NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

SUPPLEMENTAL TO THESIS
(2 of 2: ATTACK VECTOR TAXONOMY)

SPOOFED NETWORKS: EXPLOITATION OF GNSS SECURITY VULNERABILITY IN 4G AND 5G MOBILE NETWORKS

by

Matthew J. Lanoue

June 2021

Thesis Advisor: Chad A. Bollmann
Co-Advisors: James B. Michael
Darren J. Rogers

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

An Attack Vector Taxonomy for Mobile Telephony Security Vulnerabilities

Matthew Lanoue, Chad A. Bollmann, James Bret Michael, and John Roth, Naval Postgraduate School

Duminda Wijesekera, George Mason University

A simplified cybersecurity threat matrix may provide a unifying way to define the security risk posed by current and future generations of mobile telephony.

A security framework proposed by Adam Shostack simplifies threat modeling by asking four questions:[1]

1. What are we working on?
2. What can go wrong?
3. What are we going to do about it?
4. Did we do a good job?

Digital Object Identifier 10.1109/MC.2021.3057059
Date of current version: 9 April 2021

While the first and last questions may be easier to answer, the second and third questions require substantial effort to address. To answer the second question, Shostack proposes using the STRIDE method to identify potential threats. This method invites scientists and engineers to imagine how common attack methods such as spoofing, tampering, repudiation, information disclosure, denial of service (DoS), and escalation of privilege may be used to target a system. For simple systems, this may be effective, but it is not an efficient method for brainstorming threats for a complex system such as mobile telephony.

Addressing the security risks for mobile telephony is a multidisciplinary endeavor. The prevailing practice of examining methodologies and how attacks might be applied does not result in an intuitive representation of attack vectors usable by most experts in areas other than security. We believe that an improved approach to identifying potential threats to mobile telephony, consistent with the four-question framework, is to categorize attacks in terms of attack vectors and their relationship to the user plane (UP),

 0018-9162/21©2021IEEE

Authorized licensed use limited to: NPS Dudley Knox Library. Downloaded on April 18,2021 at 19:08:29 UTC from IEEE Xplore. Restrictions apply.

EDITOR JAMES BRET MICHAEL
IEEE Senior Member; bmichael@nps.edu

control plane (CP), and their interactions throughout the radio front-end and the core network. Our proposed matrix can aggregate more detailed taxonomies in the literature to enable unified views of what can go wrong and facilitate decisions on what to do about it.

5G MOBILE TELEPHONY

The fifth generation of mobile telephony, also known as new radio (5G NR), promises increased bandwidth, reduced latency, customizability, and greater cellular coverage. The apps envisioned for 5G can be broadly categorized into three use cases: enhanced mobile broadband, massive machine-type communication, and ultrareliable low-latency communication. Examples of planned apps include vehicle-to-everything (V2X) communication, smart cities with networked sensors, and mobile streaming of 3D video with ultrahigh resolution and ultra-reliable low-latency communication. 5G virtualized many services of the previous generations and created app-specific, virtual network slices that specify performance and cybersecurity requirements—a significant advance. But 5G systems have also become complex in the quest to enable the flexibility to meet anticipated functional and performance requirements over the next 10 years. 6G may introduce additional complexity with its move from centralized service provisioning to a decentralized network and service architecture.

With the rapid adoption of 5G NR, one might assume that stakeholders have already thoroughly scrubbed 5G NR artifacts (for example, standards, protocols, architecture, and designs) for security vulnerabilities, although recent studies indicate otherwise.[2,3] Others have enumerated security vulnerabilities in 5G,[4–6] but the current taxonomies and threat models suffer from drawbacks, including complex methodologies that can be difficult for nonsecurity experts to rapidly apply. Some of these other taxonomies also afford limited flexibility to adapt to changes in technology, architecture, design, and usage.

Each engineering and science discipline involved in mobile telephony contributes in some way to understanding and assessing the security risks within the overall engineering tradeoff space. A common, easily graspable taxonomy to categorize the threats, which parts of the network they affect, and where they originate is necessary to facilitate the effective engagement and investment of resources spread across the research, development, sustainment, and defense of the mobile telephony infrastructure.

Accordingly, we developed a threat-matrix-based approach to the categorization of attacks based on attack vectors instead of methods. The approach is straightforward to use and adaptable. Let's take a look at how the approach can be applied to 5G NR.

