW-C when no rule is found. It sends <*GTP-U * * * *, drop*>[9] rule signifying that the default rule is to drop the packet. We should mention that the default rule must be placed as the last entry of the last table at SGW-U, otherwise legitimate packets will be dropped. This is a common misconfiguration issue reported in packet forwarding middleboxes (liked routers and switches) [26].

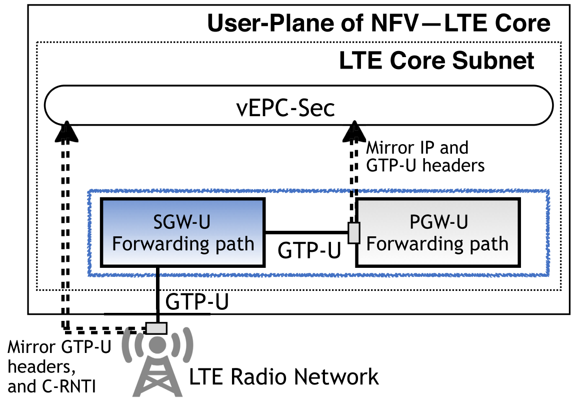


Fig. 9. Our solution to guard GTP-U traffic.

Ensuring data packets are not maliciously throttled: Although the above solution addresses fake IP packets injection problem, SGW-U can misbehave in a different way. It can replay legitimate data packets to overbill[10] the subscriber [3], and can even delay the data packets forwarding to throttle the end user data throughput. Note that, periodically delaying some TCP packets cause out of order delivery at the receiver. As a result, the TCP at the send side keeps transitioning between fast retransmit/fast recovery and congestion avoidance phases. The application data sending rate is throttled as a consequence.

To address these issues, our approach is to enable data packets' headers inspection at vEPC-sec. When the device has some data to send, it establishes its data channel with LTE network (i.e., by sending *Service Request* message). When LTE base station receives the data packets from the device, it puts GTP-U header that includes a message type, GTP-U tunnel identifier, and the packet sequence number. The sequence number field uniquely identifies the packet in an IP flow at LTE network. The value is incremented on every data packet transmission. Our idea is to enable 1:1 mapping between packets sent by LTE base station to SGW-U and the ones received by PGW-U from SGW-U. This approach isolates the malicious activity done at SGW-U. As shown in Figure 9, we require that LTE base station should mirror its interface towards SGW-U to vEPC-sec. The mirrored packets are only the GTP-U headers and the device radio network identity (i.e. C-RNTI), and does not include packets payload. Similarly, PGW-U mirrors the IP and GTP-U headers of the packets that it has received from SGW-U to vEPC-sec. By looking at same headers reported from two different entities, vEPC-sec can distinguish any missing packets, out of order packets, and even duplicate packets.

Ensuring data packets are originated by legitimate device: By correlating C-RNTI with IP packets, vEPC-sec can also avoid other attacks that have been reported in recent past.

[9] Should be read as: packet from GTP-U protocol with any source address, any source port, any destination address, and any destination port will be dropped.

[10] The subscriber pays for the data packets it has sent/received at PGW-U. So SGW-U can replay subscriber data packets to overcharge the subscriber.

These include spoofing of IP packets and injecting data packets by using the IP address of control-plane [3], [27], [28]. These attack mainly occur when the malicious device which is authenticated during the connection establishment phase may lie about itself while sending IP packets. To address this issue, `vEPC-sec` binds the C-RNTI with the IP address the P-GW has assigned. In this way, the attacker can only send IP data by using its own data-plane IP address. He can neither use control-plane IP address or spoof IP address of other subscribers.

C. Discussion

We briefly discuss how our solution works during VNFs failure recovery and scalability scenarios.

