

Received July 31, 2019, accepted August 25, 2019, date of publication August 28, 2019, date of current version September 16, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2937997

A Systematic Analysis Method for 5G Non-Access Stratum Signalling Security

XINXIN HU¹, CAIXIA LIU, SHUXIN LIU¹, WEI YOU, YINGLE LI, AND YU ZHAO

National Digital Switching System Engineering and Technological Research Center, Zhengzhou 450002, China

Corresponding author: Xinxin Hu (justinhu@hust.edu.cn)

This work was supported in part by the National Natural Science Fund for Innovative Research Groups under Grant 61521003, and in part by the National Natural Science Foundation of China under Grant 61801515.

ABSTRACT This paper proposes a systematic analysis method for 5G Non-Access Stratum Signalling security based on formal analysis, which has identified 10 new 5G protocol vulnerabilities, and an improved PKI security mechanism targeted at eliminating these vulnerabilities. Firstly, the 5G system, state transition properties and security properties were abstracted from 3GPP specifications. To mimic an attacker's behavior, a Dolev-Yao adversary model was constructed in the 5G model by empowering it with NAS signalling security testing knowledge and reasonable security capabilities in the wireless channel. Then we used the TAMARIN prover to verify all the abstracted properties one by one and discovered some protocol vulnerabilities based on the counterexamples found. We further verified these vulnerabilities on the testbed and identified 10 new 5G protocol vulnerabilities. Moreover, we analyzed the root cause of these vulnerabilities and concluded that all of them were caused by the unconditional trust between UE and gNodeB. Therefore, we propose an improved PKI mechanism based on the existing asymmetric encryption of 5G. Besides the existing public-private key pair of the home network, we introduce a new pair of asymmetric keys in the gNodeB to encrypt and sign the signalling message sent to UE. The gNodeB can be connected only when the verification succeeds and then the RRC connection can be initiated. This mechanism can effectively avoid all the vulnerabilities found in this paper.

INDEX TERMS 5G, protocol security, privacy, NAS, formal method, public key infrastructure.

I. INTRODUCTION

Since 2016, 3rd Generation Partnership Project (3GPP), the main standardization organization responsible for the development of the fifth-generation mobile communication system (5G) standards, started to work on the development of the 5G specifications. The 5G Non-Stand Alone (NSA) specifications—5G New Radio (5G NR) and Stand Alone (SA) specifications were officially frozen at the 78th and 80th plenary meeting of 3GPP in December 2017 and June 2018, respectively, which marked the official completion of the 5G first phase specifications [1], [2]. Now 3GPP is developing specifications for release 16 enabling 5G to support massive Machine Type of Communication (mMTC) and ultra Reliable Low Latency Communication (uRLLC).

Although 5G network is the evolution of 4G, it's different from 4G in many aspects. In addition to supporting

faster network connections, 5G will open the era of Internet of Things (IoT). It will be applied in the internet of vehicles, smart home, and Virtual Reality (VR) and many other scenarios. ITU takes eMBB, mMTC, and uRLLC as three major application scenarios for 5G. However, those new application scenarios, new technologies and new service methods will inevitably introduce new security challenges to 5G. According to GSMA data, the number of people using mobile phones worldwide has reached 5.03 billion by the end of 2017, accounting for almost two-thirds of the global population [3]. With the advent of 5G era, this number will continue to increase. Mobile communication network is inseparable from people's daily lives and social development, as a result of which it often becomes the target of attackers. Not only the attacks launched by attackers threaten the personal property and privacy of citizens, but also affect national security [4], [5]. It's pivotal to investigate the 5G system for detecting potential vulnerabilities. In the legacy cellular network, some attacks may prevent subscribers from accessing

The associate editor coordinating the review of this article and approving it for publication was Fan Zhang.

specific services or all services [6]–[12], while some other may seriously break the confidentiality or secrecy of user communications [13]–[20], or severely violate subscribers privacy [21]–[28]. Some of the aforementioned attacks are caused by the fact that the stakeholders do not strictly implement the 3GPP specifications, while others come from protocol defects in the 3GPP specifications. As a consequence, quite a number of these security issues need to be addressed from standard level, and 5G is no exception. 3GPP SA WG3 is set to be responsible for security and privacy related standard development in 3GPP, which determines the security and privacy requirements and specifies the security architectures and protocols. 3GPP has developed some security-related specifications for 3G, 4G, and the latest 5G [29]–[32].

Although research methods may vary, researchers in mobile network security field always focus on confidentiality, integrity, identity authenticity, anti-replay attack, availability, etc. Arapinis *et al.* [21] used formal methods to model and analyze the security properties of 3G AKA and found a protocol vulnerability can be exploited to break unlinkability of subscribers. Shaik *et al.* [27] used passive, semi-passive and active adversary models in LTE system to discover attacks that can lead to user location disclosure and denial of service. Hussain *et al.* [33] uncovered quantity of attacks related to LTE system by combining a symbolic model checker and a cryptographic protocol verifier. Rupprecht *et al.* [23] gave a comprehensive analysis on data link layer of LTE protocols and found that a resourceful adversary can perform DNS spoofing attack. Kim *et al.* [26] proposed a FUZZ based method to analyze LTE network security, through which 36 vulnerabilities were discovered. Hussain *et al.* [24] carefully analyzed the 5G paging procedure and found that an attacker could link a user's phone number with their IMSI and further track the subscriber's location by side channel information. Roger and Marojevic [41] reviewed the 5G security architecture and the main protocols of the control plane signalling at a high level. Ravishankar *et al.* [34] demonstrated a logical vulnerability used in AKA protocols, including 5G AKA, which could be exploited to reveal SQN value. Basin *et al.* [35] presented the first comprehensive formal analysis of 5G AKA using TAMARIN and found some design flaws in terms of Lowe authentication security. Adrien [36] demonstrated that the state-of-art 5G AKA is still vulnerable to some attacks and proposed AKA⁺ protocol which is σ -unlinkable. Ferrag *et al.* [37] comprehensively investigated the authentication and privacy-preserving schemes for 4G and 5G networks. Rupprecht *et al.* [38] adopted a systematization methodology for attacks and countermeasures in prior mobile networks. They divided the root causes of the attacks into four categories, among which the most serious is the unsecured pre-authentication traffic. Cremers and Dehnel-Wild [39] showed 5G AKA critically relies on unstated assumptions on the inner workings of the underlying channels by formally modeling all parties defined by 3GPP. In addition, there are also some researchers who presented comprehensive review on 5G security [40]–[44]. Although the above work has been

productive in the field of mobile network security whether in 3G, 4G, even the latest 5G, there is seldom open literature suggesting a method to systematically expose 5G NAS signalling vulnerabilities.

A. SCOPE OF THE PROBLEM

All UEs must work in accordance with the command of NAS signalling. In addition, some NAS signalling runs before the security mechanism takes effect, and it's transmitted in the air interface. So, it is often exploited actively or passively by attackers. 5G network will connect a large number of devices, so the NAS signalling security is of special importance. Before 5G UE completes the authentication and key agreement procedure, RRC and NAS signalling have no encryption protection, integrity protection and replay protection, which provide a lot of convenience for the attacker. In this paper, we assume that all the NAS signalling security mechanisms defined by 3GPP specifications for encryption, integrity, and replay protection are implemented, such as strict implementation of freshly generated ECC ephemeral public/private key pair to conceal SUscription Permanent Identifier (SUPI). As for signalling procedures, we study the registration procedure, authentication procedure, deregistration procedure, security mode command procedure, service request procedure, identification procedure, deregistration procedure, etc. In addition to considering the traditional NAS layer signalling security issues, this paper also analyzes some new security features in 5G network, such as enhanced home network control, asymmetric encryption of SUPI, etc. We don't discuss the session management signalling of the NAS layer because this part of signalling is protected by network after the authentication and key agreement is completed according to 3GPP.

B. CHALLENGES

(1) 5G system, which is an amalgamation of multiple protocols, is more complex than its predecessor. We need to abstract the protocol details and protocol role state transition process according to the specifications. (2) Converting the protocols from the natural language expression to formal language is another tough work. (3) The 5G specifications are still under developing. There are many expressions such as For Further Study (FFS) in the specifications and it is necessary to keep up with the latest version, which requires high timeliness. (4) 5G network has not been deployed on a large scale, and this has brought obstacles to commercial network test.

In order to comprehensively analyze the NAS layer signalling security of 5G network, we formally describe the 5G NAS protocols with reference to the 3GPP specifications, and model the network functions such as UE, AMF, UDM, etc. In addition, we adopt the Dolve-Yao adversary model to simulate the attacker's behavior and security capabilities. Combined with the knowledge of the NAS signalling security characteristics, we make a transition system with protocol participants as well as adversary. Subsequently, we abstract

the state transition properties that the protocols are expected to meet and the security properties that the 3GPP expects 5G protocols satisfy. Then we use symbolic modeling tools to check if all possible executions of 5G model satisfy the abstracted properties. If there was an execution which violates the abstracted property, we think that we have found a counterexample which is a possible attack path. Then we verify the attack on the testbed according to the obtained attack path. It should be noted that the work done in this paper is based on the assumptions that the security mechanisms such as encryption, integrity protection, and anti-replay algorithms can work properly and meet the requirements of 3GPP specifications. All keys in the key hierarchy are not known to attacker.

C. CONTRIBUTIONS

(1) We propose a systematic approach to analyze the security of 5G NAS protocols, which utilizes a symbolic model analysis tool to verify the expected properties of the system through model checking method. (2) We built an abstract 5G system model and the corresponding adversary model. (3) Using the proposed method, we found 10 vulnerabilities exploiting 5G NAS signalling and verified them on the testbed. (4) We propose a defense method utilizing the existing public-private key pair adopted by 3GPP, which uses the PKI to provide authenticity of NAS messages.

The reminder of the paper is organized as follows. Section II introduces the 5G network architecture and critical NAS procedures as well as signalling security comparison between 4G and 5G. In section III, we introduce the systematic method used to analyze the security of 5G protocols. Section IV presents the vulnerabilities discovered through our method in detail. Section V discusses how to verify the vulnerabilities we found on the testbed. Section VI shows the root causes of the discovered vulnerabilities and gives a potential countermeasure. A conclusion is given in section VII.