5G NR ARCHITECTURE

Figure 1 illustrates the key components of a simplified 5G architecture. The 5G

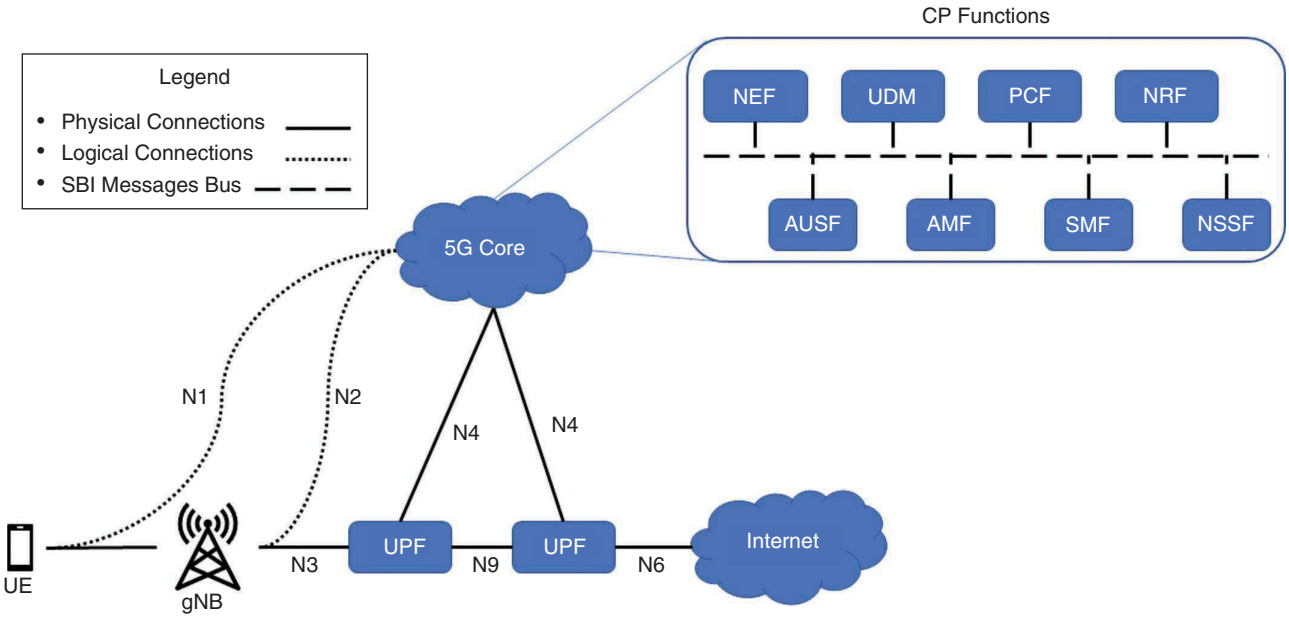

FIGURE 1. An overview of the 5G network architecture. SBI: service-based infrastructure; UE: user equipment; UPF: UP function; UDM: unified data management; AMF: access and mobility management function; SMF: session management function; NEF: network exposure function; AUSF: authentication server function; PCF: policy control function; NRF: network repository function; NSSF: network slice selection function.

Authorized licensed use limited to: NPS Dudley Knox Library. Downloaded on April 18,2021 at 19:08:29 UTC from IEEE Xplore. Restrictions apply.

architecture organizes telephony capabilities as being in the CP or UP in addition to specifying the relationships between the two planes. The specific implementation of any 5G network may differ between service providers and granularity can be added (such as an access plane).

The CP in Figure 1 consists of the 5G core and CP functions (CPFs), the logical connections N1 and N2 to the UE and logical radio nodes known as gNBs, and the N4 connections to the UPF. CPFs in the 5G core communicate with one another via HTTP/2 queries and responses over the SBI bus. For the interested reader, more details are contained in "Additional Details on 5G Architecture."

The UP of Figure 1 consists of the UE, radio access network (RAN), UPF, and the connections N3, N6, and N9. Within the RAN, the actual mechanism by which the base station (gNB) interacts with the core network depends on the mode of operation and equipment involved. It could consist of a simple gNB such as in Figure 1 or multiple radio units (RUs) working together as a single, virtual gNB to provide integrated access and backhaul (IAB), as depicted in Figure 2.

In a 5G deployment, the UP may consist of a single gNB or multiple gNBs as well as a single UPF or multiple UPFs. IAB provisioning itself is a potential UP attack (UPA) target.

ADDITIONAL DETAILS ON 5G ARCHITECTURE

RADIO ACCESS NETWORK (RAN)

Per our approach, the access plane (consisting primarily of the RAN) is rolled into the UP. The Open RAN Alliance (O-RAN ALLIANCE) splits the RAN into three parts by separating the functional modules that compose a single logical radio node (gNB): the radio unit (RU), distributed unit (DU), and centralized unit (CU). In 5G literature, fronthaul refers to the link between the RU and DU, midhaul to the link between the DU and CU, and backhaul to the link between the CU and the 5G core network. This split, illustrated in Figure 2, supports integrated access and backhaul, promotes interoperability among vendors' implementations of these modules within the RAN, and can (or will) be used to provide complementary services such as optimized traffic flow based on artificial intelligence techniques.[S1] The additional computing power available in the RAN enables the creation of new services further away from the core network—these are called mobile edge services.

