Pak. j. life soc. Sci. (2024), 22(2): 21092-21109 E-ISSN: 2221-7630;P-ISSN: 1727-4915

Pakistan Journal of Life and Social Sciences
www.pjlss.edu.pk

https://doi.org/10.57239/PJLSS-2024-22.2.000204

RESEARCH ARTICLE

Research of the Cellular Communication Protocols Security

L.V. Cherckesova[1*], O.A. Safaryan[2], E.V. Pinevich[3], I.A. Alferova[4]

[1,2,3,4] Don State Technical University, Rostov-on-Don, Russia

ARTICLE INFO

Received: Oct 24, 2024

Accepted: Dec 11, 2024

Keywords

5G
Wireless communication
Protocols
Software engineering

***Corresponding Author:**
chia2002@inbox.ru

ABSTRACT

This paper examines the security of cellular protocols by analyzing the network architecture and protocols across generations. The methodology involves tracing the evolution of cellular communication technologies from 1G to 5G, focusing on architectural features, advancements, and mechanisms while identifying vulnerabilities that persist or emerge in newer generations, such as LTE and 5G. The transition from 2G to 5G has brought increased bandwidth and data transfer demands, necessitating cloud-based flat network architectures. However, LTE and 5G vulnerabilities, such as IMSI traps, VoLTE call interception, and DDoS attacks, expose users to significant risks. The rapid evolution of equipment and software offers immense research opportunities but also expands the attack surface for exploiting wireless technology.

INTRODUCTION

Wireless communications and wireless systems have experienced phenomenal growth over the past decade and have become part of critical infrastructure. The use of wireless devices such as cell phones, computers and laptops (mobile stations – MS) has become a means of implementing viable location-based services and applications that require location information. Federal communications commissions in some countries also require extended 911 emergency services that require similar information.

The Cellular Network Architecture

The cellular network architecture is shown in the diagram (Figure 1). Although this diagram does not relate to specific standard, it provides the overview of various components of the network. Reader can refer to second section of article for more detailed information on the specific protocols architecture. In the radio access subsystem, the mobile station (MS, called user equipment), is the device whose location needs to be determined.

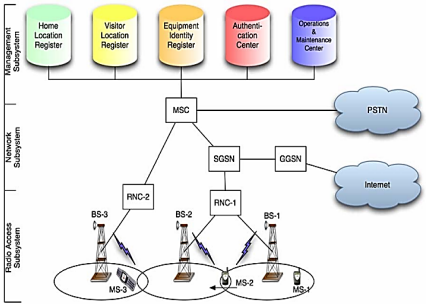


Figure 1: The architecture diagram of the cellular system

Base stations (BS, also called Node B) are the fixed stationary transmitters that are the access points to the rest network. The MS communicate with BS during idle periods (signaling), cell phone calls (voice), or other data transmission. Radio network controllers (RNC), which also manages the radio resources of each BS and MS (frequency channels, time intervals, extended spectrum codes, transmission power, etc.) control the base stations.

THE PROTOCOLS BRIEF DESCRIPTION

The Protocols of the First and Second Generations (1G and 2G)

The wireless cellular network of the 1st generation was introduced in the 1980s. Since then, the various advances were made in this direction, and after 1G new generations were developed, such as 2G, 3G, 4G and 5G networks. Standards of first generation (1G) were analog and had many disadvantages and shortcomings. There were problems with both signal quality and technology compatibility.

The Japanese developers were the first to do this in 1979, then in 1981 the analog network was launched in Denmark, Finland, Norway and Sweden, and in 1983 in the USA.

In 1982, the European Conference of Postal and Telecommunication Agencies was formed the working group called GSM (French Groupe Spécial Mobile, special mobile communications group). The purpose of this group was to study and develop of pan-European terrestrial mobile communication system for general use.

In 1989, the European Telecommunications Standards Institute continued the study and development of the second generation of mobile communications. The abbreviation GSM (Balston, 1993) then acquired the different meaning — Global System for Mobile Communications (Figure 2).

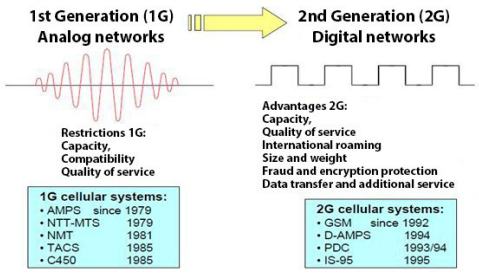


Figure 2: Transition from 1G to 2G

The GSM (Global System for Mobile Communications) protocol stack, which regulates the interaction processes between mobile phones and base stations, can be divided into 3 logical levels:

Layer 1: Physical layer. Protocols of this level describe the principles of interaction between devices in the radio broadcast air. To ensure simultaneous interaction between the network and several mobile devices, the GSM networks use two multiple access technologies: FDMA (Frequency Division Multiple Access) and TDMA (Time Division Multiple Access) (Gabrielyan et al., 2021).

FDMA involves dividing the available frequency range into the channels (ARFCN), each of which allows data to be transmitted from subscribers to the network (uplink) and from the network to subscribers (downlink).

TDMA involves multiplexing the physical channel with time division, that is, each device is given the opportunity to receive and transmit data at certain points in time. According to TDMA, the physical channel (ARFCN) is divided into several logical channels, for example, PCH (Paging channel), on which the base station notifies the phone about an incoming call, and BCCH (Broadcast Control channel), used for identification of the base station by mobile phones (Figure 3).

Layer 2: The channel (data link) layer, the main tasks of which are: establishing, maintaining and disconnecting connections between the network devices; data flows control and monitoring, error detection, as well as data transit of the third level. The protocols LAPD and LAPDm work at this level, multiple connections are provided, as well as the functionality of BCCH, PCH, AGCH and DCCH logical channels.

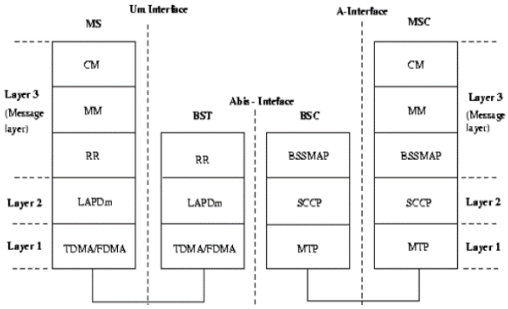


Figure 3: Logical levels of the GSM protocol

Layer 3: Network layer, divided into three sublevels: Radio Resource (RR) — the sublevel responsible for creating and releasing logical channels between devices; Mobility Management (MM) — the sublevel that authenticates users, as well as tracking subscriber movements between coverage areas (cells) of various base stations; Call Control (CC) — the sublevel responsible for phone calls.

