5G Security

Alessandro Castelli
ID:147073
E-mail: castelli.alessandro@spes.uniud.it

October 25, 2024

Abstract

In recent decades, wireless communications have undergone rapid evolution, driven by the growing demand from users for faster, more reliable, and high-performance connections. Among the most significant technological innovations, 5G technology has emerged as the new frontier in mobile telecommunications, promising greater bandwidth capacity, lower latency, and an enormous density of connections for smart devices and IoT. However, along with these extraordinary capabilities, new security challenges also arise. 5G introduces a more complex and decentralized network infrastructure, increasing the risk of vulnerabilities and cyberattacks. This article examines the key aspects of 5G security, highlighting network vulnerabilities and discussing advanced technological solutions developed to mitigate these risks.

Contents

1 Introduzione **3**

2 Threat model **4**

3 Security goals **6**

4 Security service and implementation **7**

5 Attacks and Vulnerabilities **9**
5.1 Threats to the Radio Access Network 9
5.2 Core Network Security Threats 10
5.3 Countermeasures 12
5.4 GUTI Reallocation Command Attack 14
5.4.1 Difficulties in Wide Network Scenarios 15
5.4.2 Operator Responsibilities and Countermeasure 15
5.5 Security Capabilities Bidding-Down Attack 15

6 Conclusions **16**

A Terms to Remember **17**

1 Introduzione

The history of mobile communication is a journey of continuous innovation. It all began with 1G [2] in the 1970s and 1980s when data transmission was analog. It was the beginning, but with significant limitations: the quality was low, security was non-existent, and calls could be easily intercepted. However, it enabled something revolutionary for that time: mobile connectivity and the first voice services, albeit in a rudimentary form.

With the arrival of 2G in 1991, the situation changed drastically. Communication became digital, addressing many of the issues of 1G. Now there was greater security, efficiency, and increased bandwidth. This opened the door to new services such as text messaging, making mobile communication more sophisticated.

Subsequently, 3G introduced the concept of mobile broadband. It was no longer just about calls or messages, but also about video calls, internet browsing, and much faster data transmission. 3G represented a breakthrough, although it still suffered from spectrum and latency issues, demonstrating that the technology still had room for improvement.

When 4G arrived between 2009 and 2010, the world of mobile communication made a tremendous leap forward. Thanks to technologies like LTE, data speeds increased significantly, bringing mobile streaming, online gaming, and high-definition videos within reach anywhere. Theoretical download speeds could reach up to 100 Mbps, although in practice they were often lower.

Finally, with the advent of 5G, we entered a new era. It was no longer just about incremental improvements but about a revolution. 5G promises speeds of up to 20 Gbps [4], ultra-low latency, and the ability to connect billions of devices simultaneously. It is a technology designed for an interconnected world, where there are not only smartphones but also cars, IoT devices, smart cities, and automated industrial systems. 5G is the foundation of what will be the Internet of Everything, where every aspect of daily life is connected and integrated into a global network.

Il **5G** thus represents the latest evolution in mobile communication technology, bringing significant improvements such as *high bandwidth* and *extremely low latency*. This technology supports advanced applications like augmented reality (AR), virtual reality (VR), and ultra-reliable low-latency communications (URLLC). In March 2018, the 3GPP released the 15th version of the mobile communication standards, laying the groundwork for the **5G**. The transmission speed of this new technology allows users to enjoy significantly higher data transfers, particularly for applications that require high throughput, such as high-definition video streaming [4].

The reduction of latency is another key goal, with the expectation of latency lower than 1 millisecond, paving the way for real-time use of critical applications such as telemedicine and autonomous driving. Moreover, 5G enables the simultaneous connection of a much larger number of devices compared to previous generations, a fundamental feature given the continuous growth of the IoT market.

Thanks to all this, 5G will become a foundation for a network that connects not only people but also objects, devices, and machines. We are talking, therefore, about a system that is revolutionizing the way we envision the internet—not merely as an exchange of data between people, but as a massive integration between humans and machines, and among machines themselves.

And here come into play the three fundamental scenarios of 5G: **eMBB, mMTC,** and **uRLLC**. Each of these represents a facet of the overall vision of 5G [5]: **eMBB** is designed for enhanced mobile broadband, allowing for ultra-fast downloads and high-quality streaming; **mMTC** focuses on massive machine-type communication, essential for supporting the IoT (Internet of Things); while **uRLLC** is crucial for applications requiring minimal latency and extreme reliability, such as tele-surgery or autonomous vehicles. But what are the technologies that actually make all this possible?
The 3GPP has defined more than 70 types of 5G SA1 files necessary for this purpose, and the key technologies developed for 5G include: *Massive MIMO*, *filter bank based multicarrier (FBMC)*, *Full Duplex*, *Ultra Dense networking (UDN)*, *software-defined networking (SDN)*, and *network function virtualization (NFV)* [5].
This article will present an in-depth study on security within 5G networks. Section 2 will examine the assets, attackers, and risks associated with a 5G network. Section 3 will outline the security objectives intended to be achieved for the identified assets. Section 4 will analyze the security services implemented within this technology. Finally, Section 5 will discuss the potential vulnerabilities present. To keep the text clear and concise, avoiding excessive complications and frequent digressions into the various technologies and vulnerabilities, a glossary has been included at the end of the document. This glossary provides explanations of protocols, acronyms, and other relevant terms, thereby facilitating the understanding of the content.

2 Threat model

In the literature, numerous security challenges related to 5G have been identified. The main assets involved include sensitive user data, such as personal information and browsing data, which transit through the 5G network. In addition, there is the network infrastructure itself, composed of base stations, servers, and networking devices, as well as end-user devices, including smartphones, tablets, and IoT devices, which represent an essential part of the system. It is important to note that, in addition to end devices, the network control systems responsible for traffic management and authentication require adequate protection to ensure the integrity and availability of the network.