Fault tolerance and scalability: LTE network operators aim to provide all-time service access to their subscribers. Both cloud service providers and LTE standard discuss failure recovery and scalability procedures [29]. In the failure recovery procedure, a standby EPC VNF replaces the failed VNF, and failed signaling messages are re-executed. It is possible that during the recovery process the alternative VNF cannot restore the GTP-C security keys (e.g. in fail-stop failure scenario). Similarly, scalability requirements stipulates that a new VNF instance should be prepared to handle increasing subscribers requests. This new VNF does not have the security keys to communicate with its peer VNFs.

To address these challenges, we propose that once the new VNF instance becomes active, it first contacts `vEPC-sec` by sending *Keys Information Request due to Recovery and Scalability* message. In this message, it includes its own VNF identity, and the VNF identity of the failed instance – in case of failure, or the identity of the original VNF that is being scaled. On receiving the request, `vEPC-sec` first determines all those VNF instances that have been affected due to failure/ scalability. It initiates "key change on the fly" procedure by sending *Re-keying Required* message to all affected VNFs. Once these VNFs receive *Re-keying Required* message from `vEPC-sec`, they suspend their communication and prepare to change the key by responding with *Re-keying Request Acknowledged* message. On receiving the *Re-keying Request Acknowledged* message from all these VNFs, `vEPC-sec` derives and distributes new security keys, according to the procedure discussed in §VI-A. In this way, the new VNF, as well as other affected VNFs can resume secure messages exchange over GTP-C interface.

VII. Security Analysis

We briefly discuss why available cloud security mechanisms do not protect from LTE NFV security vulnerabilities. Later, we provide the security analysis of `vEPC-sec`.

A. Limitations of Cloud Security Solutions

In the cloud, packets exchange can be cryptographically secure either by using TLS at the transport layer, or Internet Protocol security (IPsec) at the networking layer of the protocol stack. Firewall, Intrusion Detection Systems (IDS), and Intrusion Prevention Systems (IPS) middleboxes block unauthorized access to tenant VNFs, detect and prevent the malicious activities at these VNFs, respectively. We find that all these mechanisms are not sufficient to provide LTE NFV security.

GTP uses UDP/IP protocol to transfer GTP messages [16] and cannot use transport layer security mechanisms. GTP tunnels can use IPsec to secure their messages. However, IPsec does not meet high availability and fault tolerance LTE requirements [30]. It takes more than 15 seconds to re-establish new signaling bearers with the subscribers (which were served by the failed VNF instance) [30]. This delay is 18× more than LTE high availability requirement of five-nines[11] (i.e., VNF downtime should not be greater than 864.3 milliseconds per day). Moreover, IPsec does not protect against fake IP packets injection vulnerability (§IV-C).

Security middleboxes in the cloud have their own limitations. Firewall performs stateful inspection of GTP traffic entering into the tenant network. Its functions protect the mobile packet core from signaling storms and man in the middle attacks. However, it does not guard against insider attacks when an adversary VNF becomes part of victim tenant's network. The purpose of IDS/IPS is to perform signature-based packets inspection to find malicious activities between LTE VNFs. They also fail to detect discussed vulnerabilities when malicious tenant VNF fully obeys the LTE standards to alter the functionalities provided by victim tenant.

B. `vEPC-sec` Security Analysis

On secure communication between malicious and victim tenant's VNFs: Our solution does not allow the malicious VNF to establish a secure GTP-C connection with victim VNFs. We consider an adversarial model in which an adversary can communicate with `vEPC-sec` to derive the keys. Although an adversary cannot sniff TLS protected packets between `vEPC-sec` and victim VNFs, it can get the VNFs identities through other means (e.g. sniffing the ARP packets and decoding UIUD from MAC address [24]). To understand, we take an example of malicious SGW that holds victim tenant's MME and PGW VNF UIUDs and describe it in Analysis 1 pseudocode. Malicious SGW first establishes the TLS connection with `vEPC-sec` and then sends the *Keys Information Request* message. In the message, it includes UIUDs of its VNF as well as victim MME and PGW VNFs. On receiving the key generation request, `vEPC-sec` first verifies whether all these VNFs belong to the same tenant or not. It contacts the cloud database to get an answer. The cloud database has the record of all UIUDs and has mapped these identities against the operator, location, priority and weight factor. It replies `vEPC-sec` with the operator names that host and manages these VNFs. On receiving the response, `vEPC-sec` determines that all three VNFs do not belong to the same operator and hence rejects the request by sending *Keys Information Request Rejected* message back to malicious SGW. It can mention the reject cause as: *different operators*. The malicious SGW can try all different