II. BACKGROUND

A. 5G ARCHITECTURE

Fig. 1 depicts a simplified 5G network architecture, which consists of User Equipment (UE), 5G access network (5G-AN), 5G core network (5GC) and data network (DN). As can be seen from the figure, 5G network introduces many new features, such as Control and User Plane Separation (CUPS),

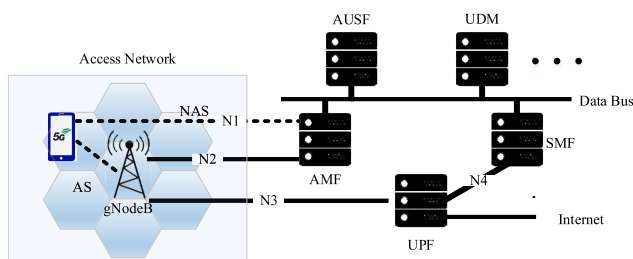


FIGURE 1. The 5G network architecture.

Service Based Architecture (SBA). The network functions in 5GC are connected through a data bus.

1) UE

UE usually consists of two parts, namely ME and USIM card. The ME is a hardware device that supports user communication, which includes a CPU, a baseband chip, a display, a battery, etc. ME stores information about the IMEI, which uniquely identifies the identity of the device. The USIM card is a universal user identity module issued by operators. The USIM card stores information such as the subscriber’s SUPI, the root key, and the operator’s public key. This information can be used to uniquely identify a legitimate subscriber and complete mutual authentication whenever UE wants to access 5G network. In 5G network, Subscription Permanent Identifier (SUPI) is a permanent identity of the user. In order to protect it from being leaked, SUPI is never transmitted in the open air interface. The permanent key K is the root key of the 5G key hierarchy, and all keys in the 5G network can be derived through this key.

2) 5G-AN

The 5G network includes multiple access methods, such as 3GPP access and non-3GPP access (N3IWF). Similar to the legacy cellular network, a large number of 5G base stations (gNodeB) form the 5G access network, and 5G-AN divides the geographical area into hexagonal cells, as a result of which UE can access to 5G network anytime and anywhere. When a UE tries to access to 5G system, the gNodeB has to allocate radio resources to it. After the RRC connection is established, the UE can continue to establish NAS connection.

3) AMF

The Access and Mobility management function is the termination of NAS signalling on the network side, which is responsible for registration management, connection management, reachability management, mobility management in 5GS as well as NAS message ciphering, integrity protection. Compared with 4G LTE, 3GPP separates the access control and mobility management function from MME to form the AMF in the 5GC. At the same time, the session management function is separated from the MME to form the session management function (SMF) in 5GC.

4) AUSF

The Authentication server function is mainly responsible for authentication of 3GPP access and untrusted non-3GPP access.

5) UDM

Unified data management function is mainly responsible for the generation of 3GPP AKA Authentication credentials, the storage and management of all user identity information (SUPI) in the 5G system as well as decryption of SUCI.

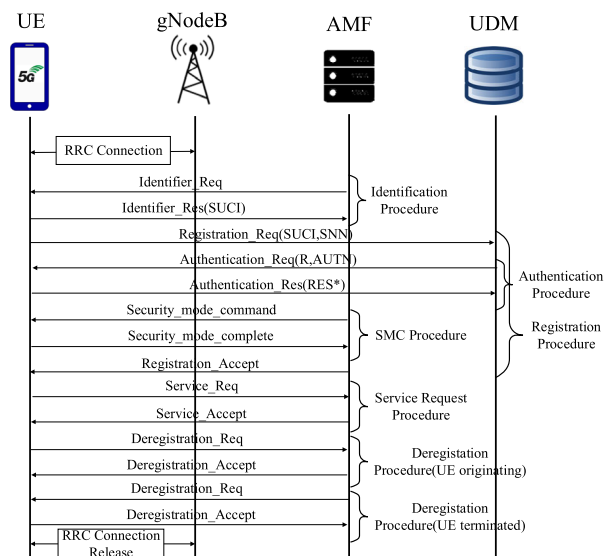


FIGURE 2. Some critical procedures in 5G system.

B. SOME CRITICAL PROCEDURES

In 5G system, the correct operation of UE requires support from many protocols. When one UE accesses the 5G system, it’s supposed to run the registration procedure to register in the network. When the UE try to register, the mutual authentication between network and the UE has to be completed. After authentication, network and the UE need to reach an agreement on the keys and algorithms used for data encryption, integrity protection, which requires the correctly running of security mode command (SMC) procedure. If the UE needs to log out and releases the wireless resources in the 5G system, the deregistration procedure needs to be executed. If the network loses the temporary identity information of the UE due to roaming or other reasons, it needs to initiate an identification procedure. The complete procedures mentioned above are shown in Fig. 2.

C. SIGNALLING SECURITY COMPARISON BETWEEN 4G AND 5G

The 5G system is evolved on the basis of 4G, and so is the Control Plane (CP) signalling system. To make the 5G system safer, many improvements have been made in control plane signalling system to avoid the vulnerabilities in 4G. In this section, we will discuss the main differences in control plane signalling between 4G and 5G. Also, the potential impact on 5G of 4G CP signalling vulnerabilities is analyzed.

1) SIGNALLING SECURITY IMPROVEMENT COMPARED TO 4G

The key differences of control plane signalling security between 4G and 5G lie in the following 6 aspects:

- In 5G security architecture, SBA domain security, which is totally a new security feature compared to 4G, requires that the control plane signalling in core network has to provide security protection that enables network

TABLE 1. Potential impact of 4G CP signalling vulnerabilities.

Attacks in 4G	Threat	Impact on 5G
Signalling abuse	DoS [7], [8]	Signalling in air interface of 5G system may be abused by attackers to launch DoS.
IMSI catcher	Subscriber identity disclosure [24], [27], [33]	The attacker obtains the subscriber’s SUPI through some attacks similar to those described in [24], or emergency services, etc.
Downgrade attack	Facing various threats in 2/3/4G networks[27], [33]	The attacker could downgrade the 5G UE to 2G/3G/4G network, causing the 5G security mechanism to fail.
User plane data tampering attack	Communication content tampering, DNS hijacking, fraud[23]	User plane data in 5G networks still lack integrity protection, which may be exploited by attackers.
Location tracking	Violate user location privacy [23], [34]	Some information related to the user’s identity, such as RNTI, SQN, etc., may be exploited by attackers.

functions of the SBA architecture to securely communicate within the serving network domain and with other network domains.

- Finer-grained control plane signalling security domain partitioning. To improve control plane signalling security, there are 6 security layers in 5G system compared to 4G in which there are only 4 layers.
- Stronger subscriber identity privacy protection. The control plane signalling in the air channel will never contain the plaintext SUPI, and instead transmit the encrypted version of SUPI—SUCI.
- Increasing home control through control plane signalling. Subscriber’s authentication response information must be passed to the HPLMN for verification through control plane signalling in 5GS.
- Security Edge Protection Proxy (SEPP) is deployed to improve inter-operator signalling protection.
- Diameter, served as application layer protocol for control plane signalling in 4G, has been replaced with more efficient HTTP/2 in 5G system.

2) POTENTIAL IMPACT OF 4G VULNERABILITIES

Prior to the development of 5G specifications, most of the LTE vulnerabilities were discussed and reported by the 3GPP security group in order to develop a more secure 5G specifications [31]. However, not all problems have been completely solved. Given the similarity of 4G and 5G control plane signalling, some security issues in 4G may continue to exist in 5G. Table 1 lists the potential impact of some main vulnerabilities in 4G on 5G.

III. THE PROPOSED METHOD

In this section, we will elaborate on the systematic approach proposed to uncover 5G network vulnerabilities. This method is a semi-automatic one that can effectively analyze the operational status of 5G protocols. Firstly, the 5G model is

abstracted from 3GPP specifications. Secondly, the adversary is constructed with Dolev-Yao model as well as NAS signalling security testing knowledge. Thirdly, the desired properties are abstracted from 3GPP specifications. To verify the properties under the adversary, we adopt the TAMARIN Prover to check all the abstracted properties under the Dolev-Yao adversary model. If there was an execution which violates the abstracted property, we think we get an attack path and will verify it on the testbed. The method mainly includes the following components.

A. ABSTRACTION OF 5G SYSTEM

In this step, we referred to 3GPP specifications for abstracting the protocols mentioned in section II-B and the protocol roles related to our research in the 5G network, such as UE, AMF, UDM. The 5G network and the protocols are abstracted into a state transition system. Although the correct operation of the protocols requires many network entities, such as gNodeB (including DU and CU), SMF, AUSF, SEAF, ARPF, etc., we simplified the network functions that are not related to NAS signalling, and classified the functions of some network entities into the NAS signalling related network entities. This kind of abstraction can simplify the problem without affecting the final analysis result. For example, the completion of the entire registration procedure requires the participation of the UE, (R) AN, AMF, UPF, SMF, PCF, AUSF in turn, and even SEAF and SEPP of the serving network are required if the UE is roaming. But the control plane signalling, especially the NAS protocols, will not be dealt with except for AMF, gNodeB, and UE. It is another problem to study the internal signalling security of other network entities inside the core network. So, when studying the registration procedure, the internal network functions of core network mentioned above were abstracted into AMF.

It should be noted that we also abstracted some of the new security features of 5G network. In order to improve the security of the 5G, 3GPP adds many new security features to it, such as enhanced home control. In the authentication procedure, it is necessary to determine whether the UE authentication is successful by the home network. At the same time, it is necessary to identify the legitimacy of the serving network by the home network and analyze the rationality of the UE roaming. Although this avoids some attacks, it also adds new attack surface. Another example is that 3GPP introduces asymmetric encryption in the 5G network to protect SUPI which is the user's permanent identity. As we will describe, this actually increases the 5G network attack surface as well.

B. ADVERSARY MODEL AND SECURITY CAPABILITY

The adversary model and its security capabilities play an important role in the approach we propose. And it relies on the reasonable modeling of the adversary's capabilities to uncover vulnerabilities.

On the one hand, we give the adversary reasonable security capabilities according to the specifications. For example, some NAS signalling lacks confidentiality, integrity,

and anti-replay protection so that an attacker can read, change, and replay these messages at will. Therefore, we check whether all NAS layer messages have confidentiality, integrity, anti-replay protection and authenticity verification. We pay special attention to such kinds of messages: (1) Messages without integrity and confidentiality protection. This kind of messages can make the attackers aware of the message contents, and change them at will. If the attacker still has some other knowledge (such as the user's legal SUCI, S-TMSI, etc.), the messages can be rebuilt and sent to network or UE which will not verify the authenticity of this message at all and even respond accordingly. Based on this, we can get some attacks. (2) Messages that have integrity protection but no confidentiality protection. Attackers can obtain some information through observing the signalling plaintext messages to prepare for subsequent attacks. (3) Messages without anti-replay protection. (4) Messages without relay protection. Finally, these security capabilities are given to the adversary in the form of knowledge, so that the attack paths that do not meet the cryptographic security settings can be ruled out. The results of NAS layer messages security verification are shown in the appendix table 4.