5G networks introduce various innovative technologies, such as *software-defined networking (SDN)* and *network function virtualization (NFV)*, which offer significant advantages in terms of flexibility and scalability. However, these technologies also expose the network to new security risks, such as resource exhaustion and vulnerabilities in programming interfaces, which can become tar-

gets for targeted attacks. Additionally, *Massive Multiple-Input Multiple-Output (MIMO)* and *millimeter-wave communications (mmWave)* increase network capacity, but they must address security issues related to resource management and information confidentiality. Cloud computing and *Multi-access Edge Computing* (MEC) can also contribute to improving network efficiency, but the storage and processing of data in the cloud increase the risk of attacks on sensitive data [1].

The potential attackers in this context can vary significantly. On one hand, there are individual hackers seeking to intercept or manipulate data for illicit purposes. On the other hand, there are organized cyber-criminal groups, and in some cases, even state-sponsored actors, whose objective might be to gain unauthorized access to sensitive information or cause damage to the network for economic or geopolitical reasons. Adding complexity to the picture are also internal malicious actors, meaning personnel with legitimate access to the network, who could abuse their permissions. In this context, it is crucial to recognize that vulnerabilities do not only come from the outside, but also from misconfigurations and inadequate security practices by authorized personnel. The attack vectors through which these actors can strike are numerous. The security of radio interfaces could be compromised if such protocols are not implemented correctly or if complex vulnerabilities are exploited. This paves the way for attacks such as communication interception or classic man-in-the-middle attacks. The integrity of the user plane also represents a critical point: despite the introduction of advanced encryption in 5G, targeted attacks could exploit any flaws in the data protection process, thereby compromising the integrity of the transmitted data without being detected.

During roaming, another critical moment for security, the protection parameters of users may not be updated correctly when transitioning between different networks, exposing users to potential data interception or manipulation attacks. This issue highlights the need for proper synchronization and updating of security measures across different networks to ensure continuous and consistent protection.

The 5G infrastructure is also not immune to DoS (Denial of Service) attacks. Despite advances in protective measures, the network control systems remain visible and can be vulnerable to service disruptions, especially if the control channels are not adequately encrypted. Here, the management of credentials and configurations becomes essential to prevent unauthorized access.

Finally, an important attack vector is represented by end-user devices. Often, these devices lack sufficient security measures at the operating system and application level, making them vulnerable to malware, DoS attacks, or manipulation of configuration data, thereby compromising the entire network. This vulnerability is exacerbated by the increasing proliferation of IoT devices, which may have intrinsic vulnerabilities due to inadequate designs or a lack of regular updates.

In conclusion, the threat landscape in the 5G network is complex and varied, requiring a multifactorial approach for asset protection and risk mitigation. Special attention must be paid not only to security technologies and protocols

but also to the continuous training of personnel and the awareness of end-users regarding risks and best practices for security.

3 Security goals

Every communication system, in order to be considered secure, needs protection protocols and technologies that safeguard the resources used in communication, ensuring that the data complies with the properties of: **Confidentiality**, **Integrity**, **Availability**, **Authenticity**, **Accountability** [8].

These objectives are essential to ensure that 5G networks can operate securely and reliably, protecting users' data and communications. Let's start with Confidentiality. This objective aims to ensure that only authorized users can access sensitive information. In 5G networks, where the volume of transmitted data is enormous and includes personal and business information, it is crucial to implement robust encryption and authentication measures. Confidentiality is not just a matter of data protection, but also of users' trust in the system. If users perceive that their data is not secure, they may be reluctant to use the services offered.

Moving on to Integrity, this objective focuses on protecting data from unauthorized modifications. In a 5G environment, where IoT devices and critical applications are increasingly interconnected, it is essential to ensure that information remains accurate and is not altered during transmission.

Availability is another key objective. 5G networks must be always available to ensure that users can access services at any time. This is particularly important for critical applications such as those in the healthcare or transportation sectors. To achieve this objective, it is necessary to implement redundancy and resilience solutions to mitigate the effects of attacks such as Distributed Denial of Service (DDoS), which can compromise the availability of services.

Authenticity is essential to ensure that the entities involved in communications are who they claim to be. In a 5G context, where interactions occur between a variety of devices and users, it is crucial to implement robust authentication mechanisms. This not only protects assets from unauthorized access but also helps to build a trust ecosystem among the various actors involved.

Finally, **Accountability** is crucial to ensure that all actions and transactions within the network can be traced and verified. This objective is particularly important for regulatory compliance and for managing responsibilities. Implementing logging and monitoring systems allows for the detection of suspicious activities and enables a rapid response to potential security incidents.

In summary, the CIAAA model provides a comprehensive framework for addressing security challenges in 5G networks. Each objective is interconnected and contributes to creating a secure and reliable environment for users and applications. Protecting assets in a 5G context requires a holistic approach that integrates advanced technologies, strict security policies, and continuous vigilance to address emerging threats.