[11] Public cloud provides four nines of high availability, that is downtime of 8.64 seconds/day is allowed [7].

vEPC-sec Analysis 1 Adversary Tries to Receive Shared Symmetric Keys to Communicate With Victim VNFs

```
Assume an adversary can sniff all VNF identities of victim tenant;
Let SGW_M = Adversarial controlled SGW-C VNF identity;
Let MME_V[n] = VNF identities of victim tenant's MMEs;
Let PGW_V [n]= VNF identities of victim tenant's PGW-C;
for $i$ = 0 to MME_V[n] do
    for $j$ = 0 to PGW_V[n] do
        SendTovEPC-sec ($KeysInformationRequest$,
        SGW_M, MME_V[i], MME_V[j])
        if $KeysInformationResponse$ == TRUE then
            return 1; // Adversary wins
        else
            return 0;// Adversary loses
        end
    end
end
end
```

vEPC-sec Analysis 2 Adversary Tries to Misuse PGW-U Resources by Sending Fake IP Packets or Delaying Packets

```
Let $receiver$ = Adversarial controlled machine over the Internet;
if SendtoPGW-U(msg, FAKE_src_IP, dest_IP) == SUCCESS
&&
ReceivefromPGW-U(msg, FAKE_src_IP,
dest_IP)==SUCCESS then
    return 1; // Adversary wins
else
    return 0;// Adversary loses
end
end
```

combinations of MME and PGW identities and can send *Keys Information Request* as many times as it wants. Every time, its request will be rejected by vEPC-sec. Note that, we can improve the implementation of vEPC-sec by raising an alarm to NFV orchestrator that can take further action against a malicious tenant.

On injecting fake IP packets: Our solution detects the fake IP packets injection by SGW-U. In Analysis 2, we show that the adversary is allowed to inject fake IP packets which are against the policy provided by PGW-C. When these fake IP packets arrive at PGW-U they are marked as resource abuse attempt packets. As there exists no forwarding table entry against these fake IP packets. PGW-U then sends an alarm signal message to PGW-C that takes further action after contacting NFV orchestrator. We should point out, it is not possible for SGW-U to send the IP packets when it recovers from the failure. This is because that the lost data bearers are required to be re-established by SGW-C first, and data forwarding policies are installed afterward.

On illegal throttling of data packets: We show that an attacker cannot illegally throttle subscriber's data packets by delaying the packets forwarding. In our analysis, as shown in Analysis 3 pseudocode, the adversarial control attacker receives the packets from LTE base station and delays their forwarding to PGW-U. When the PGW-U receives the packets (both delayed and not delayed), it mirrors packets' headers to vEPC-sec (refer to Figure 9) before forwarding them to the Internet. vEPC-sec performs 1:1 mapping of packet sequence numbers that it has received from LTE base station and PGW-U. If the sequence numbers mismatch is consistently observed for a certain period of time (as the attacker periodically delays packets forwarding to achieve throughput throttling), vEPC-sec raises the alarm towards NFV orchestrator. NFV orchestrator then needs to identify whether the pause in data forwarding by SGW-U is intentional or not. If it is intentional then it takes an action against the malicious tenant; otherwise it replaces slow performing SGW-U instance.