On the other hand, we use the Dolev-Yao adversary model to enable the adversary with reasonable security capabilities: (1) The attacker can sniff the messages transmitted in the wireless common channel without being perceived by the participating entities. Modeling is $[Out(x)] \rightarrow [K(x), [Fr(x)] \rightarrow [K(x)]$. (2) The attacker can discard and modify any message in the wireless common channel, formalized as $[!KU(x)] - [K(x)] \rightarrow [In(x)]$. (3) The attacker can impersonate a legitimate protocol entity and inject messages into the wireless common channel as the entity, interacting with other legal entities without being perceived, and formalizing it as $[K(x)] \rightarrow [In(x)]$. (4) The attacker complies with all encryption assumptions; that is to say, the attacker can decrypt the encrypted messages only if the key is mastered, and the ciphertext cannot be cracked. The above capabilities enable an attacker to initiate both passive and active attacks. In practice, an attacker can decode the airborne messages by a USRP device and packet capture tool to implement the attack capability (1). The attacker can construct the pseudo base station by loading the 5G NR protocol stack through the USRP device to realize capabilities (2) and (3). The 5G NR protocol stack is compliant with the encryption assumption, so capability (4) is guaranteed.

C. PROPERTY ABSTRACTION PRINCIPLES

To provide inspection objectives, we need to extract the properties that need to be checked from 3GPP specifications.

There are three types of properties of interest in this paper, namely model check, authentication, and privacy. We abstract these three types of properties according to the following principles.

For model check properties, our general abstraction principle is that the NAS protocol participating entities can work normally in the way that 3GPP expects. So, we first build

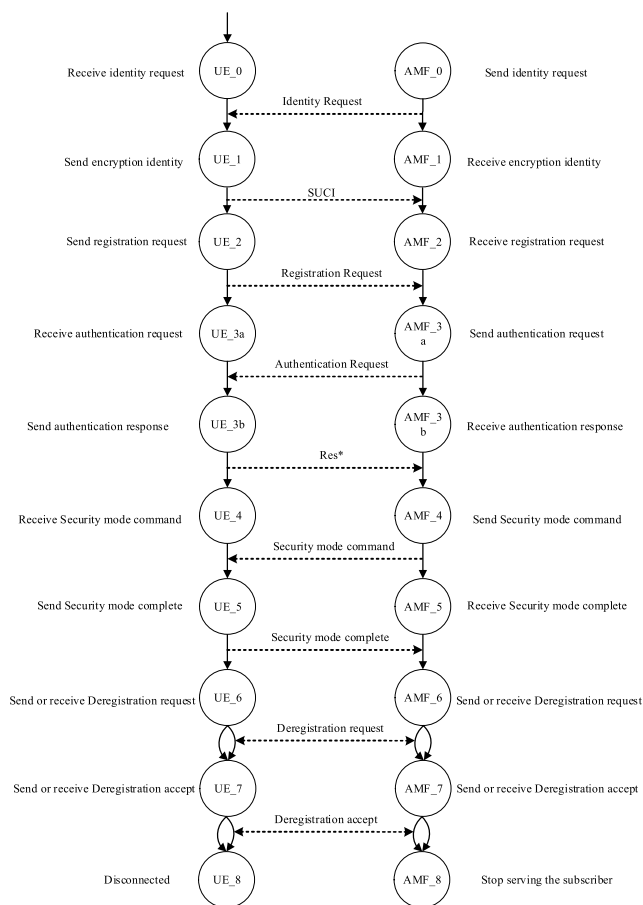


FIGURE 3. Partial state diagram for 5G NAS signalling.

the finite state machines (FSM) for the protocol participating entities as shown in Fig. 3. Then we extract the model check related properties based on the transfer relationship between the state machines as well as the following principle: The FSM of each protocol entity performs state transition in the manner expected by 3GPP without being interrupted, disturbed, or frauded. Based on this, we can easily extract the model check properties. For example, we can extract one property that a UE in the initial state will definitely enter the state of waiting for authentication request.

For authentication properties, we give the abstraction principles with reference to TS33.501-5.1.2[32]:(1) Subscription authentication: The serving network shall authenticate the SUPI in the process of authentication and key agreement between UE and network. (2) Serving network authentication: The UE shall authenticate the serving network identifier through implicit key authentication. (3) UE authorization: The serving network shall authorize the UE through the subscription profile obtained from the home network. UE authorization is based on the authenticated SUPI. (4) Serving network authorization by the home network: Assurance shall be provided to the UE that it is connected to a serving network that is authorized by the home network to provide services to the UE. This authorization is 'implicit' in the sense that it

is implied by a successful authentication and key agreement run. (5) Access network authorization: Assurance shall be provided to the UE that it is connected to an access network that is authorized by the serving network to provide services to the UE. This authorization is 'implicit' in the sense that it is implied by a successful establishment of access network security. This access network authorization applies to all types of access networks. For instance, we can extract one property that UE and UDM need to reach an agreement on UDM identity.

For secrecy properties, we give the abstraction principles with reference to TS33.501-5.2: (1) The UE shall support ciphering of NAS-signalling. (2) Confidentiality protection of the NAS-signalling is optional to use. (3) The UE shall support integrity protection and replay protection of NAS-signalling. (4) Integrity protection of the NAS-signalling is mandatory to use except for several cases. (5) The subscription credential(s) shall be integrity protected within the UE using a tamper resistant secure hardware component. (6) The long-term key(s) of the subscription credential(s) (i.e. K) shall be confidentiality protected within the UE using a tamper resistant secure hardware component. (7) The long-term key(s) of the subscription credential(s) shall never be available in the clear outside of the tamper resistant secure hardware component. (8) The SUPI should not be transferred in clear text over NG-RAN except routing information, e.g. Mobile Country Code (MCC) and Mobile Network Code (MNC). (9) The ME shall support the null-scheme. If the home network has not provisioned the Home Network Public Key in USIM, the SUPI protection in initial registration procedure is not provided. In this case, the null-scheme shall be used by the ME. For example, we can extract one property that the attacker cannot know the root key K.

We abstract the properties to be verified for UE as partly shown in table 2. The abstraction rules of the properties for other network functions is similar.

TABLE 2. Properties abstracted from UE perspective.

ID	Classification	Description
P1	Model Check	The UE in the initial state will definitely enter the state of issuing the registration request.
P2	Model Check	The UE in the state of issuing the registration request will definitely enter the state of receiving the Authentication request.
P3	Model Check	The UE in the state of receiving the Authentication request will definitely enter the state of issuing authentication response.
P4	Model Check	The UE in the state of issuing authentication response will definitely enter the state of authentication succeed.
P5	Authentication	UE and UDM need to reach an agreement on UDM identity.
P6	Authentication	UE and UDM need to reach an agreement on SNN.
P7	Authentication	UE and AMF need to reach an agreement on UDM identity.
P8	Secrecy	The attacker cannot know the root key K.

D. TAMARIN MODEL CHECK

In order to perform formal analysis automatically, we use the TAMARIN prover [45] for model checking. The TAMARIN prover is a powerful tool for symbolic modeling and analysis of security protocols. It takes the security protocol model as input and specifies the actions taken by the protocol running agent in different roles (e.g., protocol initiator, responder) and then automatically builds evidence. Even if the role of the protocol has quite a number of instances interleaved in parallel, it can be run with the actions of the adversary. The attacker and the protocol interact by updating the network messages and generating new messages. TAMARIN uses a multi-set rewrite-based expression language to define protocol participants and attackers. These rules define a labeled transition system whose state includes symbolic representations of adversary knowledge, messages on the network, newly generated information, and protocols state, etc. In the TAMARIN prover, the protocol state and its conversion are described by *rules*. The security property of the expected verification is described by *lemmas*. The rules and lemmas are input together into TAMARIN. When the protocol meets the corresponding security properties, TAMARIN will output verification success, otherwise both falsified result and attack path will be output.

We use the sptyh language to describe the 5G protocols and its participating roles as a state transition system. When the UE in the initial state sends a registration request message to the network, the UE transfers from the initial state to the waiting for authentication request state. The above state transition is described in sptyh language as follows:

```

Rule UE_send_registrationReq:
  let
    m = <SUCI, UDM, SNN>
  in
    [! LTK (SUPI, SUCI, UDM, SNN, K)]
  ->
    [St_UE_1_USIM (SUPI, SUCI, UDM, SNN, K), Out(m)]
    
```

For security properties that need to be verified, they need to be described by lemmas. For example, UE and AMF need to reach agreement on UDM identity (P7), which can be described as follows:

```

Lemma agreement_UE_AMF_UDM:
  "All a b c d t #i.
  (Commit(a, < a, b, c, d >, t, <'UE','K_AMF'>))@ #i
  &not(Ex #r.
  RevealK(a)@r & Honest(<a,d>))@ #i)"
  ==>(Ex a2 b2 c2 t2 #j.
  Running(b2, < a2, b2, c2, d>, t2, <'AMF','K_AMF'>))@ #j)"
    
```

In the above description, there is a *commit(...)* action fact in the UE for all paths, where the UE considers that the parties involved in the protocol are a, b, c, and d (instantiating as UE, AMF, AUSF and UDM). UE considers the session key K to be the term t, and the long-term key K is not known by the adversary. Then there must be at least one *Running(...)* action

fact from AMF, which means AMF reach an agreement with the UE on the identity of UDM.

By inputting the 5G model and the properties to be verified into the TAMARIN prover in the above manner, the prover will automatically check all properties.

E. VALIDATION ON THE TESTBED

After getting the attack paths in TAMARIN, we design the corresponding attack method and verify it on the testbed. The reason why verification is needed is because the attack paths obtained by TAMARIN are not necessarily completely reliable. On the one hand, the modeling of the 5G protocols may deviate from the actual situation, resulting in infeasible attack paths. On the other hand, it may be that the protocols have avoided the relevant attack paths in the implementation. So, we verify all the attack paths on the trial network to avoid the infeasible ones.

IV. ATTACKS DISCOVERED

In this section, we detail the vulnerabilities discovered by our proposed method, and introduce each attack from three aspects: adversary assumptions, attack procedure and implications. At the end of this section, we summarize all the discovered vulnerabilities in a table.

A. AUTHENTICATION FAILURE ATTACK

1) ADVERSARY ASSUMPTIONS

If the attacker's target is a specific user, it is necessary to obtain the temporary or permanent identity information of him or her. According to [24], the SUPI can be cracked. Besides, if the operator does not update the 5G-GUTI in time, the user identity information can also be obtained. This assumption is reasonable because a TMSI generally uses 8 ~ 10 times or updates follow certain rules according to the experience in prior cellular network [25]. In addition, the attacker needs to construct a malicious gNodeB.