4 Security service and implementation

The architecture of 5G is organized into 3 layers: an application layer, a service layer, and a transport layer [6]. Each layer is designed with specific security functionalities that, when combined, create a secure and resilient system against threats. The main security elements in 5G include:

- **Network access security**: Mechanisms that allow a user device (UE) to authenticate and securely access network services. The UE exchanges protocol messages through the access network with the service network (SN). Cryptographic keys are stored in the `USIM` module of the device and in the operator's environment. This ensures the security of data and communications, preventing unauthorized access.
- **Network domain security**: A set of features that enable network nodes to securely exchange control plane and user plane data within 3GPP networks and across different networks. Protection techniques include encryption and integrity measures to prevent interception and tampering.
- **User domain security**: Focuses on protecting the user's device and the data contained within it, preventing unauthorized access to the mobile terminal. Hardware mechanisms are implemented to protect the `USIM` module and prevent tampering with terminals, ensuring user authenticity.
- **Service-based architecture domain security**: Protects the registration, discovery, and authorization of network elements, as well as service-based interfaces. It enables the secure integration of new virtual network functions in 5G and supports secure roaming, involving both the service network and the home network.
- **Visibility and configurability of security**: Allows users to be informed about the presence of security functions and offers the possibility to configure security features according to their needs. The security specifications of 5G, defined by 3GPP, establish optional functionalities, providing degrees of freedom for the secure implementation and operation of the network.

The security procedures of 5G are based on a hierarchical derivation framework. The long-term key `K` is stored by the `Authentication Credential Repository and Processing Function` (ARPF) while the USIM holds the corresponding copy of that user symmetric key [6]. All other keys are derived from it.

The 3GPP introduced the `Extensible Authentication Protocol (EAP)` for Authentication and Key Agreement, defining `EAP-AKA` and `5G AKA` as mandatory authentication methods for devices (UE) and the network. These protocols ensure mutual authentication between the device and the network, as well as protecting the security and encryption of services. During registration, a 5G device sends the SUCI to initiate the authentication process based on the selected protocol.

The 5G security specifications define various security contexts for different situations: for a single 5G service network (SN), between multiple SNs, and between 5G and 4G networks. When a device is connected to two SNs, each network must independently manage and utilize its own security context. In the case where the device is registered on two SNs within the same Public Land Mobile Network (PLMN), whether they are 3GPP or non-3GPP, the device establishes two separate Non-Access Stratum (NAS) connections for each network, but shares a common NAS security context, which includes a single set of keys and security algorithms.

The procedures for maintaining or discarding a security context during state transition state that the configuration of the handover type is at the operator's discretion, based on individual security requirements. Consequently, security during handover becomes an optional function rather than a mandatory one, which may lead some operators to implement potentially insecure handover procedures.

Cryptographic separation and protection against replay attacks for two active NAS connections are ensured through a shared NAS security context, with distinct parameters for each connection. The NAS employs 128-bit encryption algorithms to ensure the integrity and confidentiality of the data. It is important to note, however, that options for null encryption and integrity protection are also provided. Additionally, if the device does not have a NAS security context, the initial NAS message is transmitted in clear text, including the subscriber identifier and the device's security capabilities.

In the control of radio resources, integrity and confidentiality are ensured by the PDCP (Packet Data Convergence Protocol) layer that operates between the device and the gNB. It is important to note that no layer below the PDCP is subject to integrity protection. Protection against replay attacks is activated when integrity protection is in function, except in cases where null integrity protection is selected. RRC integrity checks are performed both on the device and the gNB, and if an integrity check fails after protection has been activated, the corresponding message is immediately discarded.

Moving to the user plane, the Session Management Function (SMF) is responsible for providing the security policy for a PDU (Protocol Data Unit) session to the gNB during the session establishment phase. If integrity protection is not activated for the Data Radio Bearers (DRB), neither the gNB nor the device will be able to protect the integrity of the traffic of such DRBs. Similarly, if user plane encryption is not activated for the DRBs, the traffic will not be encrypted. The local SMF has the option to override the confidentiality option present in the user plane security policy received from the Home Network (HN) SMF.

Finally, regarding the privacy of the subscription ID, the SUCI represents the concealed version of the permanent subscription identifier of 5G (SUPI). This is transmitted over the air to avoid exposing the user's identity in clear text. The SUCI is generated from the SUPI using the operator's public key and a probabilistic asymmetric encryption method, which helps to prevent identity tracking. However, the null protection system of the SUPI is used during unau-

thenticated emergency sessions if configured by the Home Network (HN), or when the operator's public key has not been provided. The 5G specifications also define a temporary identifier, the 5G Globally Unique Temporary Identifier (5G-GUTI), to reduce the exposure of the SUPI and the SUCI. The 5G-GUTI must be reassigned based on device triggers, but the frequency of such reassignment is left to the discretion of the network implementation.

5 Attacks and Vulnerabilities

In this section, the vulnerabilities present in 5G networks will be analyzed, taking into account the differences related to the architectures used. 5G networks can be divided into two main categories: the Non-Standalone 5G (NSA), a transitional architecture that relies on 4G LTE infrastructure for signaling and control management, and the Standalone 5G (SA), which represents a completely autonomous network. It is essential to understand that vulnerabilities can manifest in both architectures; however, some may be more common or relevant in one than the other, depending on the specific implementations and technologies used. Below, some of the main vulnerabilities that can be easily exploited in attack contexts applicable to both NSA and SA networks will be illustrated. I will begin by examining the weaknesses that are more likely to occur in NSA architectures, before addressing more general vulnerabilities.

5.1 Threats to the Radio Access Network

Security threats to the Radio Access Network (RAN) focus on the vulnerabilities of the radio part of the network, which includes base stations and mobile devices.

Information Leakage: Information loss can include threats such as the interception of paging and the cracking of `IMSI`. Paging interception is a passive scanning technique that exploits the paging messages sent by base stations to mobile terminals. An attacker, using a `SDR device` (software-defined radio), can capture RF signals in the vicinity of the victim and detect the temporary identity of the mobile subscriber (S-TMSI) or the paging cycle. By exploiting this information, the attacker can calculate the Paging Frame Index (PFI), limiting the possible IMSI candidates for the victim to only 8. Subsequently, the attacker sends an IMSI paging with these 8 candidates and observes the responses to identify the victim's actual IMSI.