vEPC-sec Analysis 3 Adversary Tries to Throttle the Victim Tenant's Subscriber Packets

```
Let $receiver$ = Adversarial controlled machine over the Internet;
while Certain TIME has not passed do
    if ReceivefromENB(msg, src_IP, dest_IP) == SUCCESS
    then
        $WAIT(timer)$; //wait for certain time before
        forwarding to PGW-U
        SendtoPGW-U(msg, src_IP, dest_IP);
        $sent\_count = sent\_count + 1$; //count packets sent
        by SGW-U
        ReceivefromPGW-U(msg, src_IP, dest_IP);
        $received\_count = received\_count + 1$; //count
        packets received at reciver
    end
end
if $sent\_count == received\_count$ then
    return 1; // Adversary wins
else
    return 0;// Adversary loses
end
end
```

Performance: We simulate to determine the performance of vEPC-sec. First, we determine how quickly our solution can detect the throttling of data packets. The challenge we face was to distinguish between slow performing SGW-U with the malicious one. To solve this challenge, we implement a sequence number window at vEPC-sec. Our sequence number window is linearly numbered. When the packet sequence number from LTE base station arrives, we put it in the window and wait for the packet sequence number from LTE PGW-U. On receiving the sequence number from PGW-U, the difference is calculated. If the difference is zero it means there is no packet delay. If the next packet sequence number has arrived from LTE base station while our window was waiting for a packet from PGW-U, we move the window. That is, we do not record the delayed packet. In this way, our final window has churn of readings representing packet

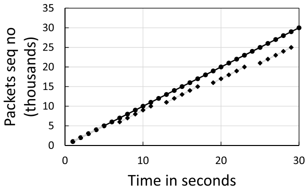


Fig. 10. Data packets throttling detection. The throttling of data means the packets sequence number skews from the linear line (above).

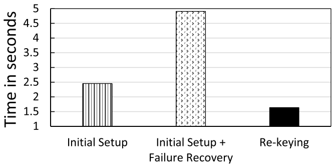


Fig. 11. Key management overhead.

sequence numbers. That is, once the window has skewed, then this skew will keep increasing over time. In this way, we detect the adversarial SGW-U that periodically delays the packets. From the Figure 10, we can see that in just 30 seconds, `vEPC-sec` can detect packets delaying malicious activity when the gap (representing lower dotted lines) has largely skewed away from the linear line.

In Figure 11, we show an overhead associated with KDF. We consider a machine with CPU of 2.5GHz and 3GB RAM. Our approach only causes one time overhead of 2.5 seconds. This overhead is also associated with the number of times the failure recovery has occurred. For every failure recovery procedure, `vEPC-sec` needs to generate fresh pairs of shared-symmetric keys. The re-keying process which explicitly asks VNFs to calculate the key has the lowest overhead. This is mainly due to the fact that these VNFs have to locally run KDF once.

VIII. Related Work

The most recent work on NFV security is SafeBricks, published last year [31]. It shields generic NFs from an untrusted cloud and proposes to encrypt all the traffic entering into the cloud. In our work, we did not provide extra encryption of already encrypted traffic (e.g., traffic ciphered by secure DNS and secure SCTP protocols). Rather, our focus is to cryptographically secure the unsecured LTE GTP-C traffic. Other works [32], [33] discuss security issues associated with multi-tenancy and live migration. References [36] and [37] use Intel Software Guard Extensions (Intel SGX) to securely isolate the states of NFV applications. [36], [37] unveils DDoS attack that comes from flexible and elastic resource provisioning in NFV. Contrary to all these works, this paper presents attacks which are unique to LTE operations. We show how an adversary by sending fake signaling messages can disrupt LTE service, and to be worse, no middlebox signature based vulnerability detection solution can detect these types of attacks. Further, all these previous works have not discussed attacks on user-plane, but our paper addresses.

A number of other works discuss LTE security issues. References [38] and [39] conduct LTE protocol vulnerability analysis and show real impacts on LTE subscribers. Reference [40] conducts experimental validation to prove that LTE temporary identity can disclose subscriber location. Reference [41] discusses privacy attacks in which signaling information is leveraged to infer user privacy information. Reference [42] shows that current cellular infrastructures exhibit security loopholes (off-path TCP hijacking) due to their NAT/firewall settings. References [3] and [43] study insecurity in mobile data charging. References [27] and [28] discuss how a subscriber can inject control-plane traffic into user-plane and can get free data service. Different to all the above works, we do not discuss security vulnerabilities originated by an adversarial device. Instead, we present first work that discusses the security issues arising from LTE core network implemented over the public cloud.