UP

The N1 and N2 connections are used to pass information to and from the CPFs by the UE and gNB, respectively. As depicted in Figure 2, the UPF can be further classified by its connection to other UP features. The packet data session anchor UPF connects to a gNB via the N3 link, and the intermediate UPF (IUPF) connects to another UPF via the N9 link. The IUPF connected to an external data network such as the Internet via the N6 link may also support inter public land mobile network UP security to protect the network from incoming malicious traffic.[S2]

CP

Important CPFs include the AMF, SMF, and UDM. The AMF controls the process for new UE and gNBs to connect to the 5G network and UE handoffs between gNBs. Within the CP, the destination for information carried over the N1 and N2 connections is the AMF. Upon the UE's request, the SMF creates, updates, and terminates sessions as permitted by the AMF and manages the session context with the UPF over the N4 connection. The UDM replaces the home subscriber server in the 4G standard; it manages user data and authentication credentials. Further information on CPFs is found in 3GPP Technical Specification 23.501; network functions and entities are listed in clause 4.2.2, and further details about each network function are in clause 6.2.[S2]

NETWORK FUNCTION VIRTUALIZATION AND NETWORK SLICING

Instead of each network function residing on separate machines, the network functions share common hardware and become virtual network functions (VNFs). While this concept has been applied to some 4G networks, it will be fully adopted in 5G, including within the RAN.[4] This is evident in the decoupling of gNB functionality, as previously discussed. A network slice can thus be thought of as all of the related VNFs servicing a certain network app. For instance, a server rack may contain a network slice for a massive machine-type communication network for the factory floor and another slice for an ultrareliable low-latency communication network for connecting supervisors to multiple geographically dispersed factory floors.

References

S1. "Open RAN explained." Nokia, Espoo, Finland, Oct. 16, 2020. [Online]. Available: www.nokia.com/about-us/newsroom/articles/open-ran-explained

S2. "System architecture for the 5G system (5GS)." 3GPP, Sophia Antipolis, Tech. Spec. 23.501. V16.6.0," June 24, 2020.

Authorized licensed use limited to: NPS Dudley Knox Library. Downloaded on April 18,2021 at 19:08:29 UTC from IEEE Xplore. Restrictions apply.

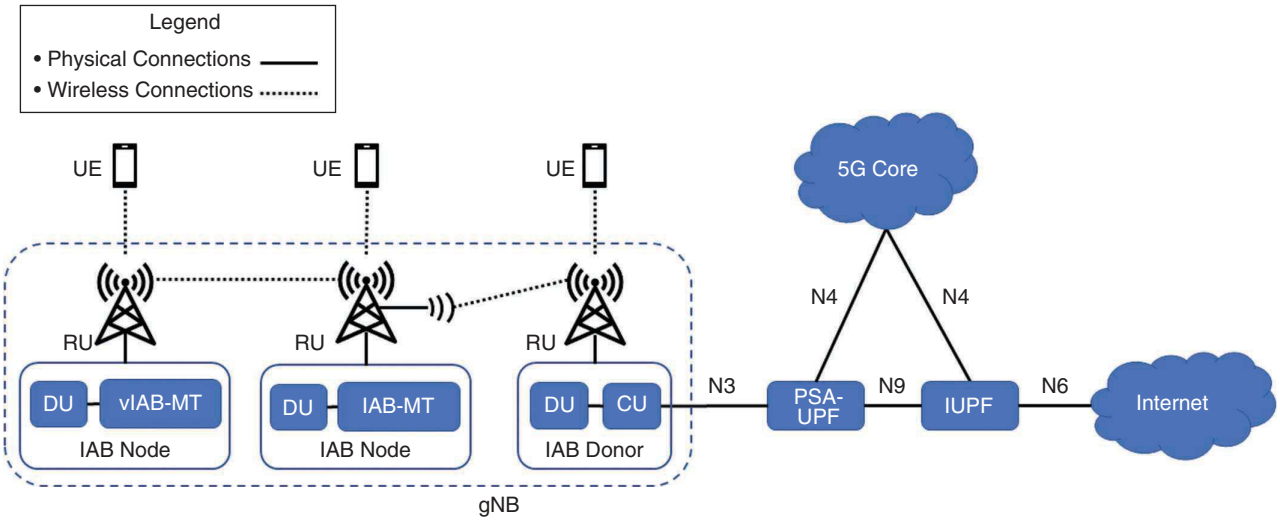

FIGURE 2. The disaggregation of RAN functions for IAB. DU: distributed unit; CU: centralized unit; PSA-UPF: packet data session anchor UPF; IUPF: intermediate UPF; IAB-MT: integrated access and backhaul mobile termination; vIAB-MT: virtual IAB-MT.

In addition to the changes in network architecture, 5G introduces or improves various security mechanisms. These measures include encrypting the International Mobile Subscriber Identity (IMSI) whenever it is required to be transmitted over the air, limiting which network functions have access to the IMSI (as part of the Subscription Permanent Identifier), and the improved 5G Authentication and Key Agreement protocol.[7,8] A deeper investigation of the security architecture for 5G is beyond the scope of this article. Instead, we focus on how our methodology can be applied to mobile telephony, using a brief overview of the 5G network architecture as a backdrop.