Denial of Service (DoS) for the User: This type of attack includes various forms, such as DoS on the `RRC` (Radio Resource Control) connection, the rejection DoS `RRC`, and the release DoS `RRC`. The DoS on the `RRC` connection is particularly severe, as it exploits the `S-TMSI` (SAE-Temporary Mobile Subscriber Identity) of the victim, which was previously obtained through an information leakage. Base stations do not implement authentication procedures for terminals, allowing attackers to interfere with the victim's wireless access. The

attacker can send a connection request message RRC using the victim's S-TMSI value, leading to the disconnection of the victim's RRC connection and establishing a connection with the attacker, who can then continue to send requests to prevent legitimate access to the service.

DoS on the Base Station: A DoS attack can also target the depletion of the base station's resources, making it difficult for legitimate users to connect. When a terminal attempts to establish a connection, it uses the RRC protocol, which involves a random access process. Attackers can exploit this process to send unauthorized requests, increasing the number of active RRC connections and overloading the base station, causing delays or disruptions in services.

Eavesdropping: In theory, intercepting wireless networks should be impossible due to the security settings between terminals and base stations. However, there are cases where it is possible to extract the security key stream and decrypt communications. Voice traffic in mobile communications uses the RTP (Real-time Transport Protocol) and is transmitted via a voice bearer, distinct from the data bearer. To ensure Quality of Service (QoS), the voice bearer is dedicated, with a separate QoS Class Identifier (QCI). During the creation of the key stream for encryption in the AS security procedure, four parameters are used: `count, direction, length, and bearer ID`. Of these, only the bearer ID (DRB ID) represents a critical variable. The DRB ID is assigned when the base station creates a voice bearer, but some base stations from certain manufacturers may assign the same DRB ID within the same RRC connection, creating a vulnerability. An attacker can exploit this flaw by intercepting the victim's encrypted voice communication using a sniffer. After the call is over, the attacker can make a voice call to the victim, intercepting both the clear and the encrypted calls. By applying the XOR operation between the plaintext and ciphertext of the second call, the attacker can extract the key stream, which can be used to decrypt the first call since both used the same DRB ID and occurred in the same RRC connection. To prevent this type of attack, the 3GPP TS 33.401 standard recommends using different DRB IDs for various bearers, thus preventing the reuse of the same DRB ID in the base station.

Unauthorized Data Usage: In mobile networks, there are predefined and dedicated bearers designed for legitimate communications. An attacker can exploit these bearers to access data unauthorizedly, for example, by establishing cost-free communications. Additionally, caller masquerading, also known as caller spoofing, is an attack in which the attacker falsifies the caller's identity, deceiving the victim.

5.2 Core Network Security Threats

Information Leakage: Information related to 5G NSA core networks can essentially be classified into two categories: that concerning the EPC (Evolved Packet Core) equipment, dedicated to data processing, and that related to the IMS (IP Multimedia Subsystem) equipment, which provides various services.

Since the EPC equipment uses the GTP protocol for communications, while the IMS equipment relies on the SIP (Session Initiation Protocol), an attacker has the option to choose the most appropriate protocol to access the desired information. The GTP protocol is divided into two components: GTP-C, used for communication between core network equipment, and GTP-U, responsible for transporting data traffic from the user terminal through a tunnel between the base station and the PGW (Packet Gateway). To extract information about the EPC equipment's IP address, the attacker can resort to a packet injection method, embedding an echo request in the payload of the data along with a GTP-C message for monitoring the health status of the network equipment. By using a program called Packit, the Android Debug Bridge (ADB) command can be executed in the Android terminal, thus generating a packet. By sending this packet to the IP address obtained through the Tracert command in tethering mode, the GTP-C packet is injected and transmitted to the mobile communication network. Once the PGW receives the packet, it verifies it and sends an echo response, allowing the attacker to identify the source IP of that message as PGWIP.

IP Address Exhaustion: By using the GTP-in-GTP technique, an attacker can exhaust the pool of IP addresses in the network by sending session requests with incremented terminal numbers, thereby preventing legitimate terminals from connecting. This technique exploits a vulnerability in the GTP (GPRS Tunneling Protocol), which is used for data transport in mobile networks. GTP-in-GTP involves encapsulating GTP packets within other GTP packets, allowing the attacker to send multiple requests in a single connection. The attacker creates GTP sessions within other GTP sessions, amplifying the attack and increasing the number of requests generated, thereby accelerating the depletion of network resources. This technique proves particularly effective as it takes advantage of the overhead associated with managing GTP sessions, causing a strain on the IP address allocation system. [0.2cm]

Denial of Service (DoS): A DoS attack can be executed by repeatedly sending connection request messages to the 5G NSA network, overwhelming the core network and causing service disruption.

NAS Manipulation: Messages from the NAS protocol, used for signaling between terminals and the core network, are not always protected by encryption. An attacker can exploit this vulnerability by installing a malicious base station to manipulate messages and alter critical parameters for data encryption and integrity.

Eavesdropping: Voice communication over a 5G network utilizes the IMS network and initiates sessions via the SIP protocol, following the guidelines established by the 3GPP standard. Consequently, the security of the SIP protocol is crucial and is primarily ensured through the Security Associations (SAs) of IPSec. However, the implementation of IPSec SAs is selectively managed by 5G

network operators, and support for VoLTE (Voice over LTE) does not necessarily imply full IPSec coverage due to the significant impact it could have on terminal performance. For instance, the Samsung Galaxy S10 supports IPSec but has an issue: the related setting can be disabled through a hidden menu. If an attacker gains remote access to this hidden menu and modifies the IPSec setting, the victim's voice communications will occur without encryption. Furthermore, if the EEA field is altered through NAS (Non-Access Stratum) manipulation and a NAS encryption algorithm is not used, the wireless communication in the AS (Access Stratum) section will also be unencrypted. In such a scenario, the attacker can intercept the wireless traffic using a man-in-the-middle (MitM) attack, allowing them to eavesdrop on the victim's unencrypted voice traffic.