2) ATTACK PROCEDURE

The attack procedure is shown in Fig. 4. To prepare for the attack, the attacker first needs to eavesdrop the signalling messages in vicinity of the victim UE. When the victim UE tries to connect to the 5G network, it has to initiate a

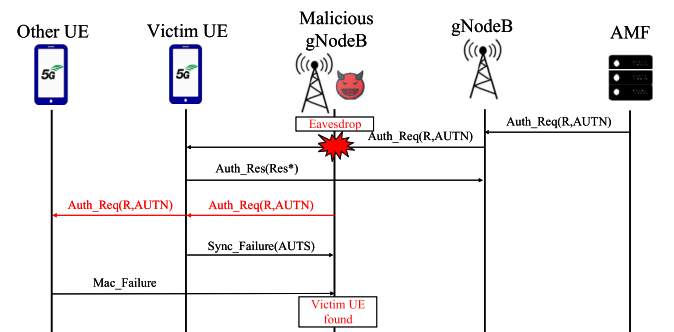


FIGURE 4. Authentication failure attack.

registration request with SUCI or 5G-GUTI included. Then the network side AMF will respond the victim UE with an authentication request message including a nonce (denoted by R) and AUTN in plaintext. As soon as the attacker captures the authentication request, he immediately saves the R and AUTN from the message.

The attacker could attract the UE attach to the malicious gNodeB with a stronger signal power whenever he wants to check the position of a target subscriber. Then he can initiate authentication procedure and replay the previously intercepted authentication request message. After receiving the authentication request, the victim UE performs MAC and SQN verification. MAC verification is carried out to verify whether the message is sent from the real network, and SQN verification is executed to verify whether the network and UE have lost synchronization. If both of the verifications are successful, the RES* is calculated and sent with authentication response message. In the authentication failure scenario, MAC verification will pass but SQN verification will fail. This is because the authentication request message comes from the real network and the SQN value in the UE side has changed in the authentication procedure during which the attacker intercepts the authentication request message. Based on the response received, the attacker can make a determination whether the subscriber is located in the current cell. If the attacker receives a synchronization failure message (Sync_Failure), he can conclude that the target victim is in the current cell because other subscribers will return a MAC verification failure message (Mac_failure).

3) IMPLICATIONS

This attack can be used by an attacker to track the location of a user. For example, the attacker can determine whether certain important people (such as ambassadors, national leaders, etc.) is in the current cell. After intercepting an authentication request message, the attacker only needs to run the above attack again whenever he wants to judge the location of the user.

B. LOCATION TRACKING WITH REPLAYED SMC

1) ADVERSARY ASSUMPTIONS

The attacker needs to forge a malicious gNodeB as well as a malicious UE using USRP to construct a communication channel between the victim UE and the legitimate network.

2) ATTACK PROCEDURE

Fig. 5 depicts the complete attack procedure. In the first phase, the attacker needs to eavesdrop in the same cell with victim UE. After the victim UE is authenticated, the network initiates a security mode command procedure for the victim UE. According to TS33.501 [32], the security mode command message has integrity protection but no encryption, and the attacker can recognize and save it locally. In the second phase of the attack, the victim UE can be forced to connect to the malicious gNodeB with stronger signal power near

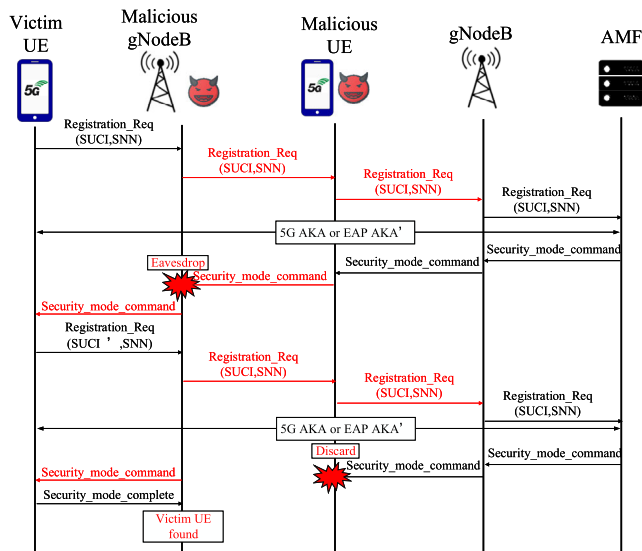


FIGURE 5. Location tracking with replayed SMC.

the target cell whenever the attacker needs to determine the location of the target subscriber. After the RRC connection is established, the victim UE will initiate registration to the network. When the real network and the victim UE perform normal authentication, the malicious gNodeB and the malicious UE need only transparent transmission without interference. When the real network initiates the security mode command procedure, the attacker discards the security mode command message and sends the previously stored security mode command message to the victim UE. If the attacker receives the encrypted security mode complete message from the victim UE, he can conclude the target UE is in the current cell. If the attacker receives the plaintext security mode reject message, he can conclude the target UE is not in the current cell.

This attack is inspired by Hussain et.al [33]. However, there are at least three differences between this attack and the counterpart of huassin et al. First, the attack is discovered in 5G system. Second, this attack implementation considers the state transition characteristic of the NAS protocols, making the attack closer to the practical situation. Third, we set MITM between the victim UE and legitimate gNodeB so that the attacker can discard the SMC sent from legitimate gNodeB timely. This will greatly increase the attack success rate.

3) IMPLICATIONS

This attack can be used to determine whether the target user is in a target cell, thereby enabling tracking of a target subscriber's location.

C. DEREGISTRATION ATTACK

1) ADVERSARY ASSUMPTION

The attacker needs to construct a malicious gNodeB using the USRP device and the 5G protocol stack. In the targeted DoS

attack, it is also necessary to have a malicious UE and the 5G-GUTI of the victim UE.

2) ATTACK PROCEDURE

In the first attack (see Fig. 6(a)), the attacker can initiate a Deregistration request (UE originating de-registration) as the victim UE. In this request message, the 5G-GUTI of the victim UE needs to be included so that the victim UE will receive the deregistration accept. If the attacker sets the third bit position of the deregistration type field to 0, this can make the victim UE stay in no service state for a long time. It is also possible to set the third bit position to 1, which enables the UE to re-initiate registration and prepare for subsequent attacks. In the second attack (see Fig. 6(b)), the attacker can set a malicious gNodeB with stronger signal power in vicinity of victim UEs to make all UEs connect to it, and then broadcast the deregistration request (UE terminated de-registration), which can make all UEs in the cell have no access to communication service. In addition, the third bit position of the deregistration type field may be set to 1 in this attack, which will cause all subscribers in the cell to initiate registration requests at the same time, resulting in a signalling storm.

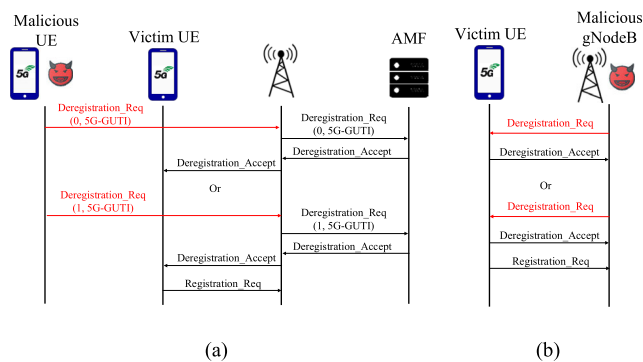


FIGURE 6. Deregistration attack.

3) IMPLICATIONS

The above attacks can cause a single user’s DoS as well as large-scale DoS in the entire cell, and can also cause signalling storm of one gNodeB. These attacks may adversely affect the reputation of the carriers. In the second attack, the registration request initiated by the real UEs through broadcasting deregistration request is simpler and more effective than repeatedly initiating the RRC connection request mentioned in [26].

D. REGISTRATION ATTACK

1) ADVERSARY ASSUMPTION

The attacker needs to obtain the RRC connection information of the target gNodeB, so that the attacker can initiate the same connection request as the UE. The information can be obtained by eavesdropping on SIB and MIB broadcast by the

target gNodeB. In addition, it is necessary to collect a lot of real SUCIs or 5G-GUTIs.

2) ATTACK PROCEDURE

Fig. 7 depicts the complete attack procedure. First, the attacker constructs a malicious UE and initiates an RRC connection to the target network. According to 3GPP specifications, the attacker piggybacks a registration request message in the RRC connection complete message so that the network initiates an authentication request to the malicious UE. The real network will wait for the malicious UE to return an authentication response message and this connection will last for several seconds. If the malicious UE does not return an authentication response, the gNodeB will release its RRC connection with the malicious UE. In order to ensure the success of the attack, the attacker needs to repeatedly initiate the registration request, and the malicious UE needs to restart to obtain a new C-RNTI, otherwise it will be regarded as the same UE by the gNodeB. In addition, the malicious UE must continuously initiate new RRC connection request before the old RRC connection is released by gNodeB, so as to ensure that the number of RRC connections released by the gNodeB is less than the re-initiated RRC connection.

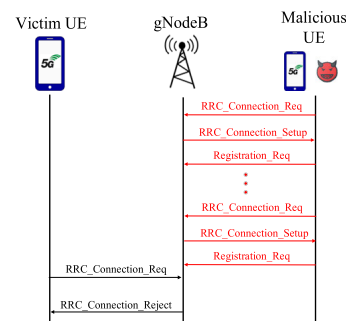


FIGURE 7. Registration attack.

3) IMPLICATIONS

This attack may lead to regional DoS of the target gNodeB, as a result of which it may cause users’ dissatisfaction and the carrier’s reputation loss.

E. LOCATION TRACKING WITH SUCI

1) ADVERSARY ASSUMPTION

The attacker needs to intercept the real registration request message (including the SUCI) of the target subscriber, and needs to construct a malicious UE as well as a properly configured malicious gNodeB. There can be a dedicated communication channel between the malicious UE and the malicious gNodeB.