The GTP protocol (Technical Report, n.d.) (GPRS Tunneling Protocol) is used to transfer traffic inside PS–core and GRX. This is tunneling protocol that runs on top of the UDP protocol and uses ports 2123 (for management, GTP–C), 2152 (for user data transfer, GTP–U), 3386 (for billing, GTP'). Message Type field is mainly used for the control in GTP–C, in GTP–U usually Message Type = 0xFF (TPDU). TEID is the tunnel identifier that is not mapped to IP address, i.e. the sender of packets with the same TEID may have different addresses at different times (if subscriber moves and switches to other SGSN). When subscriber connects to the Internet, the PDP Context Activation procedure is performed. In the simplified form, this procedure looks like Figure 4.

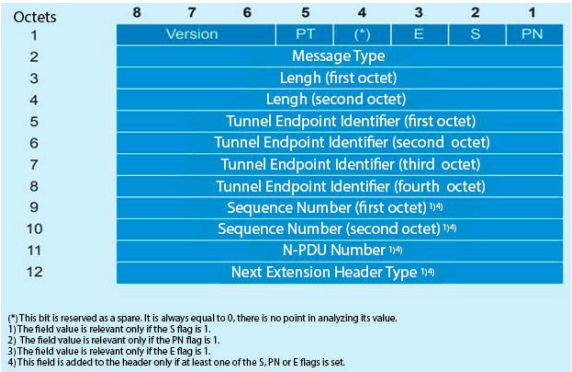


Figure 4: GTP header structure

The Protocols of the Third Generations (3G)

Work on the creation of third-generation technologies began in the 1990s, and their implementation and introduction took place only in the early 2000s (in 2002 in Russia). The standards developed by that time were based on CDMA technology (Code Division Multiple Access – multiple access with code division) (Viterbi, 1995). The third generation of mobile communications includes 5 standards (Grag, 2000a, 2000b): UMTS/WCDMA, CDMA2000/IMT–MC,TD–CDMA/TD–SCDMA, DECT and UWC–136

(Figure 5). The most common of these are UMTS/WCDMA and CDMA2000/IMT–MC standards. UMTS /WCDMA standard has gained popularity in Russia.

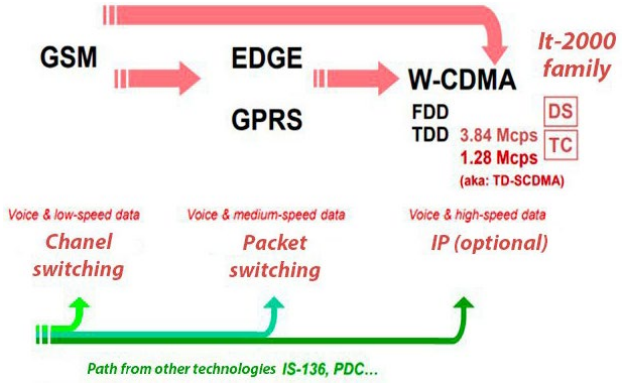


Figure 5: 3G Technology development system from GSM/GPRS/EDGE to UMNS

UMTS (Universal Mobile Telecommunications System – universal mobile telecommunication system) is the cellular communication technology developed for 3G introduction in Europe (Figure 6). The frequency range used is 2110–2200 MHz (often the channel width is 5 MHz). The data transfer rate in UMTS mode is no more than 2 Mbit/s (for the stationary subscriber), and when the subscriber is moving, depending on the speed of movement, it can drop to 144 Kbit/s.

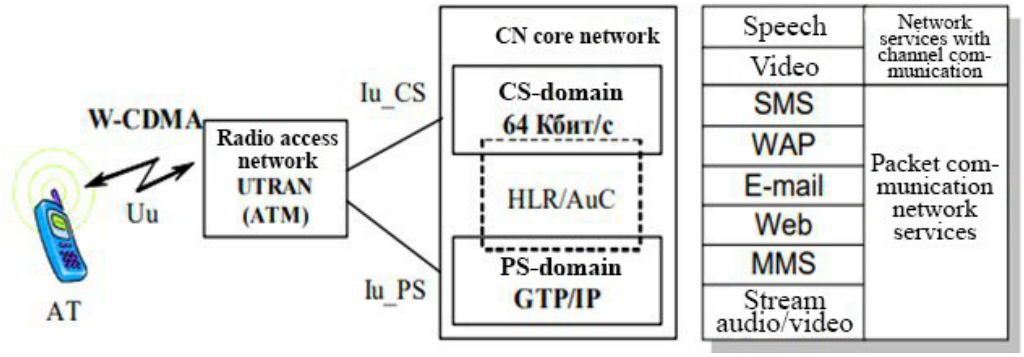


Figure 6: UNTS protocol

HSDPA (High–Speed Downlink Packet Access – high–speed packet data transfer from base station to mobile phone) is the first of the HSPA (High Speed Packet Access) family of cellular communication protocols based on UMTS technology (Figure 7). This protocol and its subsequent versions have significantly increased the data transfer rate in the 3G networks. In its first implementation, the HSDPA protocol had the maximum data transfer rate of 1.2 Mbit/s. The data transfer rate in the next implementation of the HSDPA protocol was already 3.6 Mbit/s.

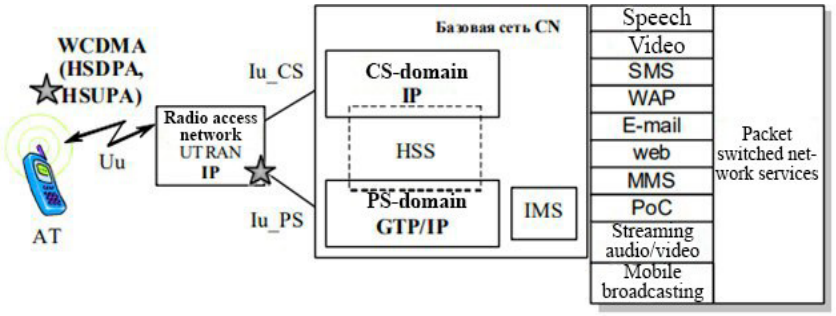


Figure 7: HSDPA protocol

To the further increase of the data transfer rate, HSPA and HSPA (Holma, Toskala, 2006; Stuhlfauth, 2012) + (high-speed packet access) were introduced. Thanks to HSPA+, networks can be upgraded to operate at broadband speeds. The concept of MIMO (Multiple Input Multiple Output) was introduced first in HSPA+. Due to this, the data transfer rate can reach 42 Mbit/s (Viterbi, 1995). HSPA and HSPA+ can be considered as 3.5G and 3.75G (generation), respectively. The modulation method used in HSPA+ was 64–bit QAM.

The Protocols of the Fourth Generations (4G - LTE)

3G, which has not yet exhausted its capabilities, is being replaced by new fourth–generation (4G) technologies, which are more responsive to the demands of that time. 4G generation technologies have outlined completely new requirements for the quality of the communication signal and its stability (Figure 8).

