Federal Communications Commission

Technological Advisory Council

5G Cybersecurity Subcommittee

September 12, 2016

Editor: Tom McGarry
tom.mcgarry@neustar.biz

Table of Contents

1. Executive Summary

This document contains draft security-related recommendations intended for 5G standards development activities. TAC members will work these draft recommendations with ATIS PTSC to develop a final list of recommendations for 3GPP. The draft recommendations are:

- Denial of service
  - It is recommended that 5G standards be defined in such a way as to enable resource isolation techniques such as network slicing to confine the effects of DoS attacks
  - It is recommended that 5G networks be able to deauthorize an individual device (or multiple devices) in such a way as the device does not continue to utilize the control plane or media plane resources
  - It is recommended that base stations have the ability to schedule the radio resource for each device in an unpredictable way
  - It is recommended that 5G network elements embed DoS detection and mitigation functions into the RAN functions via key security indicators with related dynamic resolution
- Key management
  - It is recommended that industry standard encryption techniques be used to protect data during transport
  - It is recommended that 5G networks provide options for using asymmetric key material to support diverse IoT Use Cases
  - It is recommended that 5G networks enable privacy protections to guard against using key and certificates to identify and track consumers
  - It is recommended that 5G standards development consider alternative trust models that enable flexibility in establishing trust models across heterogeneous devices, access technologies, network domains and communication modes
  - It is recommended that 5G networks support new secure enrollment processes that allow entities other than carriers to provision enrollment certificates to devices
  - It is recommended that 5G networks support robust methods for identifying and responding to misbehavior
  - It is recommended that 5G networks support multiple devices that operate at multiple levels of sensitivity/assurance
- Identity management
  - It is recommended that the 5G network that provides access to a device be able to uniquely identify, authenticate and authorize each individual device that accesses the network either directly or indirectly (e.g., via a gateway, virtual network)
  - It is recommended that an equipment or subscriber identity that is transported across networks and presented to a terminating device be authenticated and authorized
  - It is recommended that UE be able to authenticate the network before attaching
  - It is recommended that Soft SIMs deploy rigorous cybersecurity measures that can protect against attacks aimed at software applications
- Isolation mechanisms
  - It is recommended that 5G standards be defined in such a way as to enable resource isolation techniques such as network slicing to enable different levels of security among different resources
  - It is recommended that there be access to the control plane and media plane at the base station to enable security monitoring of traffic

2. FCC 2016 activity on 5G

FCC Direction to TAC

In 2016 the FCC provided the following directions to the TAC Cybersecurity Working Group:

> The next evolutionary step in wireless broadband communication, 5G, is expected to support a highly diverse range of new applications, user requirements, and connected devices, including smartphones, Sensors, Robotics, mission critical wireless communication, and automated guided vehicle systems for the automotive and automotive supply industries. All of this will be realized by the continued development of a number of existing wireless technologies (e.g., LTE) and new Radio Access Technologies (RAT).
>
> However, one of the key challenges facing 5G is to support such a wide spectrum of distinctly different use cases and user requirements in an agile, reliable, and secure manner, all at the same time. The security aspect of 5G is expected to play a critical role in a number of use cases and services under consideration, including: 1) new use cases having a direct impact on safety of life such as self-driving vehicles 2) supporting various mission critical services that require stringent security requirements; 3) supporting a number of legacy technologies that are either unsecure or not sufficiently secure; 4) supporting billions of wirelessly connected devices with wide-ranging capabilities, including low cost devices that lack the necessary tools to secure data or be upgraded in the field; 5) making use of newly conceived programmable core network architectures that raise serious security questions, not yet answered; 6) supporting vulnerable mobile communication devices (e.g., smartphones, tablet) which are expected (under 5G) to become even more powerful devices, thereby making them more alluring to hackers and more menacing not only to the device's owner but also to our global interconnected networks and economy.
>
> The first draft of 5G standards is expected to be released by the end of year 2018, meaning we are still at the pre-standards stage. We ask the Working Group (WG) to utilize what the Cyber WG has learned about IoT and programmable networks security, and any other related topics, in order to recommend to the FCC the strategy, procedures and steps necessary to help incorporate the concept of "security by design" into the very fabric of 5G, its design specifications, and consequently 5G's complex multi-product line ecosystem. Additionally, we ask the following questions:
>
> 1. What other key technical areas, if any, should be researched while exploring ways to integrate "security by design" concept into 5G design specifications?
> 2. What are the important tools and security controls that should be built into 5G design specifications in order to make 5G networks and devices sufficiently secure from the onset?
> 3. What are the SDOs most active in 5G standards development process? To what extent do TAC members participate in those SDOs? What opportunities exist for those members, either through direct voting or other advocacy mechanisms, to support the TAC's recommendations and ensure future standards incorporate security from the outset?
> 4. How do we make sure the security controls identified become integrated into 5G design specifications? Describe strategy, procedures involved and specific step to take in this regard.
> 5. How should the FCC and industry coordinate efforts in order to maximize their effectiveness in this endeavor?