Spoofing: IP spoofing is a common attack in which the attacker sends packets with falsified IP addresses, causing billing issues and potential DoS attacks. Additionally, spoofing the `from` field in SIP or MMS packets can be used for voice phishing, presenting falsified numbers on the receiving terminal.

5.3 Countermeasures

We can consider an approach to countermeasures against security threats to 5G on two levels: the standardization of security guidelines and detection through intrusion detection systems or intrusion prevention systems.

Denial of Service (DoS) of the RRC connection.
The main cause of the DoS threat on the RRC connection is that the RRC connection request, a message transmitted when a user terminal accesses the network, is sent in clear text, and the message includes the TMSI, a temporary identification information of the user terminal. To address this, it is necessary to verify the spoofing of RRC messages at the base station level, which is not easy to define in the 3GPP specifications. Furthermore, blocking attackers from discovering the temporary identification information of a specific user can be a way to verify, but even this is not simple, as there are too many known methods. The use of temporary identification information in the RRC connection request serves to prevent privacy invasion caused by the leakage and abuse of the IMSI, which is the subscriber identification information within the USIM. However, attackers can intercept the clear text RRC connection request messages and easily identify the TMSI. From this, the attacker can create and send a modulated RRC connection request message so that the base station confuses the message with one sent from the victim's terminal. In the core network, the temporary identification information is created at specified time intervals according to specific rules based on the IMSI, and even if the TMSI is changed, the attacker can identify the changed TMSI and recreate an attack message. When receiving the modulated RRC connection request, the base stations disconnect the existing victim terminal without verifying the modulation state and allow access to the attacker's terminal. In such a situation, if the base station does not disconnect the existing victim terminal or maintains the connection for a certain period of time, it can reduce the victim's

DoS threat. Since the attacker's terminal fails authentication after the RRC connection, maintaining the connection with the existing terminal only during the period when the attacker's terminal sends the RRC connection request and the authentication fails may not significantly affect the performance of the base station.

NAS Manipulation (NAS Encryption Spoofing).
The 5G standard from 3GPP supports both integrity verification and encrypted communication to strengthen the security of the NAS protocol between terminals and 5G core networks. However, since encrypted communication is not mandatory but an optional feature among the security functions of the NAS protocol, there are some cases where security functions are not used due to the policy of the country or the mobile operator. Besides emergency calls, some countries or operators may not utilize the encryption capabilities provided by 5G standards in terms of security. Furthermore, the 5G standard does not define mutual authentication between terminals (UE) and 5G networks, nor the integrity verification function for the initial messages exchanged before encryption, relying on an underlying assumption that presumes the initial messages are not modulated. This issue arises from a vulnerability that does not confirm the spoofing of the first access request message (attack request) sent to the 5G network by the UE. A malicious attacker infiltrates between the victim's UE and a base station, manipulates an access request message that includes a request for encryption and integrity verification normally sent by the terminal into an access request message with encryption disabled and unverified integrity, and sends it to a regular base station.

The equipment of the 5G core network (MME) forms an unencrypted channel with the UE according to the request to disable encryption, a manipulated message different from the original message sent by the UE, even if the activation of encryption is set by default. This issue occurs because the 3GPP standard does not incorporate an integrity authentication procedure to check the spoofing of the initial message before the mutual authentication phase, leading to unencrypted communication caused by unencrypted and unverified integrity manipulated UE messages.

Two threats may arise from this: the first is the manipulation of message content. There are vulnerabilities or risks where malicious attackers can manipulate the messages sent from a victim's terminal at will. This is due to the fact that the standard does not define any integrity verification procedure to check whether the access request messages sent from a terminal during the initial access to the communication network have been manipulated or not. The second threat is the interception of communications. Due to a request to disable encryption within the modulated initial access message, communication between the terminal and the 5G network will be formed as an unencrypted channel, allowing a malicious attacker to intercept all wireless communications and expose personal and location information contained in the exchanged messages. An attacker can disable encryption settings between the victim's terminal and the

core network using a fake base station to utilize the victim's data or intercept the message content. We have implemented an algorithm to detect this and conducted performance detection tests. By analyzing the encryption fields in the NAS protocol messages between terminals and core networks, it is possible to set up encryption bypass channels or determine if they are non-standard terminals.

Interception (SIP spoofing).
Since interception caused by SIP spoofing is possible when IPSec is disabled, it is necessary to ensure that the encryption settings for voice communication between terminals and the 5G network are managed by the mobile operator's network, rather than being processed according to the terminal's function (a selective application in the network is required for non-IPSec supporting terminals). If the IPSec setting for the 5G voice service is determined by the terminal's function, malicious users may attempt multiple attacks exploiting the settings of their terminal. Furthermore, efforts should be made to raise awareness and publicize the risk of communication detail leaks when attackers maliciously manipulate messages and engage in unencrypted communications between the terminal and the 5G network. The applicable section for IPSec is defined as local policy in the 3GPP, but it is necessary to consider making it mandatory at the 3GPP standard level. Consequently, we have attempted to present a vulnerability report on the issues described above to the 3GPP and have discussed these matters through collaboration with a research group. Guidelines for the proper configuration and activation of the IPSec protocol in the 5G network could be incorporated in the next revision of the 3GPP specifications.