MOBILE TELEPHONY ATTACK VECTOR CATEGORIZATION SCHEME

Much of the literature on the security of mobile telephony centers on attack methods instead of the origins or targets of attacks. While approaches from MITRE and the European Union Agency for Cybersecurity (ENISA) may help to consolidate attack vectors based upon objectives or methods such as outages[5] or credential access,[6] these approaches do not adequately address complex systems such as 5G in which functionality is partitioned among multiple planes. For example, a DoS attack (that is, an outage) could be utilized against a gNB to prevent access within a network cell or against a service provider's UDM to prevent access to an entire network.

Although these attacks have the same objective (outage) and method (target gNB), the attacks differ in their implementation and effect. Similarly, credential access could potentially be accomplished both in the UP as well as the CP. Thus, while binning attack vectors on the method is convenient, it does not provide sufficient clarity to the network engineers, who are constantly refining the standards, and the vendors and service providers, who are constantly refining the UE and infrastructure. Binning on methodology also requires a more complex taxonomy that is difficult to visualize.

Considering that the CP and UP split is fundamental to 5G, we instead propose the approach described in Table 1 that bins attack vectors based on their source and targeted function [that is, a UPA or CP attack (CPA)] within these two planes. This simplified top-level model provides for aggregating many of the more detailed taxonomies in the literature, such as ENISA[5] and ATT&CK Mobile,[6] and can assist stakeholders in forming their own mental models of security risks.

To account for the complexity of the multiplane system, the user of the matrix categorizes the attack vectors based on the plane of their origin (UP or CP) and the type of attack (UPA or CPA), as shown in Table 1. In a CPA, the targeted function or feature lies in the CP, whereas in a UPA, the function or feature that is targeted lies in the UP.

TABLE 1. The proposed 5G attack vector matrix.

		Type of Attack (Targeted Function)	
		UPA	CPA
Attack source	UP		
	CP		

Authorized licensed use limited to: NPS Dudley Knox Library. Downloaded on April 18,2021 at 19:08:29 UTC from IEEE Xplore. Restrictions apply.

The following sections analyze selected attack methods (recently proposed as well as legacy) and examine how they might be applied to mobile telephony, specifically against a 5G network. These attack methods are then fleshed out into more specific attack vectors and, in turn, categorized within our proposed taxonomy. While some of these legacy attack vectors presented may not be viable against a 5G network, we use them to illustrate our taxonomy. Used in conjunction with Shostack's framework, one can assess the policies and mechanisms for mitigating the security risks of 5G in addition to assessing the goodness of the underlying trust assumptions for the security architecture. The power of categorizing attacks based on attack vectors is in presenting a simplified visualization of weaknesses in the mobile telephony architecture to facilitate threat discovery. We believe that some of the first attack vectors employed against 5G networks will be derivatives of legacy attack vectors.

CPAs

The proprietary nature by which individual organizations implement CP features may deter unsophisticated or poorly funded would-be attackers. History, however, teaches us that security through obscurity is inadequate. In addition, the network functions and large repositories of information are likely to be tempting targets for attackers. The 5G CP will also be the most complex, and history also shows that each added feature typically introduces bugs (that is, vulnerabilities or weaknesses) and dependencies; some of the weaknesses may be exploitable for cyberattacks. Finally, the diversity of services and dynamic reconfigurability envisioned in 5G will require extensive virtualization in control nodes and many more edge control nodes. The net result of these changes significantly increases both the attack surface and access vectors to the 5G CP.

Network function spoofing

The use of software-defined networking and network function virtualization to create multiple app-specific network slices will rely on configurations and updates that traverse the Internet using well-documented and recognized network protocols (for example, HTTP/2, TCP, and IP). A capable adversary could insert itself into the routing chain and determine the location and functions contained in a particular network slice. Instead of probing and exploiting vulnerabilities within the network slice application programming interface and functions, they could redirect 5G traffic from a particular RAN to a spoofed network core, thus enabling the collection of information for all connected UE and the manipulation of any network function within the slice.

> But 5G systems have also become complex in the quest to enable the flexibility to meet anticipated functional and performance requirements over the next 10 years.

Database record access and manipulation

Using an initial methodology similar to network function spoofing, an adversary could exploit a vulnerability within the network slice instead of simulating the entire slice. Such attacks are potentially more damaging because the effects could be more difficult to detect and longer lasting. A spoofed slice or database is fixed by reestablishing the proper connections, but one that is manipulated requires additional effort to determine the scope of corruption of the database and restore its integrity.