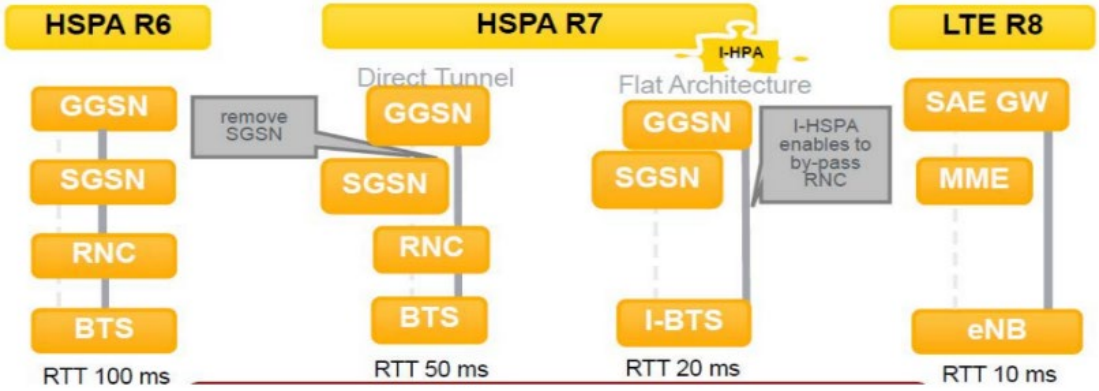


Figure 8: Evolution of the 3G to 4G technology system architecture

The brainchild of joint research by Hewlett-Packard and NTT DoCoMo in the field of developing data transmission technologies in fourth–generation wireless networks were the LTE and WiMax standards (Au, 2006; WiMAX Forum, 2006).

The WiMAX standard was developed in 2001 by the WiMAX Forum organization, which includes manufacturers such as Samsung, Huawei Technologies, Intel and other well–known companies. Conceptually, WiMAX standard is extension of the wireless Wi–Fi standard. The WiMAX standard versions are divided into fixed designed for stationary subscribers, and mobile, for moving subscribers with speed not exceeding 115 km/h. The first commercial WiMAX network was launched in Canada in 2005.

The LTE standard (Long–Term Evolution–long–term development) (Sesia et al., 2011) is, essentially, the continuation of the development of GSM / UMTS standards and originally did not belong to the fourth generation of mobile communications. Today, LTE is the main standard for fourth–generation (4G) networks. First introduced by the aforementioned NTT DoCoMo Company, the world's largest Japanese cellular operator, the LTE standard, in its tenth release of LTE Advanced, was chosen by the International Telecommunication Union as the standard that meets the requirements of fourth–generation wireless communications. The first commercial implementation of the LTE network was carried out in 2009 in Sweden and Norway.

The generalized structure of the LTE network, which shows the presence of two layers of the functional connections: the radio access layer (AS, Access Stratum) and the exterior of the radio access layer (NAS, Non-Access Stratum). The ovals with arrows shown in Figure 9 indicate access points to services.

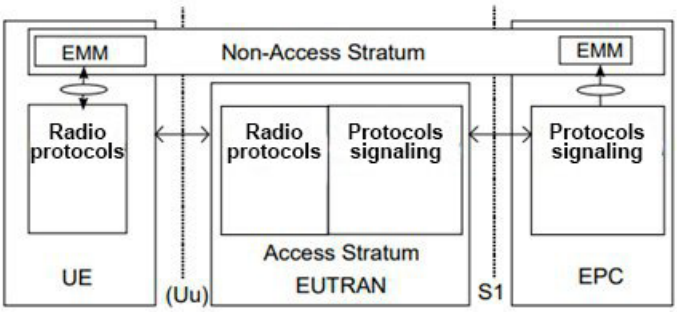


Figure 9: Generalized LTE network structure

The interface between the area of the user equipment (UE) and the area of the radio access network UTRAN (3GPP A Global Initiative, 2013) is called the Uu interface; the interface between the area of the radio access network and the area of the base network (core), EPC is called the S1 interface. The composition and functioning (operation) of the various protocols related to the Uu and S1 interfaces are divided into two so–called planes: the user plane (UP) and the control plane (CP). Mobility management mechanisms in the underlying base network (EMM, EPC Mobility Management) operate outside the access layer.

– The user plane implements protocols that ensure the transmission of user data over the radio channel. The control plane includes those protocols that, in various aspects, provide the connection between the user plane and the network. Protocols designed for the transparent transmission of messages related to the provision of various services also belong to this plane.

–The area of radio access network is divided logically into two levels: the layer of the radio network (RNL, Radio Network Layer) and the layer of the transport network (TNL, Transport Network Layer). The interaction of the basic stations (BS) entering the area of the radio access network is carried out based on the X2 interface. In addition, there is transit connection between the base stations and the base network via the mobility control unit (S1–MM interface) or the service node (S1–U interface), Figure 10.

Thus, it can be argued that S1 interface supports multiple relationships between the set of basic stations and the mobility control units (MCU) and service nodes (SN) blocks.

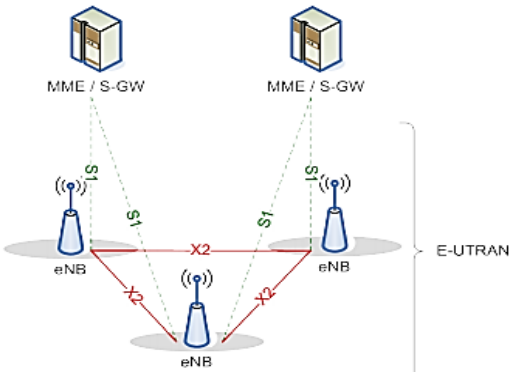
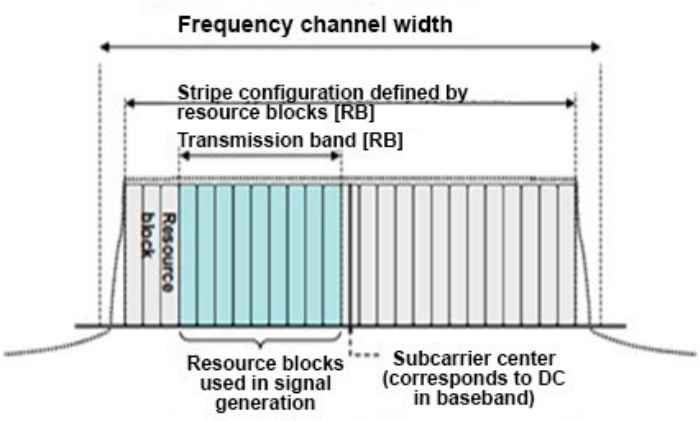


Figure 10: Connection of functional nodes of the radio access network

The maximum theoretical data transfer rate in LTE networks is 326.4 Mbit/s. In practice, the data transfer rate significantly depends on the frequency bandwidth used by the operator. The Megafon mobile operator has the largest frequency range today (40 MHz), which is serious advantage over other domestic mobile operators that use the width of 10 MHz. The maximum data transfer rate in LTE network with bandwidth of 10 MHZ is 75 Mbit/s (Figure 11). The maximum data transfer rate when using 40 MHz bandwidth can reach 300 Mbit/s.