5.4 GUTI Reallocation Command Attack

The 5G-GUTI (Globally Unique Temporary Identifier) is a temporary identifier assigned to users to ensure their privacy. In some 5G SA (Standalone) networks, this identifier is reassigned in a random and unpredictable manner. The vulnerability lies in the fact that the command for GUTI reassignment is sent without security protection, either in terms of integrity or encryption [3]. Attackers can exploit this vulnerability in various ways:

- **Man-in-the-Middle (MiTM)**: An attacker can intercept the configuration update message and modify the 5G-GUTI value. If the UE attempts to re-establish a connection using this altered value, the network will not recognize it, leading to a Denial of Service (DoS) attack.
- **GUTI Refreshment Neutralization**: This attack exploits the lack of an acknowledgment (ACK) request in the configuration update message. If the UE does not send a completion message, the attacker can block or alter the commands without consequences.
- **Victim Tracking**: The absence of encryption in the GUTI reassignment command allows attackers to track users' locations. They can send silent

messages (which do not trigger notifications) and observe the network's responses to determine the UE's presence in specific geographical areas.

5.4.1 Difficulties in Wide Network Scenarios

In a network with many connected devices, it is complicated for the attacker to isolate a single device due to the number of messages sent. However, they can improve the accuracy of the attack by sending messages at specific intervals or during low-activity times.

5.4.2 Operator Responsibilities and Countermeasure

The identified vulnerabilities stem from implementation errors in 5G networks. In particular, operators should not send NAS messages without ensuring integrity. Additionally, the lack of encryption is a design choice, according to 3GPP documents, which make encryption optional. In conclusion, to mitigate this problem, operators should send these messages ensuring their integrity through encryption.

5.5 Security Capabilities Bidding-Down Attack

The Bidding-Down Attack is a vulnerability that can manifest in certain implementations of 5G networks, particularly in non-standalone (NSA) and standalone (SA) networks of operators that do not correctly follow the security specifications defined by 3GPP, meaning they do not adequately verify the security capabilities supported by the User Equipment (UE) (i.e., they cannot distinguish between the security capabilities communicated in the initial registration request and those received in the NAS Security Mode Command (SMC) message). This attack exploits weaknesses in the security capabilities negotiation process between the user device (UE) and the network, with potential negative consequences for the confidentiality and integrity of transmitted data [3].

When a UE sends a registration request to the network, it communicates its security capabilities, specifying the supported encryption and integrity algorithms. However, an active attacker in a man-in-the-middle mode can intercept this message and modify it, replacing the communicated security algorithms with weaker ones.

Subsequently, the authentication procedure based on 5G-AKA is executed. Once the authentication is completed, the core network (CN) sends a NAS Security Mode Command (SMC) to the UE. In NSA and SA 5G network implementations, the NAS SMC command does not include a message authentication code (MAC), meaning that the UE cannot verify the integrity of the received message. Therefore, the attacker can again alter the security algorithms, causing the device to accept the weaker ones, leading to a complete compromise of security.

Thus, the Bidding-Down Attack poses a serious threat in 5G network implementations that do not adequately implement security mechanisms, highlighting

the importance of following established guidelines to ensure the protection of communications.

To avoid the Bidding-Down Attack, several measures can be taken. Firstly, it is essential to always include a message authentication code (MAC). Another important aspect is the implementation of robust verification mechanisms. It is crucial for the network to perform accurate checks on the security capabilities communicated by the UE, comparing what the device claims to support with what the network is capable of handling. This way, it is possible to detect and prevent the acceptance of weaker security algorithms. Additionally, it is fundamental to use advanced security algorithms. Networks should avoid using obsolete or less secure algorithms, opting instead for the most recent and robust ones that offer adequate protection against such attacks.

6 Conclusions

The 5G technology represents one of the latest frontiers in the field of wireless communications. It has the potential to revolutionize many human activities and our way of conceiving modern communications, thanks to high transmission speeds, low latency, and high throughput. However, with the advancement of mobile technology, new threats also develop that can compromise the security of these networks. This article has analyzed the main vulnerabilities associated with 5G. Many of these vulnerabilities are inherited from 4G, while others are completely new. Research on security and the identification of innovative solutions is a crucial task for all stakeholders involved, from network service providers to those defining communication protocols. This becomes particularly necessary, considering that 5G is increasingly fundamental for sectors of everyday life that can become extremely sensitive if not managed properly, such as telemedicine and autonomous vehicles. The evolution of cybersecurity in the context of 5G is not only a response to existing threats but also a challenge to anticipate and mitigate future risks, given the growing global interconnectedness. Key technologies such as artificial intelligence and machine learning offer significant potential to enhance network protection by identifying sophisticated threats and large-scale attacks in real-time. Additionally, cryptography must evolve to address new forms of attack, including those that could emerge from the use of quantum computers. Looking ahead, it is clear that the security of 5G will become an increasingly central aspect, especially with the increase of connected IoT devices and the use of 5G in critical infrastructures. Without adequate protection, the consequences could be devastating, jeopardizing not only user privacy but also the stability of economic and social systems. International cooperation, the development of new technologies, and the definition of increasingly stringent standards will be essential to ensure that 5G is not only a driver of progress but also a secure platform for future generations.

A Terms to Remember

- **Massive MIMO**: Wireless Multiple Input Multiple Output systems utilize multiple transmission and reception antennas to increase network capacity, enhancing data throughput and serving a larger number of users. MIMO divides the signal into low-speed sub-signals, transmitted on spatially separated antennas over the same frequency channel. Through multi-path propagation, the receiver separates the signals into parallel streams to recover the original signal. MIMO increases channel capacity without consuming additional bandwidth or power, and speed can grow by adding more antennas. In 5G, Massive MIMO technology goes beyond the 2×2 configuration of 4G, utilizing numerous simultaneous streams to enhance network capacity and spectral efficiency. The larger antenna array allows coherent signal processing, rapidly adapting to changes in the propagation channel [7].