IX. Conclusion

We propose `vEPC-sec` that secures LTE NFV over the public cloud. It cryptographically protects LTE control-plane traffic on virtualized instances and enforces data forwarding policies at every forwarding module. `vEPC-sec` enables encryption and integrity protection in LTE core network through a distributed key management scheme. Its design ensures that communication between LTE NFs must be secure even during NF scalability and failure recovery scenarios. `vEPC-sec` provides lightweight data forwarding monitoring component that only checks one type of header from two different sources to identify whether subscriber packets were delayed or duplicated. The security analysis confirms that `vEPC-sec` shields LTE core network traffic from the adversarial model over the public cloud.

Memorial

This paper is dedicated to the last author, Mario Gerla, who was a Professor in the Computer Science department at UCLA. He has passed away on February 9, 2019, as a victim to pancreatic cancer. He was 75 years old.

References

[1] *HP: Hybrid Cloud Solutions for LTE NFV*. Accessed: Sep. 14, 2018. [Online]. Available: https://h20195.www2.hpe.com/V2/getpdf.aspx/4aa6-8334enw.pdf

[2] *SAE; Security Architecture*, document TS33.401: 3GPP, 3GPP, Sep. 2013.

[3] C. Peng, C. Y. Li, H. Wang, G. H. Tu, and S. Lu, “Real threats to your data bills: Security loopholes and defenses in mobile data charging,” in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2014, pp. 727–738.

[4] R. Mijumbi *et al.*, “Network function virtualization: State-of-the-art and research challenges,” *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 236–262, 1st Quart., 2016.

[5] *Bringing Network Function Virtualization to LTE*. Accessed: Mar. 7, 2019. [Online]. Available: http://www.5gamericas.org/files/1014/1653/1309/4G_Americas_-_NFV_to_LTE_-_November_2014_-_FINAL.pdf

[6] Z. A. Qazi, M. Walls, A. Panda, V. Sekar, S. Ratnasamy, and S. Shenker, “A high performance packet core for next generation cellular networks,” in *Proc. ACM SIGCOMM*, 2017, pp. 348–361.

[7] B. Nguyen *et al.*, “A reliable distributed cellular core network for public clouds,” Microsoft Res., Cambridge, U.K., Tech. Rep. MSR-TR-2018-4, 2018.

[8] M. T. Raza, D. Kim, K.-H. Kim, S. Lu, and M. Gerla, “Rethinking LTE network functions virtualization,” in *Proc. IEEE 25th Int. Conf. Netw. Protocols (ICNP)*, Oct. 2017, pp. 1–10.

[9] *Domain Name System Procedures, Release 13*, document TS29.303, 3GPP, Jun. 2016.

[10] *Azure Load Balancer*. Accessed: Jan. 20, 2019. [Online]. Available: https://docs.microsoft.com/en-us/azure/load-balancer/load-balancer-overview

[11] *Amazon EC2: Spot Fleet*. Accessed: Jan. 20, 2019. [Online]. Available: https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/spot-fleet.html

[12] *Secure SCTP*, document IETF draft-hohendorf-secure-sctp-25, 2019. [Online]. Available: https://datatracker.ietf.org/doc/draft-hohendorf-secure-sctp/

[13] V. Fajardo, J. Arkko, J. Loughney, and G. Zorn, *Diameter Base Protocol*, document RFC 6733, 2012.

[14] V. Varadarajan, Y. Zhang, T. Ristenpart, and M. Swift, "A placement vulnerability study in multi-tenant public clouds," in *Proc. USENIX Secur. Symp.*, 2015, pp. 913–928.