Channel width $BW_{Channel}$ [MHz]	1.4	3	5	10	15	20
Configuration of transmission channels-number of resource blocks N_{RB} system LTE	6	15	25	50	75	100

Figure 11: Data transfer rate

The Protocols of the Fifth Generations (5G)

The main purpose of mobile networks previous generations was simply to offer its users fast and reliable mobile data services. In the case of 5G, the situation is changing, as the range of wireless services is offered to the end user based on variety of access platforms and multi–level networks.

– In essence, the 5G is the dynamic, consistent and flexible set of diverse modern technologies supporting variety of applications. A more intelligent 5G architecture removes the restrictions on the proximity of base stations from radio access networks (RAN) and allows to abandon of the complex architecture. 5G is step towards distributed, flexible and virtual RAN, where new interfaces create additional data access points.

– The 3rd Generation Partnership Project (3GPP) is engaged in telecommunications technologies, including RAN (Heine, 2002), basic transport networks and service capabilities. It has prepared complete system specifications for the 5G network architecture, which is much more service–oriented than previous generations. Thus, services are provided on the base of common mechanism to those network functions that are allowed to use them. The additional qualities of the 5G network architecture, which are described in the 3GPP specifications, are modularity, reuse and autonomy of network functions.

– From 2010 to 2015, research was carried out, the first prototypes were built and various developments in this field were tested. From 2016 to 2018, standardization work was carried out, and in August 2018, the 15th release (Technical Report) was implemented, which defined the fifth–generation communication standard. Three months after its release, the 5G Toolbox appeared in the MATLAB environment. Updated releases 16 – 18 was prepared, and then commercial implementation of this standard began from 2021. Diagram illustrating the development and result of 5G standard relative to the LTE standard is shown in Figure 12. It should be noted that, despite the development of new standards, support for the standards of the previous generation does not stop.

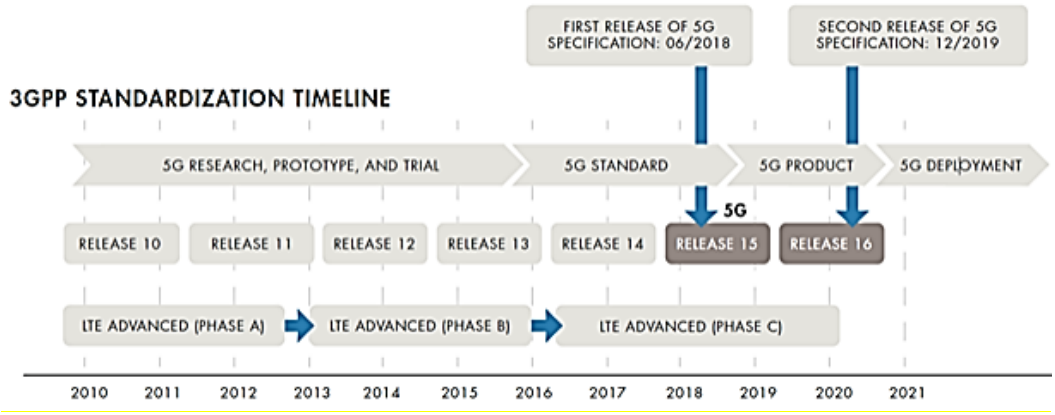


Figure 12: Time diagram of the development of standards of the 4th and 5th generations

Different frequency ranges are allocated for the new generation radio access network (5G NR). The part of the spectrum with frequencies from 30 to 300 GHz is known as millimeter waves, since the wavelengths in it range from 1 to 10 mm. Frequencies from 24 to 100 GHz are now allocated for 5G in various regions of the world. In addition to millimeter waves under 5G, little used UHF frequencies from 300 MHz to 3 GHz are allocated. The variety of frequencies involved can be adapted to unique applications, as higher frequencies also provide higher bandwidth, albeit over shorter distances. Millimeter frequencies are ideal for densely populated areas, but are not suitable for long–range communications. After the low– and high–frequency bands were allocated to 5G, each operator began to develop its own separate part of the 5G spectrum.

The new 5G specification is based on the 5G network core architecture, which provides support for the bandwidth required from 5G. The new 5G core, as defined by 3GPP, uses the cloud–based, service– oriented architecture (SBA) that covers all 5G functions and interactions, including authentication, security, session management, and traffic aggregation from end devices. The 5G core further emphasizes NFV as integral design solution with virtualized software features that can be deployed using the MES infrastructure, indispensable for the principles of the 5G architecture, as it shown in the Figure 13.

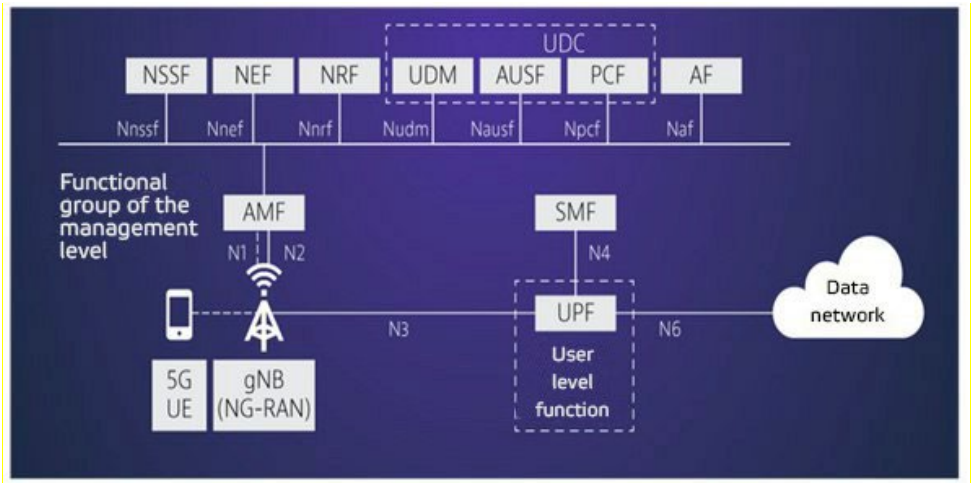


Figure 13: 5G architecture

The purpose of this paper is to examine the security of cellular protocols, which requires the examination of the network architecture and protocols of all generations. The complexity and large

number of protocols, the different types of cellular networks, and the numerous standards that apply in different circumstances make it difficult the documenting accurately such information.

METHODOLOGY

The methodology employed in the study centers on a comprehensive analysis of the security vulnerabilities inherent in cellular communication protocols across generations. It begins with a historical and technical examination, tracing the evolution of protocols from the first generation (1G) to the fifth generation (5G). This includes an in-depth discussion of the architectural features, technological advancements, and mechanisms that define each generation's protocols. Particular emphasis is placed on understanding how improvements in newer generations, such as the introduction of LTE and 5G's dynamic architecture, address or perpetuate existing vulnerabilities. This approach highlights both the progression in communication standards and the persistent gaps that remain in ensuring robust security.