Signaling interception, manipulation, and jamming

The radio frequencies associated with 5G are tremendously diverse. Frequency bands are deployed differently by country, but low- and midbands with longer propagation paths start at about 600 MHz (for low bands) and 2.5 GHz for the midrange. The highest bands, millimeter-wave (mmWave), span the 20–60 GHz range. The highest frequencies typically have much shorter and less robust propagation paths due to higher attenuation. However, the increased directionality can support ultradense deployments in urban areas and permit high-data-rate and low-latency apps that were simply not feasible with the available 4G spectrum. 5G design calls for significantly greater base station density and pushing many core services of the network to the edge to handle short propagation paths and novel apps.

The CP is subject to jamming attacks from the middle and both ends. End devices, legitimate or rogue, can target the CP for DoS and certain other types of cyberattacks. At the end of the day, the edge node is just another piece of commodity silicon running an operating system hosting many virtualized apps.[9] Similar attacks, targeting data in transit of user information such as location and other privacy data, can originate from the Internet (used for global backhaul). Regarding "the middle," dense deployments in urban areas require radio frequency (RF) backhaul that will enable classes of attacks not available in 4G.

While the encryption and directionality of backhaul links will mitigate many types of interception and manipulation attacks, jamming is possible through both RF and physical means. The high attenuation and directionality of mmWave signals permit low-tech jamming solutions such as antenna or path blockage. In safety- or time-critical apps, such as the Industrial Internet of Things (IIoT) or

Authorized licensed use limited to: NPS Dudley Knox Library. Downloaded on April 18,2021 at 19:08:29 UTC from IEEE Xplore. Restrictions apply.

connected vehicle communications, DoS jamming attacks on the CP present a real threat to public safety and have been proven feasible in the 4G and, by extension, the 5G CPs.[9]

In the context of applying the proposed classification taxonomy, jamming a link in the chain of nodes used to facilitate IAB (see Figure 2) could be used to manipulate artificial intelligence (AI) logic that routes user data to the CU connected via N3 to the PSA-UPF. If a compromised gNB was inserted in the new path determined by AI and used to harvest user information, then the targeted jamming would be considered a CPA launched from the UP.

UPAs

While the transition from 4G to 5G may alter the technical details of UPAs, UPAs tend to be the most heavily researched and publicized subset of 5G attack vectors. Examples of UPAs that originate in the UP include DoS via cell jamming, key stealing, and hardware vulnerability exploitation. Just as a DoS attack can be accomplished via multiple methods, a spoofed base station (gNB) can be used to accomplish a variety of attacks. Using a compromised gNB to deny service to a group of UE within the cell would represent a UPA launched from the UP. However, using the compromised gNB to create a database of user credentials and security keys by simulating access requests to the core network represents a CPA launched from the UP.

UE spoofing

Part of the 5G NR appeal lies in its billing as the "one standard to rule them all" by enabling the convergence of the previously isolated silos composed of V2X, industrial control systems (ICSs), the IoT/IIoT, and mobile broadband (MBB) communications (among others). Although service providers try to offer robust authentication methods to thwart the use of spoofed (or rogue) UE, the added complexity of 5G NR creates opportunities for developing new classes of UE spoofing attacks. For instance, one might leverage the integration of ultralow-latency (ULL) communications provisioned for safety and industrial apps by exploiting known vulnerabilities in legacy manufacturing stations. As noted previously, industrial or safety use cases present new consequences for DoS-type attacks against devices in the UP.

ICSs typically have stringent timing, availability, and reliability constraints that require the ULL promised by 5G. These use cases require the addition of quality-of-service (QoS) provisions to the 5G NR standards or slices thereof. These new specifications can, in turn, be leveraged to induce new classes of DoS attacks.[10] For instance, the 5G NR standards permit ULL UE to preempt transmissions from MBB UE. In one recent study, authors simulated both throughput degradation via MBB preemption and breaking typical ULL QoS guarantees using as few as five pieces of rogue UE.[10] Ultimately, these attacks warp 5G's flexibility to preempt normal device-to-base-station communications with spoofed higher priority requests; the resulting DoS does not have to be highly effective to break QoS guarantees for safety applications.

Cell jamming

The novelty of 5G ensures that many of the proposed 5G jamming attacks are still theoretical. But, in principle at least, individual frequencies or mobile devices (UE) are subject to many of the same attacks that have already been proven against 4G and are well described by Lichtman et al.[9]

The frequency diversity of 5G NR makes it much more difficult for an adversary to jam all possible downlinks from an entire cell. In turn, this means broad-based noise jamming will remain effective but require high transmission powers and be readily observable to counterjamming sensors. Selective jamming strategies will be stealthier and likely more effective at targeting smaller sets of UE. Selective jamming strategies target specific cell-to-UE connection information carried by cell downlinks such as the primary synchronization signal (PSS) and physical broadcast channel (PBCH). These attacks would typically leverage the cell's own broadcast information to identify and select the specific time and frequency blocks to target. Note that there are many more potential targets than the PSS and PBCH alone; Lichtman et al. provide a thorough discussion of these attacks.[9]

> Used in conjunction with Shostack's framework, one can assess the policies and mechanisms for mitigating the security risks of 5G in addition to assessing the goodness of the underlying trust assumptions for the security architecture.