2) ATTACK PROCEDURE

The complete attack procedure is shown in Fig. 8. In the first phase of the attack, the attacker needs to continuously listen in the cell where the target user may appear to obtain

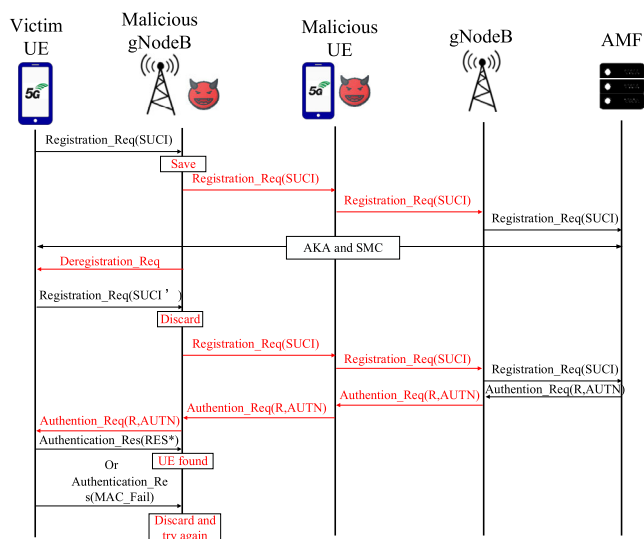


FIGURE 8. Location tracking with SUCI.

the real registration request message of the target subscriber. This message is prepared for the subsequent replay attack because the request message contains the identity information of the target subscriber. Although the attacker even does not know what the identity information is, it does not affect the attack. What the attacker only needs to know is that the message corresponds to the target user. In the second phase, the attacker deploys a malicious base station near the target cell with a stronger signal. The victim UE will release the previous RRC connection and connect to the gNodeB with the strongest signal power, which is the malicious gNodeB. Then the attacker sends a deregistration request with the third bit position of deregistration type field to 1 so that the victim UE will initiate a registration request. In the third phase, when the victim UE initiates the registration request, the attacker discards the message through the malicious gNodeB, and sends the previously intercepted registration request message to the malicious UE. The malicious UE sends the message to the legal gNodeB through the air interface. According to the normal authentication procedure, the network will initiate an authentication request (including AUTN and R, etc.) to the victim UE, and the malicious UE transmits the message to the victim UE through the malicious gNodeB. After receiving the message, the victim UE will perform MAC verification and return the authentication response message. Since the response message is plaintext, the attacker can judge whether the authentication is successful by observing the content of the message. If the authentication response content is RES*, the authentication succeeds, and the victim UE is the target UE. If the authentication response content is Mac_failure, it proves that the victim UE is not the target UE.

3) IMPLICATIONS

Through this attack, an attacker can verify whether the target user is in the current cell. Whenever an attacker wants to

determine the location of the target user, the attacker can launch the attack to achieve the target subscriber’s location tracking. Further, if the victim UE is the target UE, the attacker can allow the target UE to complete the subsequent registration procedure. As all traffic of the target UE passes through the attacker’s network (malicious gNodeB and malicious UE), the traffic can be analyzed by attackers to build target user service usage habits profile (such as phone calls, text messages, and data services). Unlike a typical man-in-the-middle attack, this attack does not require the attacker to have cryptographic knowledge, which means that there is no need to crack the encrypted data or modify the real signalling message.

F. INCREASING HOME CONTROL ATTACK

1) ADVERSARY ASSUMPTION

The attacker needs to construct a malicious gNodeB and several malicious UEs through USRP devices. In addition, he has to make the malicious UEs connect to different visiting networks.

2) ATTACK PROCEDURE

In order to improve the security of 5G network, 3GPP enhances the home control of AKA protocols. The authentication response generated by UE not only needs to be verified at serving network, but also at home network. In addition, the UE is supposed to return the RES* and the corresponding SUCI or SUPI in the authentication response to the AUSF and UDM of the HN so that the HN can comprehensively judge whether the authentication is legal according to authentication time, authentication type and the serving network name (SNN), etc. For example, a subscriber has just registered in a visiting network of the US, while another subscriber with the same SUPI registers again in another visiting network of the UK in just a few seconds or a few minutes. It is obvious that there is an abnormality. Even if the authentication response is completely correct, the authentication of the UE should not be passed (see 3GPP TS 33.501-6.1.4.2 [32]. When a new Nudm_UECM_Registration Request arrives from a visited network, the home network checks whether there is a recent authentication of the Subscriber by this visited network. If not, the Nudm_UECM_Registration Request is rejected. It is up to the home network to set the time threshold to define what 'sufficiently recent' is.). This property can be exploited by attackers.

Fig. 9 shows the complete attack procedure. Firstly, the attacker tries to make the victim UE connect to the malicious gNodeB with highest signal power. Secondly, the attacker transparently transmits the message through the malicious gNodeB and one of the malicious UE. The malicious UEs can appear in various visited places which may have a long physical distance. Attacker can reach this goal by constructing the hardware of the malicious UE through USRP with 5G protocol stack. Then he places these malicious UEs in different geospatial location and makes them

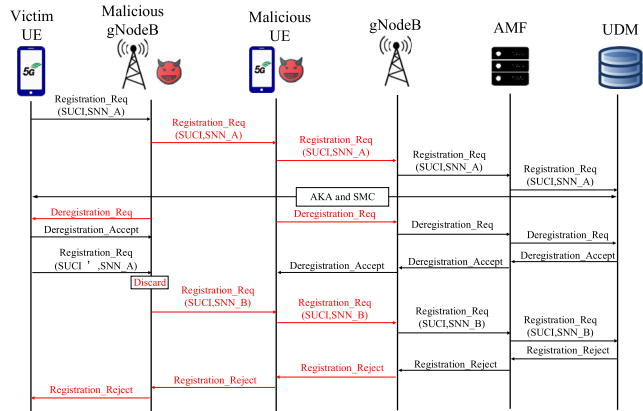


FIGURE 9. Increasing home control attack.

connected to the local PLMN. Attacker can remotely control those malicious UEs via the internet or IoT. When the attacker needs to make a malicious UE “appear” in a different visited location, he only needs to forward the registration request signalling of the victim UE to the malicious UE through the network. And the malicious UE sends the signalling to the local cellular network. The attacker only needs to constantly switch the “serving” UE to make the victim UE “appear” in different visited places. And the switching interval is no longer than 5 minutes so that the real registration request of the victim UE is finally judged by the home network as a fraudulent one. Thus, the HN denies its access to the network, causing the victim UE suffering a DoS attack without being noticed.

3) IMPLICATIONS

The victim UE will be denied access to the network without any notification. The duration of the DoS attack depends on the operator’s setting. So, it may vary from hours to days.

G. ENERGY DEPLETION ATTACK

1) ADVERSARY ASSUMPTION

The attacker uses a USRP device to construct a malicious gNodeB.

2) ATTACK PROCEDURE

In order to enhance security, 3GPP introduces an asymmetric key architecture to encrypt SUPI so that 5G network no longer transmits the user’s permanent identity information in the clear text. The subscriber identification mechanism may be invoked by the serving network when the UE cannot be identified by means of a temporary identity (5G-GUTI). To begin with, AMF sends an identity request message to the UE through gNodeB. After receiving the request, the UE generates a temporary shared key based on the network side public key and the private key in the temporary asymmetric key pair generated by itself. Then the UE uses the temporary shared key to encrypt the SUPI to obtain the SUCI, and sends the newly generated SUCI to the network in plaintext. The

asymmetric encryption is used to enhance the security of 5G and solves the IMSI catcher problem to a large extent that has been existing from 2G, 3G to 4G. However, the energy consumed by asymmetric encryption is much larger than that of symmetric encryption. Attacker can initiate the identification procedure by continuously sending identifier request, causing the victim UE to continuously perform asymmetric key generation and encryption, which will eventually lead to energy depletion of the victim UE. Fig. 10 presents the complete attack procedure.

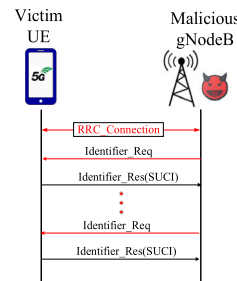


FIGURE 10. Energy depletion attack.

3) IMPLICATIONS

This attack may cause the UE to consume too much power or even shut down.

H. AUTHENTICATION REJECT ATTACK

1) ADVERSARY ASSUMPTION

The attacker only needs to construct a malicious gNodeB.

2) ATTACK PROCEDURE

When an RRC connection is established between the victim UE and the malicious gNodeB, the malicious gNodeB directly sends an authentication reject message to the victim UE (see Fig. 11). Upon receipt of the message, the victim UE will automatically disconnect the RRC connection and remain in out of service state for a long time.

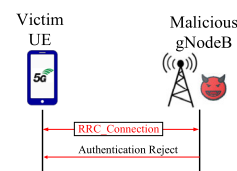


FIGURE 11. Authentication reject attack.

3) IMPLICATIONS

This attack can cause the UE to suffer severe service disruption.

I. REGISTRATION REJECT ATTACK

1) ADVERSARY ASSUMPTION

The attacker needs to construct a malicious gNodeB with USRP.

2) ATTACK PROCEDURE

Fig. 12 depicts the complete attack procedure. Firstly, the malicious gNodeB signal power is made large enough to make the victim UE attach to it. After the victim UE initiates the registration request message, the attacker piggyback a registration reject message in the DL information transfer message. The attacker can fill the 5G MM cause field with “0 0 0 1 0 1 1,” indicating that the PLMN is not allowed, and this will cause the victim UE to remain in out of service state forever unless the victim UE reboots or reinstalls the SIM card.

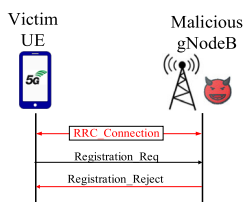


FIGURE 12. Registration reject attack.

3) IMPLICATIONS

This attack can cause the victim UE to be in a severe denial of service state. And the attacker can selectively fill in the 5G MM cause field for different reasons, so that the victim UE can be in the permanent DoS state or re-initiate the registration immediately.

J. SERVICE REJECT ATTACK

1) ADVERSARY ASSUMPTION

The attacker needs to build a malicious gNodeB as well as a malicious UE.

2) ATTACK PROCEDURE

When the victim UE needs certain services (such as call, SMS, or receiving paging messages, etc.), the victim UE will initiate a service request. If the attacker connects the victim UE and the legal gNodeB through attacker-controlled gNodeB and UE, all user plane data and NAS signalling will pass through the attacker network. If the attacker wants to reject certain services of the victim UE such as turning off the call and SMS function, the malicious gNodeB can respond to him with service reject message when the UE initiates the service request, which will cause the victim UE suffer from local DoS.

3) IMPLICATIONS

The victim UE will be unable to access certain services without notice.

A complete list of the discovered attacks is given in the Table 3. We sort out the attacks with attack name, adversary assumption, implication and exploited message.

TABLE 3. Summary of our findings.