The study further adopts a critical analytical perspective by classifying and detailing specific vulnerabilities within each protocol generation. It explores real-world examples, such as the exploitation of GSM's one-way authentication and the susceptibility of the SS7 and GTP protocols to attacks like spoofing and denial-of-service. By contrasting the improvements in 5G with the inherited weaknesses from earlier standards, the research underscores the necessity for continuous security enhancement. Additionally, it evaluates alternative solutions, like LoRaWAN for IoT, and emphasizes the importance of encryption and mutual authentication in safeguarding data and communication integrity. This methodological framework combines historical review, vulnerability assessment, and mitigation strategies to provide a holistic understanding of the security landscape in cellular protocols.

RESULTS

THE VULNERABILITIES OF CELLULAR COMMUNICATION PROTOCOLS

To date, there are huge number of vulnerabilities in the wireless communication systems. If do not take measures to ensure the security of the wireless network, the potentially dangerous situations are possible, as the result of which an attacker can gain unauthorized access to the personal information of the user.

Let us look at some vulnerabilities.

GSM Protocol Vulnerabilities

Today GSM is the most widespread cellular communication standard in terms of territory and number of subscribers. At the same time, the principles of information security, although they have been developed over many years of using the standard, are subject to threats in practice, since often the newly appeared specifications of the standard were simply not applied by operators (for example, A5/3 encryption).

Vulnerable point of the GSM standard is one–way authentication. The subscriber's device with SIM card is authenticated by the base station, but not vice versa. This vulnerability has led to the emergence of vulnerabilities through so-called "fake" base stations that imitate the real base station.

The attacker passes the subscriber's traffic through "fake" base station with encryption turned off and voice traffic (Porter, Gough, 2007), SMS messages (GSM Association, 2005), and USSD requests become available to him. There are such devices, both in industrial design for law enforcement agencies and available to everyone intruder with the necessary level of knowledge in the field of cellular communications and computer technology.

A5/1 encryption is also not guarantee of data protection – with the help of "rainbow tables" and the "Kraken utility", the encryption key can be found in just minutes in 90 percent of cases.

Telecommunication operators do practically not use the more secure A5/3 encryption – the algorithm works unstable on many models of subscriber communication devices.

SS7alarm system is of particular interest. This is set of signaling telephone protocols used to configure most telephone stations exchanges around the world based on time–division channel networks. Attacker, having gained access to this system, can controls all traffic and location of subscribers around the world.

GSM network security was initially based on the principle of "security through obscurity", but in 1994, main details of algorithm became known: crypto analysts demonstrated the possibility of hacking the A5 algorithm in less than second on the home computer with 128 MB of RAM and 73 Gb of hard disk space.

Lack of mutual authentication between mobile device and network implies that attacker has the ability to install fake base stations and convince mobile devices to connect to it. International Mobile Subscriber Identity (IMSI) has been introduced in 2G systems (individual subscriber number). Nevertheless, in the absence of mutual authentication, fake base stations can be used as "IMSI traps" to collect information and track user behavior.

Vulnerabilities of the Protocol UMTS (3G)

UMTS does not provide sufficient protection against the user's identity interception. Although IMSI is replaced with TMSI after the initial connection request, IMSI is sent in its pure form during the first rrcConnectionRequest, as well as in cases such as database VLR failure and inability VLR to identify TMSI. In the multi–service delivery system where mobile phone will be used for e–commerce, banking transactions and wireless mobile payments, the confidentiality of the user's identity and location is the necessity. Therefore, the above reasons are enough to attract the attacker's attention to identify the user and even to track him.

In particular, the first connection request message is sent over the channel of random access, which is common channel and can be easily intercepted. Many studies have been conducted on GSM to extract the International Mobile Subscriber Identity (IMSI) during connection (Uttmark, 2017). GSM–based IMSI catchers are currently available on the market, but their prices reach $500,000 and are sold only to government agencies.

Some researchers are interested in creating the capture device IMSI that has all the capabilities of commercial product. UMTS was developed after GSM and adopted the function of sending IMSI in the clear text at the beginning of connection request. Therefore, the above devices can also be used to receive the IMSI of UMTS subscriber. In addition, some UMTS debugging and testing tools, such as Air Protocol Analyzer AP–6000 and Protocol Tester K1297–G20, have the ability to receive IMSI. There are also reports that police can track suspects successfully, using mobile phone simply by using IMSI.

Consequently, the user's identity / location privacy and user traceability are violated clearly. A simple scenario for receiving the IMSI of UMTS subscriber is shown in the Figure 14 below.

Attacker impersonates UMTS VLR / SGSN. During the rrcConnectionRequest, the victim can use TMSI. If TMSI cannot be resolved, the network can make the identification request. In this case, R should send its IMSI in clear text. After receiving the IMSI, the attacker disconnects (Figure 14).

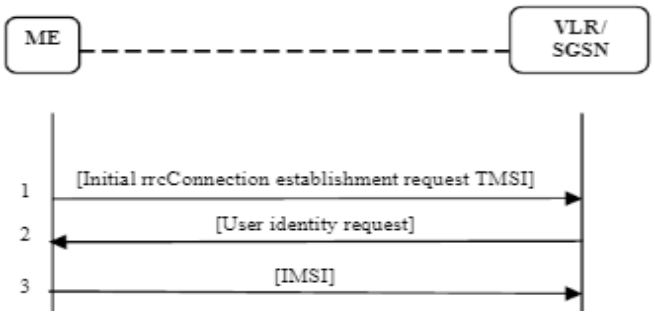


Figure 14: The scheme of IMSI obtaining

DOS HLR/AuC Attack

This is very dangerous attack by specific mobile operator. With this attack, the services of particular mobile operator can be blocked (Lee, 2009). The attack can be launched in two stages:

Stage I: At this stage, the attacker creates IMSI database corresponding to the victim's operator. Attacker can easily obtain IMSI using the procedure described in the section above. The intruder can identify the operator from the IMSI, since the digits 4 and 5 (or 4, 5 and 6) represent the operator in the IMSI structure.

Stage II: At this stage, the attacker quickly generates rrc ConnectionRequests corresponding to each IMSI using automatic procedure. For each request (except for the one already connected), VLR / SGSN sends the IMSI to HLR/AuC, which checks the validity of the IMSI. Since the attacker obtained all valid IMSI in stage I, therefore, all IMSI pass this confidence test. Now HLR calculates 5 AV corresponding to each IMSI.

This is cumbersome process of calculating RAND, MAC, XRES, CK, IK, AK for each IMSI. The AV corresponding to each IMSI are then sent to the VLR / SGSN. For each IMSI, the VLR/SGSN selects one AV and sends the RAND and AUTN for authentication. This is the stage at which the attacker will not be authenticated, since he cannot calculate RES without knowing the USIM secret key. But this is not the attacker's aim, in fact he has already done his job. His purpose – to exhaust the computing resources of HLR/AuC by avalanching more and more valid rrcConnectionRequests and using it to calculate AV.

It can also cause the exhaustion of bandwidth between VLR / SGSN and HLR/AuC. This exhaustion of resources will lead to DOS attacks for new users who try to connect.