Base station spoofing

Base station spoofing (both Wi-Fi and 3G/4G) is a proven attack used by both law enforcement and criminals to target individuals and devices via the UP. Rogue base stations can be configured to enable different types of attacks. If configured to forward connection information from UE to the CP and connect to a legitimate cellular provider, the base station acts as if it was the target's device. By reading or recording the intercepted traffic, the base station can impose nearly any attack on the connected UE, including confidentiality and integrity compromise, geolocation, and DoS.[11,12] Less-complex attacks have also been demonstrated that do not require the breaking of encryption, such as collecting UE identifying data, service downgrade, and battery-draining attacks.[5]

Authorized licensed use limited to: NPS Dudley Knox Library. Downloaded on April 18,2021 at 19:08:29 UTC from IEEE Xplore. Restrictions apply.

Hardware vulnerability exploitation

The popularity of location-based apps has resulted in many providers of telecommunications chips bundling global navigation satellite system (GNSS) functionality into their mobile telephone system as a system-on-a-chip solution. GNSS, being a set of well-documented standards designed to work with a low signal-to-noise ratio, enables a myriad of terrestrial spoofing attacks. For instance, CVE-2019–2254 represents a vulnerability applicable to numerous Qualcomm chip sets that uses the spoofed commands sent to the GNSS chip to conduct a buffer overflow and execute arbitrary code on a system.[13,14] The prevalence of nonmobile telephone features on modern smartphones represents a significant attack vector for hardware vulnerabilities by increasing the device's attack surface, as discussed previously.

> Just as a DoS attack can be accomplished via multiple methods, a spoofed base station can be used to accomplish a variety of attacks.

Key catching and stealing

Encryption between UE and RAN utilizes a permanently stored key in the UE's subscriber identification module (SIM) card. The IMSI and International Mobile Equipment Identity of the UE are protected by temporary IDs once the device is connected to the RAN. However, an attacker can force a situation where the temporary identifiers (IDs) stored in the UE and network fall out of synchronization. In this case, the network may request the UE to use its permanent ID to reestablish communications. In 4G networks, this mechanism could be exploited by the attacker to force all UEs served by a particular gNB to broadcast their IMSI, which was done in the clear, referred to as IMSI catching. Since the attack uses a CP authorization mechanism to steal user credentials, it would be a UPA launched from the CP.

Additionally, the use of permanent symmetric keys has fallen under scrutiny after The Great SIM Heist revealed a large-scale attack against the SIM provisioning process.[4,15,16] Other methods for storing permanent authorization credentials include embedded universal integrated circuit cards or some other form of "soft SIM" that can be remotely provisioned. Adversarial manipulation of any remote provisioning process could be used in a DoS attack or to enable the connection of unauthorized devices to a network. Such key stealing operations are UPA launched from the UP because the credentials stored in a SIM card are part of the UE.

Timing attacks

Precise time synchronization is essential for mobile apps, such as the synchronization of base stations to enable call handoff.[17] While clock synchronization can be accomplished via cellular signals, it can also be accomplished via network-based protocols such as the Network Time Protocol (NTP) and Precision Time Protocol (PTP) or by wireless signals, such as those from GNSS constellations supporting real-time kinematic positioning. Each method has an associated accuracy and interface—which introduces its own associated set of vulnerabilities. Clock synchronization protocols such as GNSS are one way and are thus susceptible to man-in-the-middle (MITM) attacks. PTP is a two-way clock synchronization protocol, where the round-trip time delay can be measured to detect MITM attacks. However, this alone is insufficient to guarantee protocol security.[17]

Timing attacks could be utilized to increase network delays, degrade communication links, or

TABLE 2. A partially populated 5G attack vector matrix.

		Type of Attack	
		UPA	CPA
Attack source	UP	• Cell jamming used to conduct DoS attack • Spoofed base station conducts DoS to a subset of attached UE • Spoofed UE used to attach an unauthorized device to the network • Hardware vulnerability exploitation • Key stealing	• Spoofed gNB obtains user credentials by simulating multiple access requests • Jam select links in IAB path to reroute user data through compromised gNB • Modified gNB conducts DoS against multiple gNBs attached to 5G channel by falsely advertising favorable channel characteristics
	CP	• IMSI catching • Timing attacks • Use privileged access to core network to authorize attachment of modified gNB that captures user credentials	• Use privileged access to core network to manipulate billing • Use privileged access to core network to facilitate DoS attack by modifying routing policies, QoS, and so on

Authorized licensed use limited to: NPS Dudley Knox Library. Downloaded on April 18,2021 at 19:08:29 UTC from IEEE Xplore. Restrictions apply.

create opportunities for UE to connect to a spoofed network infrastructure. Some levels of security against this type of attack are offered by the variety of timing synchronization options available to the gNB. An attack that manipulates the timing inputs to the gNB to increase delays or degrade service is a UPA launched from the CP since timing is a service provided through the CP.