- **mmWave**: Millimeter wave communications refer to the use of very high electromagnetic waves, typically ranging between *30 GHz and 300 GHz*. These waves are called millimeter waves because their wavelength varies between 1 mm and 30 mm, which are much shorter than the traditionally used radio waves. They enable very high speeds and low latency but have limited range and poor penetration capability, requiring dense infrastructure such as small cells and advanced technologies like beamforming. When used in conjunction with lower frequencies to ensure complete coverage, mmWave is crucial for enhancing network capacity in high-density user areas.

- **EAP-AKA**: is an authentication protocol designed to enable secure authentication between a mobile device and a network, based on the concept of symmetric keys. When a device wishes to connect to a network, it sends an authentication request using a `SUCI`, an encrypted version of its permanent identifier, the `SUPI`, allowing the network to identify the device without revealing the user's real identity and ensuring their privacy. Upon receiving the request, the network uses the SUCI to access the credentials stored in the `ARPF` database, where the secret key `K` is kept, shared between the device and the network. If the `EAP-AKA` protocol is chosen, the network initiates the authentication process by generating a `challenge` for the device consisting of three elements: a random number `(RAND)`, an authentication message `(AUTN)`, and the expected response `(XRES)`, which are sent to the device. The device verifies the `AUTN` value using the key `K` in its `USIM`, and if the `AUTN` is valid, it calculates a response `(RES)` based on `RAND` and the key `K`, then sends it to the network, which compares the received `RES` with the previously generated `XRES`; if they match, the authentication is successful and both parties have verified each other's identity. ubsequently, the session key is generated, derived from the key `K` and the materials generated in the challenge, with keys such as `CK` (for data encryption) and `IK` (for message integrity), essential for protecting

communications. Once the authentication process is complete and the session keys are derived, the device and the network can begin to exchange data securely, knowing that the communications are protected by robust security measures.

- **5G AKA**: The `5G AKA` process begins with the device sending an authentication request to the network via the `SUCI`, an encrypted version of the `SUPI`, which protects the user's identity during registration. The network receives this request and sends it to the credential database, which contains the secret key `K` shared between the device and the network. If the `5G AKA` protocol is chosen, the network generates a `challenge` consisting of a random number `(RAND)`, a temporary authentication `(AUTN)`, and an expected response `(XRES)`, which are sent to the device. The device, using the key `K` stored in the `USIM`, verifies the `AUTN` to confirm the identity of the network; if the verification is valid, it calculates a response `(RES)` based on `RAND` and `K`, which is then sent to the network. The network compares the `RES` with the previously generated `XRES`, and if they match, the authentication is successful. At this point, session keys for encryption and the integrity of communications are generated, ensuring the security of transmitted data. An important feature of `5G AKA` is the enhancement of user privacy, thanks to the separation of encryption and integrity keys and more robust mechanisms to prevent tracking and identity correlation attacks, significantly improving security compared to previous generations.

- **PLMN (Public Land Mobile Network)**: It is a public land mobile network. A PLMN is a mobile network that provides mobile communication services to the public. Each mobile operator has its own PLMN identified by a unique code, which allows devices to identify and connect to that network.

- **FBMC**: The Filter Bank Multicarrier (FBMC) is a modulation technique that divides a signal into multiple subchannels, applying filters to reduce interference between them, thus improving spectral efficiency compared to OFDM.

- **FullDuplex**: Full duplex is a mode of communication in which data can be transmitted and received simultaneously between two devices or points. In other words, both parties can send and receive information at the same time, without having to wait for one of them to finish transmitting.

- **Ultra Dense Networking**: Ultra Dense Networking (UDN) is a network architecture designed to improve network capacity and coverage in high-density user or device environments. UDN is based on the idea of increasing the number of cells or small base stations (small cells) in a geographic area, reducing the distance between these stations and the connected devices. This reduces the load on each individual base station, improving bandwidth capacity and signal quality.

- **Software-Defined Networking**: Software-Defined Networking (SDN) is an approach to network management that separates the control plane from the data plane. Traditionally, routers and switches perform both the task of routing traffic (data plane) and deciding how to do so (control plane). SDN moves the control plane to a central software entity called a `controller`, which has a global view of the network and can dynamically program how packets should be handled by the switches, simplifying network management and enhancing its agility.

- **Network Function Virtualization**: Network Function Virtualization (NFV) is a technology that virtualizes network functions, such as firewalls, routers, load balancers, and other network devices, on standard servers, eliminating the need for specialized hardware. NFV allows for the deployment and management of network functions as software, enhancing scalability, speed of implementation, and reducing costs.

- **S-TMSI**: It is a temporary identifier used in 4G LTE and 5G networks to represent a subscriber without revealing their IMSI (International Mobile Subscriber Identity). It is generated by the network and changes periodically, reducing the risk of tracking and cyber attacks.

- **RRC**: The RRC protocol is a fundamental part of the mobile network architecture, used to manage communication between the user terminal (UE) and the radio network in the context of mobile technologies. RRC handles radio resource management, ensuring that the terminal can access and utilize network resources efficiently. The protocol performs several key functions, including: configuring and managing radio connections, monitoring network conditions, managing transitions between different operating modes (for example, from idle state to connected state), and allocating the necessary resources for data transmission. During the handover process, RRC ensures service continuity by transferring control of radio resources from one cell to another as the user moves. RRC is divided into different operating modes: RRC_IDLE, in which the terminal is not actively connected to the network, and RRC_CONNECTED, where the terminal is in active communication with the network. The protocol also plays an important role in ensuring the security of communications by managing encryption and authentication keys.