The Vulnerabilities of the Protocol LTE (4G)

International Mobile Subscriber Identity (IMSI) has been introduced in 2G systems – the international mobile subscriber identifier (individual subscriber number). However, in the absence of mutual authentication, fake base stations can be used as “IMSI traps” to collect information and track user behavior. The 3G specification introduces mutual authentication and the use of more robust and well-analyzed cryptographic algorithms. LTE specifications further strengthen signaling protocols by requiring authentication and encryption. Since IMSI is permanent LTE subscriber identifier, the specification needs to minimize its transmission over radio channels for security and privacy reasons. Instead, the Globally Unique Temporary Identifier (GUTI) is used, which allows to identify subscribers in the radio network. The vulnerabilities of the SS7 system and the DIAMETER protocol, as well as vulnerabilities related to DPI, are considered for investigation.

The Vulnerability of the SS7 System

The standard Signaling System 7 is used for the exchange of service information between network devices in telecommunication networks (Dryberg, 2006). The SS7 system was developed forty years

ago, and has certain disadvantages in terms of security (for example, there is no encryption and authentication of service messages).

The network supports the needs of mobile communications and the provision of additional services.

In the early 2000s, the specification SIGTRAN was proposed, which allowed service information SS7to be transmitted over IP–networks. The signal network has ceased to be isolated. The SS7 gateway is required to access the signaling network. The intruder can get operator license, get into the network through hacked operator equipment, GGSN or femtocell. Due to the vulnerabilities, it is possible to obtain the subscriber's balance, intercept incoming message, listen to the call, steal information about the subscriber, determine his location and carry out the fraud: redirecting incoming call, transferring funds via USSD, changing the subscriber's profile, redirecting outgoing call, carrying out DOS attacks.

Most attacks on SS7 networks were possible due to the lack of verification of the subscriber's real location. The second and third places in the list of reasons are the inability to verify the subscriber's network affiliation and the lack of filtering of unused alarm messages. In the fourth position, there are SMS Home Routing configuration errors. Most of the disadvantages that allow to determine the location of the subscriber and to steal the data can be eliminated by changing the configuration of network equipment. At minimum, it is necessary to prohibit the processing of messages of type *AnyTimeInterrogation* and *SendIMSI* on HLR.

The architectural problems of protocols and systems are solved by blocking unwanted messages. First, it is necessary should pay attention to the messages of type SendRoutingInfoForSM, SendRoutingInfo, SendRoutungInfoForLCS, SendIMSI. Filtering will help to avoid the risks associated with denial of service, interception of SMS messages, redirection of calls, listening to calls, changing the subscriber's profile.

However, not all of SS7 messages of the network may be dangerous. It is necessary to implement filtering in such way that only unwanted messages used in attacks are cut off. To do this, it is recommended to implement additional security measures, for example, intrusion detection systems. Such systems do not affect network traffic, but they allow identifying the actions of the intruder, and determining the necessary message filtering settings.

The Vulnerability of the Protocol DIAMETER

The DIAMETER is the protocol of the AAA family, and is development of the RADIUS system. It is intended, mainly, for evaluating services in communication networks. In particular, in 3G networks, it is used to evaluate the data transmission services, and in IMS3/LTE, the protocol is one of the main control elements. The important feature of the protocol is its extensibility and the ability to create not only your own attributes, but also applications. The basic protocol implements the requirements for AAA protocols, the details of which are reflected in RFC2989, and describes the connection establishment process, compatibility verification, rules for sending messages and routing them, and disconnection. TCP and SCTP can be used as transport.

Protocol security should be provided at the transport layer and the recommendations state also that the protocol should not be used without TLS, DTLS, or IPsec. In trusted network it is possible to do without them, in particular, if the internal network of the enterprise can be considered reliable.

The specification defines several types of DIAMETER nodes. To understand the role of nodes, it is necessary to introduce two terms, which will be discussed in more detail below.

Session – controls the subscriber's status and includes only those messages that relate to single subscriber.

Connection – monitors communication status between DIAMETER nodes. Client is network device that directly handles subscriber's traffic. Role of server is clear; it should control the status of subscriber's sessions.

Agent. DIAMETER agents are intermediate nodes between client and server and perform traffic management functions. They can aggregate messages from devices on the same site, act as load balancer, modify DIAMETER packets, and act as security gateways when switching from trusted network to public one.

Functionally, agents are divided into several types:

DIAMETER Relay agent (DRL): Nodes of this type receive DIAMETER–messages from network devices and redirect them to the following nodes based on data contained in the messages based on Realm and the list of known neighbors. DRL can be used to aggregate messages from multiple nodes located in the same geographical area. DRL does not modify the significant fields in message body, so they can work with any type of DIAMETER applications. When establishing connection, they must announce the Relay Application Id.

Proxy agent. Proxies are similar to Relay, but they can modify the payload of DIAMETER–messages, for example, to control access to certain services, modify field values, etc. Most often, Relay and Proxy are combined into one entity, because Proxy without conversion rules is Relay agent shown in Figure 15.

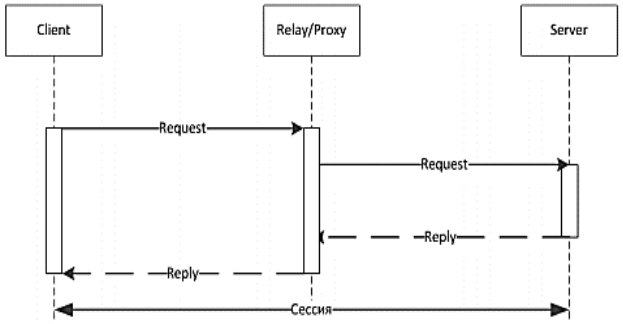


Figure 15: Main nodes of the DIAMETER protocol

Redirect Agent (DRD)

DRDs are used if DIAMETER routing rules for messages must be controlled from single point (Diametriq, LLC, n.d.). DRD receives request, determines node to which it should be sent and tells where it should be redirected (Figure 16).

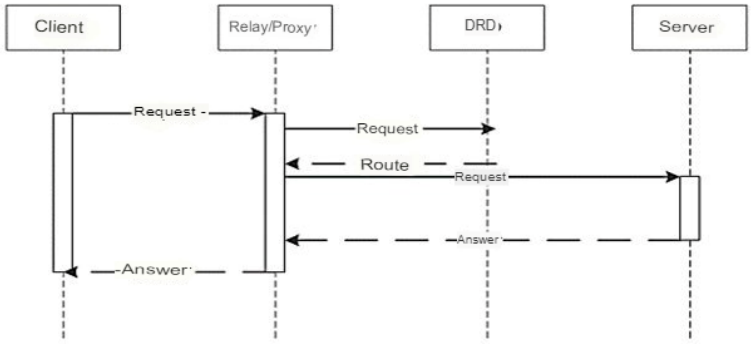


Figure 16: Control of DIAMETER messages from the single point

Node Class Translation Agents (TLA) – Protocol translators (Figure 17)

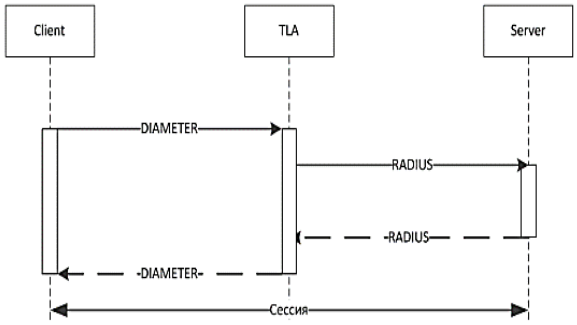


Figure 17: DIAMETER protocol translator

They are used if the physical devices are incompatible at the level of AAA protocols. For example, to implement the conversion from RADIUS to DIAMETER and back, or to support conversion between two incompatible versions of DIAMETER. TLA must ensure the transactionality and storage of the session state during its processing. TLA should announce only those ApplicationID whose support is implemented, because only in this case the agent will be able to process correctly the incoming message.