[15] Z. Xu, H. Wang, and Z. Wu, "A measurement study on co-residence threat inside the cloud," in *Proc. USENIX Secur. Symp.*, 2015, pp. 929–944.

[16] *General Packet Radio System (GPRS) Tunnelling Protocol User Plane (GTPv1-U)*, document TS29.281, 3GPP, 2013.

[17] *Tunnelling Protocol for Control Plane (GTPv2-C)*, document TS29.274, 3GPP, 2014.

[18] *LTE Restoration Procedures*, document TS23.007, 3GPP, 2014.

[19] J. Huang, F. Qian, A. Gerber, Z. M. Mao, S. Sen, and O. Spatscheck, "A close examination of performance and power characteristics of 4G LTE networks," in *Proc. 10th Int. Conf. Mobile Syst., Appl., Services*, 2012, pp. 225–238.

[20] A. Markuze, I. Smolyar, A. Morrison, and D. Tsafrir, "DAMN: Overhead-free IOMMU protection for networking," *Proc. 23rd ACM Int. Conf. Architectural Support Program. Lang. Oper. Syst.*, 2018, pp. 301–315.

[21] Y. Chiba, Y. Shinohara, and H. Shimonishi, "Source flow: Handling millions of flows on flow-based nodes," *ACM SIGCOMM*, vol. 41, no. 4, pp. 465–466, 2011.

[22] *Policy and Charging Control Architecture*, document TS 23.203, 3GPP, 2013.

[23] J. Sherry *et al.*, "Rollback-recovery for middleboxes," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 45, pp. 227–240, Oct. 2015.

[24] *VMware Workstation 5.0 Virtual Machine Identifier—UUID*. Accessed: Sep. 14, 2018. [Online]. Available: https://www.vmware.com/support/ws5/doc/ws_move_uuid.html

[25] *LTE Generic Authentication Architecture (GAA) and Generic Bootstrapping Architecture (GBA)*, document TS33.220, 3GPP, Sep. 2012.

[26] T. Fiebig *et al.* (2016). "SoK: An analysis of protocol design: Avoiding traps for implementation and deployment." [Online]. Available: https://arxiv.org/abs/1610.05531

[27] H. Kim *et al.*, "Breaking and Fixing VoLTE: Exploiting hidden data channels and Mis-implementations," in *Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Secur.*, 2015, pp. 328–339.

[28] C.-Y. Li *et al.*, "Insecurity of voice solution VoLTE in LTE mobile networks," in *Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Secur.*, 2015, pp. 316–327.

[29] *LTE Architecture Description*, document TS36.401, 3GPP, 2011.

[30] *LTE Security for Mobile Service Provider Networks*. Accessed: Mar. 7, 2019. [Online]. Available: https://www.juniper.net/us/en/local/pdf/whitepapers/2000536-en.pdf

[31] R. Poddar, C. Lan, R. A. Popa, and S. Ratnasamy, "SafeBricks: Shielding network functions in the cloud," in *Proc. 15th USENIX Symp. Netw. Syst. Design Implement. (NSDI)*, Renton, WA, USA, 2018, pp. 201–216.

[32] M. D. Firoozjaei, J. Jeong, H. Ko, and H. Kim, "Security challenges with network functions virtualization," *Future Gener. Comput. Syst.*, vol. 67, pp. 315–324, Feb. 2017.

[33] S.-J. Moon, V. Sekar, and M. K. Reiter, "Nomad: Mitigating arbitrary cloud side channels via provider-assisted migration," *Proc. 22nd ACM SIGSAC Conf. Comput. Commun. Secur.*, 2015, pp. 1595–1606.

[34] M.-W. Shih, M. Kumar, T. Kim, and A. Gavrilovska, "S-NFV: Securing NFV states by using SGX," in *Proc. ACM Int. Workshop Secur. Softw. Defined Netw. Netw. Function Virtualization*, 2016, pp. 45–48.