ID	Attack name	Adversary assumption	Implication	Exploited message			
				Message name	Non-integrity	Non-encryption	replayable
A1	Authentication failure attack	Malicious gNodeB	Location exposure	Authentication request	√	√	√
A2	Location tracking with replayed SMC	Malicious gNodeB, Malicious UE	Location exposure	Security mode command	×	√	√
A3	Deregistration attack	Malicious UE or Malicious gNodeB	DoS	Deregistration request	√	√	√
A4	Registration attack	Malicious UE	DoS	Registration request	√	√	√
A5	Location tracking with SUCI	Malicious gNodeB, Malicious UE	Location exposure	Registration request	√	√	√
A6	Increasing home control attack	Malicious gNodeB, Malicious UE	DoS	-	-	-	-
A7	Energy depletion attack	Malicious gNodeB	Energy depletion	Identity request	√	√	√
A8	Authentication reject attack	Malicious gNodeB	DoS	Authentication reject	√	√	√
A9	Registration reject attack	Malicious gNodeB	DoS	Registration reject	√	√	√
A10	Service reject attack	Malicious gNodeB, Malicious UE	DoS	Service reject	√	√	√

V. VALIDATION OF ATTACKS WITH TESTBED

In this section, we will explain how to configure the testbed and validate the discovered attacks on the testbed.

A. TESTBED SETUP

In the testbed, we used two USRP devices, of which USRP1 was used as a malicious gNodeB and USRP2 as a malicious UE. Two types of commercial 5G mobile phone

J. Service Reject Attack

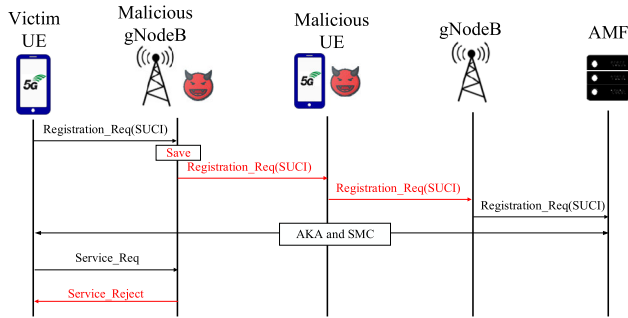


FIGURE 13. Service reject attack.

released by different equipment vendors were used as victim UEs. We used a 5G precommercial test network, coming from vender C, to work as a real commercial network. As for moral and legal considerations, our experiments were not carried out in a shielding box or a Faraday cage because there is no commercial 5G network in the experimental location. Our experiments will not harm the legitimate rights and interests of operators and ordinary subscribers.

Malicious gNodeB and malicious UE settings: In some attacks, only one malicious gNodeB needs to be constructed, while other attacks require both malicious gNodeBs and malicious UEs to cooperate with each other. We hereby explain how to construct malicious gNodeBs and malicious UEs. In constructing a malicious gNodeB, we need a USRP hardware device (denoted by USRP1) and a PC with Intel processor (denoted by PC1) running the Ubuntu operating system as well as gNodeB protocol stack. The configuration of a malicious UE is similar to that of a malicious gNodeB. It requires a USRP hardware device (denoted by USRP2) and a PC with an Intel processor (denoted by PC2) running the Ubuntu operating system as well as the 5G-UE protocol stack.

Commercial 5G UE and 5G precommercial test network: The devices used to act as victim UEs are commercial 5G terminals issued by equipment vendors, which are now available on the market. 5G SIM card is programmed in our laboratory with empty SIM. At the same time, the experimental 5G pre-commercial test network is provided by a communication equipment manufacturer. The test network is slightly different from the 5G commercial network equipment that is available for carriers, and the test network is constantly being upgraded.

The complete experimental environment is shown in Fig. 14. The registration procedure package captured on the testbed under above configuration is shown in Fig. 15. All the attacks aforementioned are validated in our testbed.

B. VALIDATION WITH PRECOMMERCIAL TEST NETWORK

In this section, we describe in detail how each attack was successfully verified. All the discovered attacks are classified into three categories, namely location exposure attacks, DoS, and other attacks.

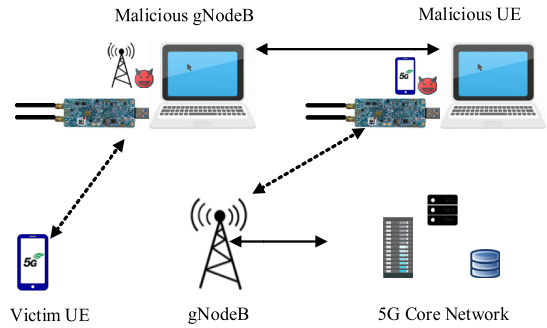


FIGURE 14. Testbed environment.

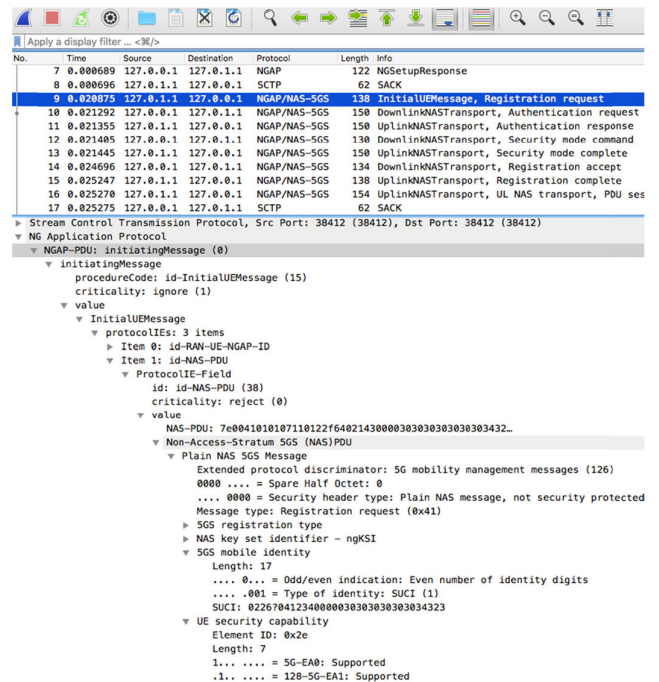


FIGURE 15. Data packets captured in our experimental environment.

1) LOCATION EXPOSURE ATTACKS

As shown in table 3, A1, A2, and A5 can be exploited by attackers to expose subscribers' location. In essence, all three of the above attacks have one thing in common, that is, they utilize the real messages transmitted between subscriber and network. And the used messages contain information uniquely corresponding to a certain subscriber. Moreover, the subscriber will respond differently to other users for this real message, which becomes the basis for attacker to determine whether the target subscriber is at a specific location.

A1: In order to verify the attack, we used several UEs, and one of them is the target user. We first used the sniffer to capture the authentication request message sent by the network to the victim UE during a benign authentication procedure. Then we decoded the message and extracted all the fields. Subsequently, we constructed the authentication request message using the extracted fields and sent the message to all UEs using the malicious gNodeB. We found that

only the target user responded with Sync_Failure, while other UEs responded with Mac_Failure.

A2: This attack is verified in the similar way as A1. At first, we set up a sniffer to eavesdrop on the security mode command message received by the victim UE during the authentication procedure. Then we make several UEs including the victim UE attach to the malicious gNodeB through attack A9. The malicious gNodeB is connected with the malicious UE through Internet, and the latter accesses the legal cellular network. In this way, we can precisely control the timing of sending the captured security mode command message. Then we reboot all UEs and let them initiate the registration procedure. Once the UEs is detected to send RES* to the network, we injected the previously intercepted security mode command message to the UEs by malicious gNodeB. Through the security mode complete message, we identified the target UE from the others.

A5: The verification of this attack is similar to the above two. We use the sniffer to capture the registration request message of the target user. Then several UEs were set to connect to the malicious gNodeB. When these UEs sent the registration request, we replaced these requests with the previously captured registration request and sent it to the core network. After receiving the authentication request from the core network, we forward the message to the requested UEs. It turns out that only the victim UE sends RES*, while other UEs send Mac_failure.

2) DOS ATTACKS

There are 6 attacks that can be exploited by an attacker to interrupt user services, namely A3, A4, A6, A8, A9, A10. The principles of these attacks are not exactly the same. For example, A3, A8, A9, and A10 utilize some 5G NAS signalling, while A4 rejects other legitimate subscriber services by excessively occupying the gNodeB resource. A6 exploits the new security mechanism of the 5G control plane signalling.

A3, A8, A9, and A10: To verify A3, we need to determine the user's 5G-GUTI as described in [46]. Then we constructed the deregistration request message using the extracted fields such as the victim's 5G-GUTI, deregistration type, etc. And the message is sent to all UEs through the malicious gNodeB. We observed that only the victim UE stayed in "no service" state for a long time unless we restarted it or re-inserted the SIM, while other UEs were not affected. A8, A9, A10 were verified the same way as A3.

A4: We use USRP to simulate many UEs and make them continuously initiate registration request. In the experiment, we had to ensure that the new registration request generation rate was greater than the RRC release rate. In this way, the total number of simulated devices that maintain an RRC connection with the legitimate NodeB would continue to increase. After reaching the service threshold of gNodeB, we reboot the victim UE to initiate a registration request. Then the connection request is rejected by gNodeB.

A6: In order to verify attack A6, we configured the core network of the pre-commercial 5G network, so that the home control security mechanism was effective. In the experiment, the switching time threshold of different visited networks was configured as no less than 5 minutes. Once it was detected less than 5 minutes, the SIM registration service would be terminated. The service would resume after 24 hours. In this configuration, we connected the victim UE to the malicious gNodeB, and then connected the malicious gNodeB with several malicious UEs located in different visited networks. The attack was carried out according to IV-F, and as a result the victim UE could not obtain the service within 24 hours.

3) OTHER ATTACKS

A7: To verify A7, we used the malicious gNodeB to continuously send the identity request message to the victim UE, and the victim UE will respond with its newly calculated SUCI. The experimental results show that the energy consumption of SUCI in 5G system was about 109.211mJ. We sent one identity request per second, so the extra energy consumption of the victim UE in one hour was roughly $109.211 \times 10^{-3} \times 3600 = 393.1596$ J.

VI. ROOT CAUSE AND COUNTERMEASURE

In this section, we will discuss the root causes of the vulnerabilities mentioned above and propose a countermeasure.

A. ROOT CAUSE ANALYSIS

Technically analyze the aforementioned attacks, we find the root causes mainly remain in the following aspects: (1) Lack of integrity, confidentiality, anti-replay, anti-relay protection of the initial RRC message. (2) Lack of integrity, confidentiality, anti-replay, anti-relay protection of the pre-authentication messages. (3) UE unconditionally trusts any gNodeB. (4) Attacker can eavesdrop on the air interface messages at will. The most serious reason for the above four is that the UE unconditionally trusts gNodeB, which may cause attacker to initiate a connection with the victim UE at will. From the perspective of carriers and equipment vendors, the above technical compromises adopted by 3GPP are likely due to the trade-offs between security and availability, functionality, efficiency, cost, etc.