The 4G network based on the DIAMETER signaling protocol has vulnerabilities that allow attacks related to subscriber location detection, SMS interception and denial of service. One of its functions is to provide roaming among subscribers. To intercept SMS, the attacker depicts usual scheme of operators work when moving the subscriber to roaming. It sends notification to subscriber's home network that the subscriber–victim is in range of another (fake) cellular network. In response, home network forwards all incoming messages and calls to attacker.

Deep Packet Inspection Systems (DPI) of Mobile Operators

Deep Packet Inspection (DPI) systems are software and hardware complexes for classifying passing Internet traffic by data type (web page, document, audio, video), protocol (HTTP, BitTorrent, VoIP / SIP) and specific programs (Skype, WhatsApp), often with additional functionality (Roth, Schillinger, 2014).

DPI systems are widespread and used by providers all over the world. Operators used DPI since about beginning of 2000th, with the advent of UMTS (3G), in order to share limited bandwidth wireless channel.

Mobile operators use other DPI features, such as TCP and HTTP traffic acceleration (TCP PEP, Performance–Enhancing Proxy) also, to speed up the Internet in mobile networks and identify users by websites. If to try to log into the operator's personal account from the phone, it will open immediately on many operators, without having to enter the username and password. Alternatively, what could have been found some years ago, simple visit to suspicious website or click on advertising banner from Android game turned into automatic subscription to paid service, which could be found out from SMS message. Deep traffic analysis system is configured so that it adds service HTTP headers when executing HTTP request to sites (hosts) from the list determined by the operator. Headers may contain subscriber's internal IP address, phone number (MSISDN), IMEI and IMSI identifiers, and the identifier of base station to which subscriber is connected (ECI /TAC).

The Vulnerabilities of the Protocol 5G

Outdated GTP protocol is the reason for vulnerability of 5G networks (Ivanov, 2020).

The ancient vulnerabilities plaguing already outdated broadband standards will migrate to relatively new 5G technology. The specialists of Positive Technologies stated this. According to the researchers' report, the main culprit for the vulnerability of fifth generation networks is the GTP protocol (GPRS Tunneling Protocol). Due to GTP, 5G networks will be open to such forms of attacks as spoofing, Man-in-the-Middle (MitM) and DOS. The developers introduced the GTP protocol in 2G generation, and it is also used in the current version – 4G. This standard allows data transfer between GSN nodes in the packet network. For example, if the user is roaming, GTP will make calls through local telecom operator. The main task of the protocol is to create easy way of interaction between telecom operators, which allows data to be transmitted through different networks and countries. Unfortunately, GTP itself contains the number of significant security problems – for example, the protocol cannot correctly verify the geolocation and subscriber's credentials. In fact, this means that potential attacker can fake traffic and hide his number by impersonating another person.

This type of attack is used often to force unsuspecting user to subscribe to paid services.

The Protocol PFCP

According to information security researchers from Positive Technologies, some 5G networks are at risk of attacks due to "long-standing vulnerabilities" in the core protocols. The HTTP/2 protocol, used to perform critical network functions, including registering and storing user profiles, contains vulnerabilities that can allow attackers to carry out denial-of-service (DOS, DDOS) attacks against mobile phone users.

Draft Standard for LoRaWAN as Method of Cellular Communication Protocols Security

LoRaWAN (Long Range Wide–area Networks, long–range global network) is the most well–known LoRa hardware protocol, which is designed to control the communication between LPWAN gateways and end devices (Economov, 2019). LoRaWAN IoT networks have multi–stage system for protecting transmitted information. In addition, data security solutions in LoRaWAN networks can be customized according to the needs of customers. At first glance, there are the ready–made and already tested solutions for LTE, 3G protection (Figure 18).

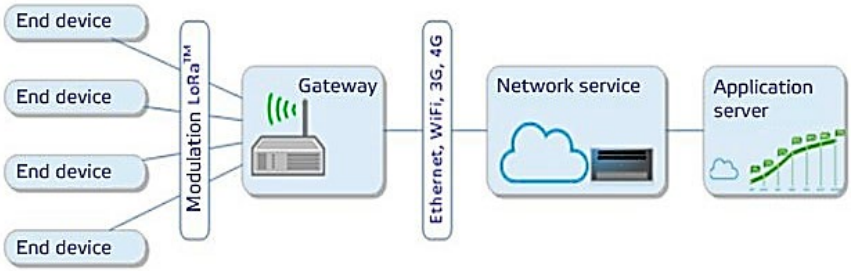


Figure 18: Architecture of LoRa networks

In LoRaWAN network, the node does not communicate with specific gateway, but transmits data to several gateways. Each gateway forwards the received packet from destination node via transport (cellular network, Wi–Fi, Ethernet or other) to the cloud server. Server manages the network, discards redundant packets, performs security checks, plans optimal route for transmitting confirmation message, and controls the data transfer rate. Using such architecture allows to get rid of *handover procedure* when moving mobile sensors within the network. Nodes in the network operate asynchronously and transmit data as it accumulates or by interruption.

STANDARD MEASURES OF SECURITY

The LoRaWAN IoT network is used multi–level security system of data transmission (Figure 19).

Level 1. AES–encryption at *application level* (between subscriber terminal and application server) using 128–bit variable session key AppSKey. This key is stored in subscriber's terminal and on application server and is not available to network operator. Only the client, the owner of application server, has access to AppSKey.

Level 2. AES–encryption and integrity verification of messages at the *network level* (between the subscriber terminal and the network server) using 128–bit NwkSKey variable session key.

This key is stored in the subscriber terminal and on the network server and is not available to the client. Only the network operator, the owner of the network server, has access to NwkSKey.

Level 3. Standard authentication and encryption Internet protocol methods (IPsec, TLS, etc.) when transferring data over transport network between network nodes (base station, servers – network, Join, application).