- **eNB**: It represents the radio access node of `4G LTE` networks. eNB is responsible for transmitting and receiving radio signals between the end user and the `LTE` core network and has functions for radio resource management, encoding and decoding, power control, mobility management, and data scheduling.

- **gNB**: The gNB is the base station used in 5G networks and represents the evolution of the eNB from LTE networks. It plays a crucial role in providing radio connectivity to user equipment (UE) and manages communications between mobile devices and the 5G core network. The gNB

is designed to support higher capacity and lower latency, enabling advanced services such as massive IoT (Internet of Things), augmented reality, and critical communications. Unlike the eNB, the gNB can simultaneously handle multiple radio access modes, including NR (New Radio), and supports both standalone (SA) and non-standalone (NSA) configurations. The gNB communicates with the core network via the NG protocol, which includes interfaces for signaling and data transport. One of the distinguishing features of the gNB is its coordination capability with other gNBs to implement advanced techniques such as Coordinated Multi-Point (CoMP), enhancing coverage and spectral efficiency. Additionally, the gNB implements advanced security mechanisms, such as authentication and encryption, to ensure secure and protected communications.

- **IMSI**: The International Mobile Subscriber Identity (IMSI) is a unique identification number, consisting of 15 digits, associated with each subscriber in mobile networks. This number is used to identify and authenticate a user within a mobile network. The IMSI is stored in the SIM card and includes the Mobile Country Code (MCC), the Mobile Network Code (MNC), and a subscriber identification number (MSIN).

- SDR: A Software-Defined Radio (SDR) device is a radio where most of the signal processing operations, such as modulation and demodulation, are performed via software instead of fixed hardware. This allows the device to adapt to different frequencies and radio protocols simply by updating the software, making it extremely flexible and versatile for applications such as telecommunications, research, and security analysis.

- RTP: The RTP protocol is used for the transport of real-time multimedia data, such as audio and video, over IP networks, and is essential for applications like VoIP telephony and videoconferencing. RTP provides functionalities to synchronize the multimedia stream and ensure the correct sequence of packets, even when they traverse different network paths. During transmission, multimedia data is encapsulated in RTP packets that include essential information such as the timestamp (for stream synchronization) and the sequence number (for packet ordering). The protocol is often used in conjunction with RTCP (RTP Control Protocol), which monitors the quality of service by providing feedback on delay, jitter, and packet loss. RTP does not natively provide security features, so to ensure communication protection, it can be used in combination with protocols like SRTP (Secure RTP), which offers encryption and authentication. RTP is particularly suited for real-time transmissions as it is optimized to minimize delays and maintain consistent quality even over unreliable networks.

- GTP: GTP is a protocol used in mobile networks for the transport of data and signaling between network nodes. It is divided into two main components: **GTP-C** (Control Plane), which manages signaling and session

control, and **GTP-U** (User Plane), which handles the actual transport of user data. When a user connects to the network, GTP-C creates a tunnel for signaling between the core network nodes. Once the tunnel is established, GTP-U transports the user data packets by encapsulating and sending them through the network. If the user changes cells or the connection needs to be updated, GTP-C intervenes to modify the session, ensuring continuity of data transmission. At the end of the session, GTP-C closes the tunnel and terminates the connection. Although GTP is essential for efficient data transport, it may present vulnerabilities that could be exploited by attacks, necessitating the adoption of appropriate security measures.

- NAS: It manages the communication between the user's device and the core network, handling authentication and registration to ensure that only authorized devices access the network. Additionally, it deals with mobility, allowing users to move from one cell to another without interruptions. It also manages data sessions, establishing and terminating connections while maintaining the quality of service.

- MME: The Mobility Management Entity (MME) is a software component in the core network of 4G and 5G mobile networks, responsible for managing user mobility and signaling. It performs key functions such as user authentication, session management, location tracking, and routing service requests to other network entities. The MME communicates with terminal devices and other network elements using protocols like NAS (Non Access Stratum) and GTP (GPRS Tunneling Protocol).

References

[1] Ijaz Ahmad et al. "Security for 5G and Beyond". In: *IEEE Communications Surveys & Tutorials* 21 (2019), pp. 3682–3722. DOI: 10.1109/COMST.2019.2916180.

[2] Ramraj Dangi et al. "Study and investigation on 5G technology: A systematic review". In: *Sensors* 22.1 (2021).

[3] Stavros Eleftherakis et al. "Demystifying Privacy in 5G Stand Alone Networks". In: *arXiv preprint arXiv:2409.17700* (2024).

[4] Iqra Javid and Sibaram Khara. "5G Network: Architecture, Protocols, Challenges and Opportunities". In: *2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*. IEEE. 2022, pp. 1–5.

[5] Xinsheng Ji et al. "Overview of 5G security technology". In: *Science China Information Sciences* 61 (2018). DOI: 10.1007/s11432-017-9426-4.

[6] Roger Piqueras Jover and V. Marojevic. "Security and Protocol Exploit Analysis of the 5G Specifications". In: *IEEE Access* 7 (2018), pp. 24956–24963. DOI: 10.1109/ACCESS.2019.2899254.

[7] Akash R. Kathavate et al. "Critical Aspects of 5G Technology- A Study on the Drivers, Technology Enhancement, Performance, and Spectrum Usage". In: *Asian Journal of Advanced Research and Reports* (2021). DOI: 10.9734/ajarr/2021/v15i530396.

[8] Jaya Preethi Mohan, Niroop Sugunaraj, and Prakash Ranganathan. "Cyber security threats for 5G networks". In: *2022 IEEE international conference on electro information technology (eIT)*. IEEE. 2022, pp. 446–454.