In the current network deployment situation, the deployment cost may be greatly increased if some network mechanisms are further improved, and the backward compatibility problem needs to be considered as well. In addition, the impact to the entire system of adding additional mechanisms needs to be reconsidered. Based on past experience, upgrades on an already deployed network is unlikely to be accepted by existing stakeholders. Fortunately, the current 5G specifications are still under development so it's still possible for 3GPP to improve security of 5G system from standard level. Besides, communication equipment vendors can make unified changes accordingly, which can greatly improve efficiency and reduce the cost. For 5G, the countermeasure

proposed has to be a reliable, lightweight, and quality guaranteed one.

B. COUNTERMEASURE

We now discuss the potential countermeasure against the vulnerabilities we discovered. By a very clear idea, Hussain et al. [47] proposed an authentication scheme which leverages the precomputation-based digital signature generation algorithms. However, their solution is a bit complicated and does not take full advantage of the existing PKI mechanism of 5G. We propose a more simplified countermeasure which utilizes the existing PKI mechanism and adds just one pair of public-private key to provide authenticity authentication for gNodeB messages as well as its identity. The PKI mechanism in 5G network is designed to protect the user’s permanent identity information SUPI. The mechanism adopts the elliptic curve algorithm (ECC) to generate a pair of public and private keys. The public key is stored in the SIM card of the 5G UE, and the private key is stored in the core network of carriers. When the network needs to request the UE identity information, the UE combines the public key stored in the SIM with the private key of the temporarily generated public-private key pair to form a shared key, which is used to encrypt the SUPI to generate the SUCI. The carrier core network then uses the temporary public key sent by the UE and the private key stored by itself to form the same shared key, which can be used to decrypt the SUCI and obtain SUPI. The carrier public-private key pair used in the above process is permanent. So, we can use the private key stored in the core network to sign the identity of the base station. In order to prove the legitimacy of gNodeBs, we give every gNodeB a certificate with the signature of HN:

$$cert_{BS_i} = P_{BS_i}, CELL_ID, Loc_{BS_i}, Ext_{cert_{BS_i}}, Sig_{HN}$$

where, P_{BS_i} , $CELL_ID$, Loc_{BS_i} , respectively denote the public key, unique identities, precise physical location of one base station. $Ext_{cert_{BS_i}}$ indicates the validity period of the certificate $cert_{BS_i}$. The signature of HN and SIB2 are computed as follows:

$$sig_{HN} = sign \left(\left(P_{BS_i}, CELL_ID, Loc_{BS_i}, Ext_{cert_{BS_i}} \right), SK_{HN} \right)$$

$$sig_{SIB2} = sign \left(\langle SIB1 \parallel SIB2 \rangle, SK_{BS_i} \right)$$

where the signature function $sign(m, k)$ induces the digital signature of m with a secret key k . SK_{HN} is the private key of HN.

The certificate and signature are attached to the original SIB1 message, and the UE determines whether to trust the gNodeB according to the signature information. In order to prevent relay attacks, we have adopted a similar philosophy as [47]. We add delay parameter to SIB2. Operators can set a reasonable threshold Δ_t according to the local wireless environment. The network side SIB message generation time is recorded as T_{gen} and the receiving time of SIB is recorded as T_{rec} . The UE calculates the difference between T_{gen} and T_{rec} , and compares it with the threshold Δ_t . If and only if

$T_{rec} - T_{gen} < \Delta_t$, the UE considers that the SIB message is not relayed. This can greatly reduce the attack window of malicious gNodeB. In order to prevent the attack from passively eavesdropping on the air interface information, we also add an initial RRC and initial NAS layer message encryption scheme. Specifically, the UE encrypts the initial RRC connection message and the initial NAS connection message with PK_{BS_i} . The specific plan is shown in Fig. 16, Fig. 17.

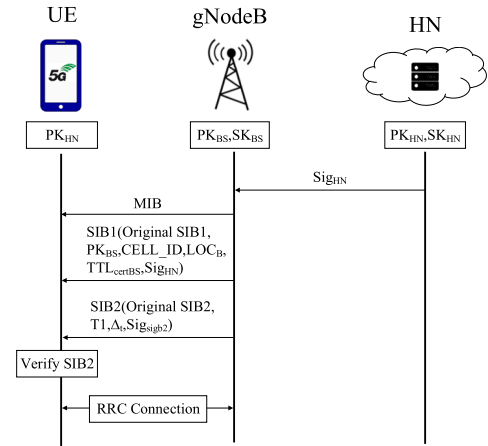


FIGURE 16. Our proposed countermeasure.

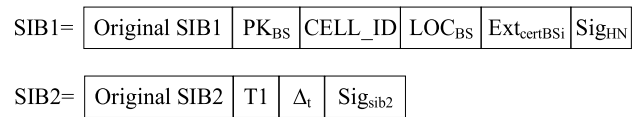


FIGURE 17. Modified SIB message.

On the one hand, our proposed scheme reduces the number of network entities and reduces the complexity comparing with [47]. On the other hand, it reduces the number of public/private key pairs required by the authentication scheme, and fully utilizes the existing PKI mechanism of the 5G network without additional overhead. It is also convenient for operators to deploy quickly.

VII. CONCLUSION

In this paper, we present a systematic approach to analyzing the 5G NAS signalling security. The symbolic model checking method was used to formally analyze the 5G NAS layer protocols, such as registration, authentication, identification, deregistration, security mode command procedures, etc. Through verification of the abstracted properties, we discovered 10 new attacks and verified them on the testbed. Moreover, we proposed a countermeasure to the vulnerabilities utilizing the PKI mechanism to essentially eliminate the connection between UE and untrusted gNodeB.

Our experiments showed that although the 5G specifications have made great progress in respect to security there are still many security issues. In the study, we also found that

TABLE 4. NAS layer message security verification results.

ID	NAS message name	Encryption	Integrity protection	Replay protection	Relay protection
1	Registration request	×	×	×	×
2	Registration accept	√	√	√	×
3	Registration complete	√	√	√	×
4	Registration reject	×	×	×	×
5	Deregistration request (UE originating)	×	×	×	×
6	Deregistration accept (UE originating)	×	×	×	×
7	Deregistration request (UE terminated)	√	√	√	×
8	Deregistration accept (UE terminated)	×	×	×	×
9	Service request	×	×	×	×
10	Service reject	×	×	×	×
11	Service accept	√	√	√	×
12	Configuration update command	√	√	√	×
13	Configuration update complete	√	√	√	×
14	Authentication request	×	×	√	×
15	Authentication response	×	×	×	×
16	Authentication reject	×	×	×	×
17	Authentication failure	×	×	×	×
18	Authentication result	×	×	×	×
19	Identity request	×	×	×	×
20	Identity response	√	×	×	×
21	Security mode command	√	×	×	×
22	Security mode complete	√	√	√	×
23	Security mode reject	×	×	×	×
24	5GMM status	√	√	√	×
25	Notification	√	√	√	×
26	Notification response	√	√	√	×
27	UL NAS transport	√	√	√	×
28	DL NAS transport	√	√	√	×

there is no integrity protection of user plane data in the case of ng-eNB, which provides 4G subscribers with user plane and control plane protocols and functions, according to the latest 3GPP specifications (TS33.501-6.10.4, V16.4.0). This may be exploited by an attacker to launch an attack. It should be noted that the 5G network has not been deployed on a

large scale and 3GPP is still developing the specifications of Release 16. Therefore, we didn't report the aforementioned vulnerabilities to GSMA or 3GPP according to the responsible disclosure procedure. Compared to the cost of upgrades on large-scale deployment of 5G networks in the future, analyzing and solving security problems from 5G standard level will be much cheaper at this stage.

In the future, we will continue to model the current 5G system and perform a more detailed analysis of the protocols, such as considering the instantiation of SQN, or using a stronger adversary model (such as eCK). Furthermore, we will discuss the security issues of RRC signalling.

APPENDIX

See Table 4.

ACKNOWLEDGMENT

The authors would like to thank Simon Meier and his team for their efficient TAMARIN Prover.

REFERENCES

- [1] 3GPP. (2017). *System Architecture Milestone of 5G Phase 1 is Achieved*. [Online]. Available: http://www.3gpp.org/news-events/3gpp-news/1930-sys_architecture
- [2] 3GPP. (2018). *Webinar-Working Towards Full 5G in Rel-16*. [Online]. Available: http://www.3gpp.org/news-events/3gpp-news/1966-webinar2_ran
- [3] *Global Mobile Trends*, GSMA, London, U.K., 2017.
- [4] P. Paganini. (2014). *Op AURORAGOLD—NSA Hacks Cellphone Networks Worldwide*. [Online]. Available: <http://securityaffairs.com/wordpress/30813/intelli-gence/op-auroragold-nsa-hacks-cellphone.html>
- [5] *Hackers Take Down the Most Wired Country in Europe*. Accessed: Sep. 4, 2019. [Online]. Available: <https://www.wired.com/2007/08/ff-estonia/>
- [6] P. P. C. Lee, T. Bu, and T. Woo, "On the detection of signaling DoS attacks on 3G/WiMax wireless networks," *Comput. Netw.*, vol. 53, no. 15, pp. 2601–2616, Oct. 2009.
- [7] R. Bassil, I. H. Elhadj, A. Chehab, and A. Kayssi, "Effects of signaling attacks on LTE networks," in *Proc. 27th Int. Conf. Adv. Inf. Netw. Appl. Workshops*, Mar. 2013, pp. 499–504.
- [8] W. K. Leong, A. Kulkarni, Y. Xu, and B. Leong, "Unveiling the Hidden Dangers of Public IP locations in 4G/LTE Cellular Data Networks," in *Proc. ACM Workshop Mobile Comput. Syst. Appl. (HotMobile)*, Feb. 2014, Art. no. 16.
- [9] G. Kambourakis, C. Kolias, S. Gritzalis, and J. H. Park, "DoS attacks exploiting signaling in UMTS and IMS," *Comput. Commun.*, vol. 34, no. 3, pp. 226–235, Mar. 2011.
- [10] D. Yu and W. Wen, "Non-access-stratum request attack in E-UTRAN," in *Proc. IEEE Comput., Commun. Appl. Conf. (ComComAp)*, Jan. 2012, pp. 48–53.
- [11] M. Lichtman, J. H. Reed, T. C. Clancy, and M. Norton, "Vulnerability of LTE to hostile interference," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Dec. 2013, pp. 285–288.
- [12] J. Xiao, X. Wang, Q. Guo, H. Long, and S. Jin, "Analysis and evaluation of jammer interference in LTE," in *Proc. ACM Int. Conf. Innov. Comput. Cloud Comput. (ICCC)*, Dec. 2013, p. 46.
- [13] J. R. Rao, P. Rohatgi, H. Scherzer, and S. Tinguely, "Partitioning attacks: Or how to rapidly clone some GSM cards," in *Proc. IEEE Symp. Secur. Privacy*, May 2002, pp. 31–41.
- [14] Y. Zhou, Y. Yu, F.-X. Standaert, and J.-J. Quisquater, "On the need of physical security for small embedded devices: A case study with COMP128-1 implementations in SIM cards," in *Proc. Int. Conf. Financial Cryptogr. Data Secur.* Berlin, Germany: Springer, 2013, pp. 230–238.
- [15] J. Liu, Y. Yu, F.-X. Standaert, Z. Guo, D. Gu, W. Sun, Y. Ge, and X. Xie, "Small tweaks do not help: Differential power analysis of MILENAGE implementations in 3G/4G USIM cards," in *Proc. Eur. Symp. Res. Comput. Secur. (ESORICS)*. Berlin, Germany: Springer, 2015, pp. 468–480.