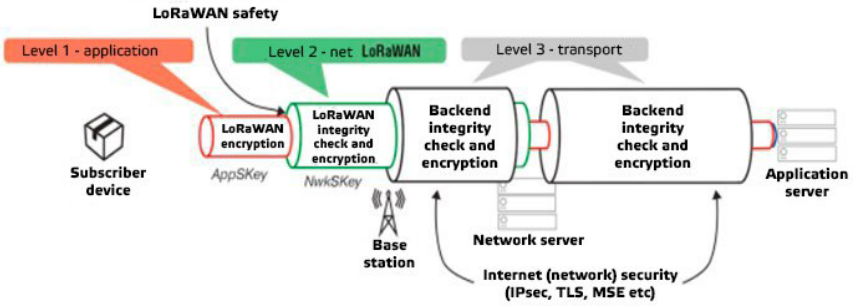


Figure 19: Data transmission security system

CONCLUSION

Evolution of networks from 2G to 5G is associated with increased requirements for bandwidth and data transfer rates from 9.6 kbit/s to 100 Mbit/s, which required the flat network architecture creation using cloud technologies. Exploiting LTE/5G network vulnerabilities are:

- Attackers can monitor subscribers, collect statistics;
- Listen to VoLTE calls, intercept Internet traffic and SMS messages (for example, with one-time bank passwords), use IMSI traps, thanks to which attackers can monitor users' devices;
- Disclosure of information about user's network, determine users location, realize DDOS attacks.

In addition, there is large number and actively developing both equipment and special software with high potential for practical application and research of wireless communication protocols.

It provide wide opportunities for research and practical implementation of wireless network. This, in turn, opens up wide opportunities for attackers to use illegally the capabilities of wireless technology stacks.

LIST OF ABBREVIATIONS

OF LTE – Long Term Evolution ABBREVIATIONS,

W–CDMA – Wideband Code Division Multiple Access),

UMTS – Universal Mobile Telecommunications System,

GSM – Global System for Mobile Communications,

QoS – Quality of Service,

EPS – Evolved Packet System,

E–UTRAN – Evolved Universal Terrestrial Radio Access Network,

E–UTRA – Evolved Universal Terrestrial Radio Access,

FD–LTE – Frequency Division,

TD–LTE – Time Division,

OFDM – Orthogonal Frequency-Division Multiplexing,

QPSK – Quadrature phase shift keying,

GPRS – General Packet Radio Service,

EDGE – Enhanced DataRates for GSM Evolution,

SAE – System Architecture Evolution,

DFT – Discrete Fourier Transform,

BEP – Bit Error Probability,

FEC – Forward Error Correction,

SNR – Signal Noise Ratio,

QAM – Quadrature Amplitude Modulation,

MIMO – Multiple Input Multiple Output,

UE – User Equipment,

SC–FDMA – Single Carrier Frequency Division Multiple Access,

OFDMA – Orthogonal Frequency Division Multiple Access,

BER – Bit Error Rate,

LPWAN – Low Power Wide Area Network.

REFERENCES

3GPP A Global Initiative. Mobile Radio lnterface Layer3 Specification; Core network protocols; Stage 3. 2013. https://www.arib.or.jp/english/html/overview/doc/STD-T63v10_10/5_Appendix/Rel10/24/24008-aa0.pdf

Au M. Mobile WiMAX certification labs coming on line. April 26, 2006. https://www.telecomasia.net/content/mobile-wimax-certification-labs-coming-line/

Balston DM. The Pan–European system: GSM. In: Balston DM, Macario RCV (Eds.), Cellular Radio Systems. Boston: Artech. House; 1993.

Diametriq, LLC. DIAMETER Routing Agent. n.d. https://diametriq.com/diametriq-products/diameter-routing-engine-diametriqs-diameter-routing-agent

Dryberg L. Alarm System. No. 7 (SS7/OX7). Protocols, Structure, Application. London: Williams; 2006. 752 p.

Economov A. Set' LoRaWAN: Bezopasnost' obespechivayetsya [LoRaWAN Network: Security is provided]. March 26, 2019. https://www.iksmedia.ru/articles/5573226-Set-LoRaWAN-bezopasnost-obespechiva.html?ysclid=m4crdsou5j472687015

Gabrielyan D, Kostoglotov A, Safaryan O, Cherckesova L, Dvornikov O. Method for estimating time length using simultaneous phase measurements in the system of simultaneously and independently operating generators. Advanced Engineering Research 2021,21(1):105–110.

Grag VK. IS-95 CDMA and CDMA 2000: Cellular/PCS Systems Implementation. Prentice Hall; 2000a.

Grag VK. Soft handoff and power control in IS-95 CDMA. In: IS-95 CDMA and CDMA 2000: Cellular/PCS Systems Implementation. Prentice Hall; 2000b.

GSM Association. SMS SS7 Fraud 3.1. February 16, 2005. https://www.gsma.com/newsroom/wp-content/uploads/2012/12/IR7031.pdf

Heine G. GPRS Signaling and Protocol Analysis. Vol. 1: RAN and Mobile Station. Bern: Artech House Publishers; 2002. 242 p.

Holma H, Toskala A (Eds.). HSUPA for UMTS: High Speed Radio Access for Mobile Communications. Chichester; Hoboken: John Wiley; 2006.

Ivanov O. Staryy protokol GTP — Prichina uyazvimosti setey 5G [Outdated GTP protocol is the reason for vulnerability of 5G Networks]. June 11, 2020. https://www.anti-malware.ru/news/2020-06-11-1447/32918

Lee PPC, Bu T, Woo TYC. On the detection of signaling DOS attacks on 3G/WiMAX wireless networks. Comput Netw 2009,53:2601-16. http://dx.doi.org/10.1016/j.comnet.2009.05.008

Porter T, Gough M. How to Cheat at VoIP Security. Rockland: Syngress; 2007.

Roth C, Schillinger R. Detectability of deep packet inspection in common provider. In: Consumer Relations: 2014 25th International Workshop on Database and Expert Systems Applications (DEXA), Munich, Germany. IEEE; 2014. p. 283-287. https://doi.org/10.1109/DEXA.2014.64

Sesia S, Toufik I, Baker M. LTE the UMTS Long Term Evolution: From Theory to Practice. Chippenham: John Wiley & Sons Ltd; 2011.

Stuhlfauth R. High Speed Packet Access: Technology and Measurement Aspects of HSDPA and HSUPA Mobile Radio Systems. Munich: Rohde & Schwarz; 2012.

Technical Report. TR-25.930vl.2.1. 3 Generation Partnership Project (3GPP), Technical Support Group (TSG - SA) Working Group2 (Architecture). RAN – meeting#5. n.d. http://www.3gpp.org/ftp/tsg_sa/%20G2_Arch/TSGS_05/tempdoc/S2-99410.doc/

Uttmark M. LTE IMSI Catchers. May 30, 2017. https://hackaday.com/2017/05/30/

Viterbi AJ. CDMA: Principles of Spread Spectrum Communication. Reading: Addison-Wesley Pub. Co; 1995.

WiMAX Forum. Mobile WiMAX. Part 1: Technical Overview and Performance Evaluation. August 2006. https://wimaxforum.org/news/downloads/Mobile_WiMAX_Part1_Overview_and_Performance.pdf