[35] H. Duan, C. Wang, X. Yuan, Y. Zhou, Q. Wang, and K. Ren. (2017). "LightBox: Full-stack protected stateful middlebox at lightning speed." [Online]. Available: https://arxiv.org/abs/1706.06261

[36] S. K. Fayaz, Y. Tobioka, V. Sekar, and M. Bailey, "Bohatei: Flexible and Elastic DDoS Defense," in *Proc. USENIX Secur. Symp.*, 2015, pp. 817–832.

[37] A. H. M. Jakaria, W. Yang, B. Rashidi, C. Fung, and M. A. Rahman, "VFence: A defense against distributed denial of service attacks using network function virtualization," in *Proc. IEEE 40th Annu. Comput. Softw. Appl. Conf. (COMPSAC)*, vol. 2, Jun. 2016, pp. 431–436.

[38] S. Hussain, O. Chowdhury, S. Mehnaz, and E. Bertino, "LTEInspector: A systematic approach for adversarial testing of 4G LTE," in *Proc. Symp. Netw. Distrib. Syst. Secur. (NDSS)*, 2018. doi: 10.14722/ndss.2018.23313.

[39] M. T. Raza, F. M. Anwar, and S. Lu, "Exposing LTE security weaknesses at protocol inter-layer, and inter-radio interactions," in *Proc. Int. Conf. Secur. Privacy Commun. Syst.* Cham, Switzerland: Springer, 2017, pp. 312–338.

[40] B. Hong, S. Bae, and Y. Kim, "GUTI reallocation demystified: Cellular location tracking with changing temporary identifier," in *Proc. Symp. Netw. Distrib. Syst. Secur. (NDSS)*, 2018, pp. 1–15.

[41] A. Shaik, R. Borgaonkar, N. Asokan, V. Niemi, and J.-P. Seifert, "Practical attacks against privacy and availability in 4G/LTE mobile communication systems," in *Proc. Netw. Distrib. Syst. Secur. (NDSS) Symp.*, 2015. doi: 10.14722/ndss.2016.23236.

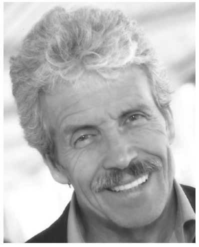
[42] Z. Qian and Z. M. Mao, "Off-path TCP sequence number inference attack—How firewall middleboxes reduce security," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2012, pp. 347–361.

[43] C. Peng, C.-Y. Li, G.-H. Tu, S. Lu, and L. Zhang, "Mobile data charging: New attacks and countermeasures," in *Proc. ACM Conf. Comput. Commun. Secur.*, 2012, pp. 195–204.

Muhammad Taqi Raza received the master's degree in computer science from the University of California at Los Angeles, Los Angeles, CA, USA, in 2017, where he is currently pursuing the Ph.D. degree with the Computer Science Department.

His research interests include network security, network function virtualization, and networked systems and their applications with an emphasis on cellular security.

Songwu Lu is currently a Professor with the Computer Science Department, University of California at Los Angeles, Los Angeles, CA, USA. His research interests cover wireless networking, mobile systems, sensor networks, and data center networking. He was on the Editorial Board of the IEEE/ACM TRANSACTIONS ON NETWORKING. He was on the Boards of the IEEE TRANSACTIONS ON MOBILE COMPUTING, *ACM Wireless Networks*, and the *IEEE Wireless Communications Magazine*.



Mario Gerla received the Engineering degree from the Politecnico di Milano, Milano, Italy, in 1966, and the M.S. and Ph.D. degrees in engineering from the University of California at Los Angeles (UCLA), Los Angeles, in 1970 and 1973, respectively.

From 1973 to 1976, he was with the Network Analysis Corporation, New York, NY, USA. He was affiliated with the Faculty of the Department of Computer Science, UCLA, until he passed away. He had designed and implemented various network protocols (channel access, clustering, routing, and transport) under the Defense Advanced Research Projects Agency and National Science Foundation grants. His work was cited more than 21 000 with an h-index of 119.