- [16] M. Briceno, I. Goldberg, and D. Wagner. (2012). *GSM Cloning*. [Online]. Available: <http://www.isaac.cs.berkeley.edu/isaac/gsm.html>
- [17] M. Briceno, I. Goldberg, and D. Wagner. (1998). *An Implementation of the GSM A3A8 Algorithm. (Specifically, COMP128)*. [Online]. Available: <http://www.scard.org/gsm/a3a8.txt>
- [18] U. Meyer and S. Wetzel, "On the impact of GSM encryption and man-in-the-middle attacks on the security of interoperating GSM/UMTS networks," in *Proc. IEEE 15th Int. Symp. Personal, Indoor Mobile Radio Commun.*, Sep. 2004, pp. 2876–2883.
- [19] C.-K. Han and H.-K. Choi, "Security analysis of handover key management in 4G LTE/SAE networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 2, pp. 457–468, Feb. 2014.
- [20] E. Barkan, E. Bihant, and N. Keller, "Instant ciphertext-only cryptanalysis of GSM encrypted communication," *J. Cryptol.*, vol. 21, no. 3, pp. 392–429, 2008.
- [21] M. Arapinis, L. Mancini, E. Ritter, M. Ryan, N. Golde, K. Redon, and R. Borgaonkar, "New privacy issues in mobile telephony: Fix and verification," in *Proc. ACM Conf. Comput. Commun. Secur. (CCS)*, Oct. 2012, pp. 205–216.
- [22] R. Borgaonkar, L. Hirshi, S. Park, A. Shaik, A. Martin, and J.-P. Seifert, "New adventures in spying 3G & 4G users: Locate, track, monitor," BlackHat, San Francisco, CA, USA, Tech. Rep., 2017. [Online]. Available: <https://www.blackhat.com/docs/us-17/wednesday/us-17-Borgaonkar-New-Adventures-In-Spying-3G-And-4G-Users-Locate-Track-And-Monitor.pdf>
- [23] D. Rupprecht, K. Kohls, T. Holz, and C. Pöpper, "Breaking LTE on layer two," in *Proc. IEEE Symp. Secur. Privacy (SP)*, 2019, pp. 1–16.
- [24] S. R. Hussain, M. Echeverria, O. Chowdhury, N. Li, and E. Bertino, "Privacy attacks to the 4G and 5G cellular paging protocols using side channel information," in *Proc. NDSS*, 2019, pp. 1–15.
- [25] 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.
- [26] H. Kim, J. Lee, E. Lee, and Y. Kim, "Touching the untouchables: Dynamic security analysis of the LTE Control plane," in *Proc. IEEE Symp. Secur. Privacy*, 2019, pp. 1–16.
- [27] 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. Symp. Netw. Distrib. Syst. Secur. (NDSS)*, 2016, p. 15.
- [28] C. Liu, X. Ji, J. Wu, and X. Qin, "A proactive defense mechanism for mobile communication user data," *Sci. China Inf. Sci.*, vol. 61, no. 10, 2018, Art. no. 109303.
- [29] *3G Security: Security Architecture*, document TS 33.102, 3GPP, 2014.
- [30] *3GPP System Architecture Evolution (SAE); Security Architecture*, document TS 33.401, 3GPP, 2011.
- [31] *Study on the Security Aspects of the Next Generation System*, document TR 33.899, 3GPP, 2017.
- [32] *Security Architecture and Procedures for 5G System (Release 15)*, document TS 33.501, 3GPP, 2019.
- [33] S. R. 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, pp. 1–15.
- [34] R. Borgaonkar, L. Hirschi, S. Park, and A. Shaik, "New privacy threat on 3G, 4G, and upcoming 5G AKA protocols," in *Proc. Privacy Enhancing Technol. (PETS)*, 2019, pp. 1–20.
- [35] D. Basin, J. Dreier, L. Hirschi, S. Radomirovic, R. Sasse, and V. Stettler, "A formal analysis of 5G authentication," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Oct. 2018, pp. 1383–1396.
- [36] A. Koutsos, "The 5G-AKA authentication protocol privacy," Apr. 2019, *arXiv:1811.06922v2*. [Online]. Available: <https://arxiv.org/pdf/1811.06922.pdf>
- [37] M. A. Ferrag, L. Maglaras, A. Argyriou, D. Kosmanos, and H. Janicke, "Security for 4G and 5G cellular networks: A survey of existing authentication and privacy-preserving schemes," *J. Netw. Comput. Appl.*, vol. 101, pp. 55–82, Jan. 2018.
- [38] D. Rupprecht, A. Dabrowski, T. Holz, E. Weippl, and C. Pöpper, "On security research towards future mobile network generations," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 2518–2542, 3rd Quart., 2018.
- [39] C. Cremers and M. Dehnel-Wild, "Component-based formal analysis of 5G-AKA: Channel assumptions and session confusion," in *Proc. Symp. Netw. Distrib. Syst. Secur. (NDSS)*, Feb. 2019, pp. 1–15.
- [40] D. Fang, Y. Qian, and R. Q. Hu, "Security for 5G mobile wireless networks," *IEEE Access*, vol. 6, pp. 4850–4874, 2017.
- [41] R. P. Jover and V. Marojevic, "Security and protocol exploit analysis of the 5G specifications," *IEEE Access*, vol. 7, pp. 24956–24963, 2019.
- [42] R. P. Jover, "The current state of affairs in 5G security and the main remaining security challenges," Apr. 2019, *arXiv:1904.08394*. [Online]. Available: <https://arxiv.org/abs/1904.08394>
- [43] X. Ji, K. Huang, L. Jin, H. Tang, C. Liu, Z. Zhong, W. You, X. Xu, H. Zhao, J. Wu, and M. Yi, "Overview of 5G security technology," *Sci. China Inf. Sci.*, vol. 61, no. 8, Aug. 2018, Art. no. 081301.
- [44] X. Hu, C. Liu, S. Liu, W. You, and K. Qiao, "Overview of mobile communication network authentication," *Chin. J. Netw. Inf. Secur.*, vol. 4, no. 12, 2018, Art. no. 2018096.
- [45] S. Meier, B. Schmidt, C. Cremers, and D. Basin, "The TAMARIN prover for the symbolic analysis of security protocols," in *Proc. Int. Conf. Comput. Aided Verification*. Berlin, Germany: Springer, 2013, pp. 696–701.
- [46] D. F. Kune, J. Koelndorfer, N. Hopper, and Y. Kim, "Location leaks over the GSM air interface," in *Proc. Symp. Netw. Distrib. Syst. Secur. (NDSS)*, Feb. 2012, pp. 1–19.
- [47] S. R. Hussain, M. Echeverria, A. Singla, O. Chowdhury, and E. Bertino, "Insecure connection bootstrapping in cellular networks: The root of all evil," in *Proc. 12th Conf. Secur. Privacy Wireless Mobile Netw.*, May 2019, pp. 1–11.



XINXIN HU received the B.E. degree from the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan, China, in 2017. He is currently pursuing the M.S. degree with the National Digital Switching System Engineering and Technological Research Center (NDSC), Zhengzhou, China. His major research interests include mobile network security and next-generation mobile networks.



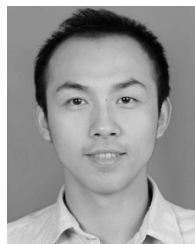
CAIXIA LIU received the Ph.D. degree in information and communication system from PLA Information Engineering University, Zhengzhou, China. She is currently a Professor and a Supervisor of postgraduate student with National Digital Switching System Engineering and Technological Research Center, Zhengzhou. Her major research interests include wireless mobile communication networks, information secrecy, and novel network architecture.



SHUXIN LIU received the B.E. degree in communication engineering from the Information Engineering University of PLA, Zhengzhou, China, in 2009, and the M.S. and Ph.D. degrees from the National Digital Switching System Engineering and Technological Research Center (NDSC), Zhengzhou, in 2012 and 2016, respectively. He is currently an Assistant Research Fellow with NDSC and the Director of the Laboratory of Network Architecture and Signaling Protocol Analysis. His research interests include network evolution, link prediction, network behavior analysis, and communication network security.

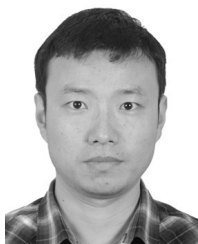


WEI YOU received the B.E. and M.S. degrees in communication engineering from the Information Engineering University of PLA, Zhengzhou, China, in 2006 and 2009, respectively, and the Ph.D. degree from National Digital Switching System Engineering and Technological Research Center (NDSC), Zhengzhou, in 2013, where he is currently a Lecturer. His major research interests include new-generation mobile communication systems and mobile communication network security.



YU ZHAO received the B.E. degree in communication engineering from the Information Engineering University of PLA, Zhengzhou, China, in 2007, and the Ph.D. degree from National Digital Switching System Engineering and Technological Research Center (NDSC), Zhengzhou, in 2017, where he is currently an Assistant Research Fellow. His research interests include social networks and telecommunication network security.

...



YINGLE LI received the B.E. degree in communication engineering from the Information Engineering University of PLA, Zhengzhou, China, in 2008, and the M.S. degree from National Digital Switching System Engineering and Technological Research Center (NDSC), Zhengzhou, in 2013, where he is currently an Assistant Research Fellow. His research interests include network evolution, link prediction, network behavior analysis, and communication network security.