MITIGATIONS AND THE WAY AHEAD FOR IMPROVED SECURITY IN 5G AND BEYOND

Agreement on a common scheme for categorizing security vulnerabilities is a required first step to enabling us to be efficient and effective at improving the security of mobile telephony. For example, the scheme can help us identify areas of overlap where investing our scarce time and people resources will maximize our return on investment in addressing potential threats. Vulnerabilities can then be identified and investigated by using industrial-strength formal methods, simulations, and experimentation, which, in turn, feed into threat modeling and security risk reduction. The goal here is to systematically improve 5G security and inform the research and development of follow-on work regarding architectures, standards, protocols, and implementations. Let's not repeat security missteps that occurred in prior generations of mobile telephony.

DISCLAIMER

Any mention of commercial products or references to commercial organizations are for information only. It does not imply recommendation or endorsement by the U.S. Government, nor does it imply that the products mentioned are necessarily the best available for this purpose. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the U.S. Government.

Table 2 illustrates our vector matrix, which is partially populated using the attack examples described previously. This scheme is invariant to the evolution of equipment, standards, and implementations, making it possible to apply it to 6G and beyond—which may introduce new functional planes. It also helps consolidate multiple attack methods, making common threat vectors within the 5G architecture easier to discern and assess.

Each threat warrants a discussion of mitigations that are beyond the scope of this article. But we can identify a couple of overarching mitigation considerations by considering two specific attacks discussed previously: cell jamming and base station spoofing.

Regarding selective jamming attacks, because most of this "free" targeting information is codified in the 5G NR standards, mitigating potential attacks in scalable, compatible ways is a challenging and open area of research. Hardware changes will have to be avoided to ensure backward compatibility. Software changes for security may unacceptably reduce the data rate or battery life. Because of the difficulty of changing a mature technology, the most effective mitigation—restricting the amount of timing and frequency detail that is provided to UE prior to authentication with a base station and encryption of the link—will have to be reserved for 6G and beyond.

Some of these same mitigations would help to address the security risks related to base station spoofing. Additional service provider mitigations could include automated cellular network sensing and anomaly detection based on unusual gNB control messages (for example, frequent RRC_REJECT and RRC_IDLE), unusually high signal strength reports, or inconsistent blacklisted cell data.[11] UE solutions could involve additional checks on certain control messages from the base station as well as reports to the CP.[11] Both UE and edge reports would effectively increase the number of sensors available to a provider to detect malicious and accidental network issues.

Many of the rogue base station mitigations could also be implemented in software in the CP. While imposing computational and data overhead, providers can implement these fixes relatively quickly. But to ensure smooth roaming and protection across cellular network operators, some fixes would require changes to the standards. Standards changes are the source of urgency for fully identifying 5G threats and mitigations; 6G standards are currently being refined, and the window for accepting changes may close before all of the useful changes are identified.[18]

> Timing attacks could be utilized to increase network delays, degrade communication links, or create opportunities for UE to connect to a spoofed network infrastructure.

We've presented a novel scheme for categorizing attack vectors based on a dual-plane architecture, such as that specified for 5G. Given that mobile telephony is part of the Internet, applying our scheme requires thinking about how other parts of the Internet interact with mobile telephony. The Internet can be thought of as an extension of the RAN, carrying CP and UE messages. With this perspective, how do we classify the cyberattacks emanating from outside of the mobile telephony environment in the proposed scheme? For example, in Table 2 we listed timing attacks as UPA from the CP. To us, this makes sense if NTP or PTP is the feature we are looking

Authorized licensed use limited to: NPS Dudley Knox Library. Downloaded on April 18,2021 at 19:08:29 UTC from IEEE Xplore. Restrictions apply.

at. If, instead, we looked at a border router peered with a 5G network, one could argue that the router is part of the UP since it is part of the data-routing path.

Regardless of whether an attacker attempts to leverage weaknesses of the UP, CP, or both, improving society's trust in the dependability of deployed 5G NR networks and the apps they enable will be challenging. By embracing the concept of the functional-plane split, our taxonomy is usable for 6G and beyond while remaining simple enough to facilitate engagement and understanding by experts across technical domains. By focusing on vectors rather than specific methods, our scheme can aggregate more detailed methodologies and accommodate the evolution of mobile telephony technology and cyberattacks. The matrix can also be scaled as desired, while improved threat visualization enables allocating limited resources to obtaining the most fruitful security enhancements. C

REFERENCES

1. J. Beyer, "Adam Shostack on threat modeling," *IEEE Softw.*, vol. 37, no. 6, pp. 110–112, 2020. doi: 10.1109/MS.2020.3017406.
2. F. Rashid. "5G networks will inherit their predecessors' security issues," *IEEE Spectrum*. June 23, 2020. https://spectrum.ieee.org/tech-talk/telecom/security/5g-networks-will-juggle-legacy-security-issues-for-years
3. A. Shaik, R. Borgaonkar, S. Park, and J.-P. Seifert, "New vulnerabilities in 4G and 5G cellular access network protocols: Exposing device capabilities," in *Proc. 12th Conf. Security Privacy Wireless Mobile Netw.*, 2019, pp. 221–231. doi: 10.1145/3317549.3319728.
4. P. Schneider and G. Horn, "Towards 5G security," in *Proc. IEEE Trustcom/BigDataSE/ISPA*, 2015, pp. 1165–1170. doi: 10.1109/Trustcom.2015.499.
5. "ENISA threat landscape for 5G networks," European Union Agency for Cybersecurity, Attiki, Greece, Dec. 2020. doi: 10.2824/802229.
6. "Mobile matrices," MITRE Corporation. https://attack.mitre.org/versions/v8/matrices/mobile/ (accessed Jan. 4, 2021).
7. "Security architecture and procedures for 5G system," 3GPP, Sophia Antipolis, Tech. Spec. 35.501. V16.5.0, Dec. 16, 2020.
8. A. Koutsos, "The 5G-AKA authentication protocol privacy," in *Proc. 2019 IEEE Euro. Symp. Security Privacy*, pp. 464–479. doi: 10.1109/EuroSP.2019.00041.
9. M. Lichtman, R. Rao, V. Marojevic, J. Reed, and R. Jover, "5G NR jamming, spoofing, and sniffing: Threat assessment and mitigation," in *Proc. IEEE Int. Conf. Commun. Workshops*, 2018, pp. 1–6. doi: 10.1109/ICCW.2018.8403769.
10. C. Chen, G. Hung, and H. Hsieh, "A study on a new type of DDoS attack against 5G ultra-reliable and low-latency communications," in *Proc. Euro. Conf. Netw. Commun.*, 2020, pp. 188–193. doi: 10.1109/EuCNC48522.2020.9200956.
11. S. Jha et al., Private communication, Oct. 2020.
12. S. Hussain, M. Echeverria, O. Chowdhury, N. Li, and E. Bertino, "Privacy attacks to the 4G and 5G cellular paging protocols using side channel information," in *Proc. Netw. Distrib. Syst. Security Symp.*, 2019, pp. 1–15. doi: 10.14722/ndss.2019.23442.
13. "CVE-2019-2254 detail," NIST, Gaithersburg, MD, NVD-CVE-2019-2254, 2019. [Online]. Available: https://nvd.nist.gov/vuln/detail/CVE-2019-2254
14. "Multiple vulnerabilities in Google Android OS could allow for arbitrary code execution," Center for Internet Security, East Greenbush, NY, 2019. [Online]. Available: https://www.cisecurity.org/advisory/multiple-vulnerabilities-in-google-android-os-could-allow-for-arbitrary-code-execution_2019-069/
15. J. Scahill and J. Begley. "How spies stole the keys to the encryption castle," The Intercept, Feb. 19, 2015. https://theintercept.com/2015/02/19/great-sim-heist/
16. "Rationale and track of security decisions in Long Term Evolved (LTE) RAN/3GPP System Architecture Evolution (SAE)," 3GPP, Sophia Antipolis, Tech. Spec. 33.821. V9.0.0," June 12, 2009.
17. L. Narula and T. E. Humphreys, "Requirements for secure clock synchronization," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 4, pp. 749–762, 2018. doi: 10.1109/JSTSP.2018.2835772.
18. X. Giao, Y. Huang, S. Dustdar, and J. Chen, "6G vision: An AI-driven decentralized network and service architecture," *IEEE Internet Comput.*, vol. 24, no. 4, pp. 34–40, 2020. doi: 10.1109/MIC.2020.2987738.

MATTHEW LANOUE is a lieutenant in the U.S. Navy and a graduate student in the Naval Postgraduate School's Department of Electrical & Computer Engineering, Monterey, California, 93943, USA. Contact him at matthew.lanoue@nps.edu.

CHAD A. BOLLMANN is a permanent military professor in the Naval Postgraduate School's Department of Electrical & Computer Engineering, Monterey, California, 93943, USA. Contact him at cabollma@nps.edu.

JAMES BRET MICHAEL is a professor in the Naval Postgraduate School's Department of Computer Science and Department of Electrical & Computer Engineering, Monterey, California, 93943, USA. Contact him at bmichael@nps.edu.

JOHN ROTH is an assistant professor in the Naval Postgraduate School's Department of Electrical & Computer Engineering, Monterey, California, 93943, USA. Contact him at jdroth@nps.edu.

DUMINDA WIJESEKERA is a professor in the Department of Computer Science at George Mason University, Fairfax, Virginia, 22030, USA. Contact him at dwijesek@gmu.edu.

Authorized licensed use limited to: NPS Dudley Knox Library. Downloaded on April 18,2021 at 19:08:29 UTC from IEEE Xplore. Restrictions apply.