Based on the FCC's direction the Cybersecurity Working Group created a 5G Subcommittee, which created this document and the recommendations herein.

FCC Report and order on 5G

On July 14, 2016 the US became the first country to open up spectrum for 5G service when the FCC voted unanimously to release spectrum above 24GHz for 5G service (FCC 16-89)[1]. They also require a Notice of Inquiry to be issued by 10-31-16 on the topic of 5G security.

5G licensees are required to provide an overview of their 5G security strategy[2] including plans for safeguarding their networks and devices from security breaches. The FCC goes on to say that requiring licensees to submit that information at that juncture creates an incentive for them to engage in the development of security measures at an earlier stage.

The FCC defines specific links to be secured[3]:

- Device to the licensee's network
- One element of the licensee's network to another element on the licensee's network
- Licensee's network to another network
- Device to device

The FCC also defines use cases for security[4]:

- Communications between a wireless device and the licensee's network
- Communications within and between each licensee's network
- Communications between mobile devices that are under end-to-end control of the licensee
- Communications between mobile devices that are not under the end-to-end control of the licensee

3. Security in the 5G environment

5G will enable a fully mobile and connected society. Drivers of 5G include IoT, mobile broadband and mission critical systems. This will require new service delivery models that involve new actors in the ecosystem. Cloud and virtualization technologies will be deployed to provide flexibility and the ability to deliver richer services quickly. Networks will provide greater API access to users and third party service providers. This environment will generate many changes in the existing mobile networks and create new security challenges. We've identified four categories of security that warrant review and recommendations. These are not listed in any priority order.

- Denial of service (DoS)
- Key management
- Identity management
- Isolation mechanisms

While evaluating security requirements for 5G we developed three categories for how devices will attach to the network.

---

[1] https://apps.fcc.gov/edocs_public/attachmatch/FCC-16-89A1.pdf
[2] https://apps.fcc.gov/edocs_public/attachmatch/FCC-16-89A1.pdf, para. 262
[3] https://apps.fcc.gov/edocs_public/attachmatch/FCC-16-89A1.pdf, para. 263
[4] https://apps.fcc.gov/edocs_public/attachmatch/FCC-16-89A1.pdf, para. 263

1) **3GPP direct network access** – In this category the user equipment attaches directly to the 3GPP network. For example, a smart phone attaching to a carrier's network.
2) **Generic Access Network** – In this category an API provides IP access into a carrier's core network, usually tunneled over IPSec. Data from the API is fed into the phone network as if it were coming from an antenna or a tower. Carriers enable WiFi-calling for their customers using generic access network connectivity.
3) **Capillary network access** – In this category the user equipment connects to a gateway that is on the 3GPP network. The device is authenticated by the gateway not the 3GPP network. The gateway is connected to the 3GPP network as described in category 1 or 2. The gateway can thereby have its own identity and multiplex the user equipment traffic over one data connection, or it can have multiple identities that map to the devices on the capillary network. It is more common for the gateway to have one identity with the 3GPP network.

4. Denial of Service (DoS)

5G is expected to enable diverse services in the mobile operators' networks, which include enhanced mobile broadband (eMBB), Internet of Things (IoT) and mission critical services. Mission critical services for UAVs (e.g., drones), vehicular networks or industrial systems (e.g., factory automation, process automation) in particular require highly available, low-latency, and highly reliable communication systems. In addition, IoT will introduce a large number of devices that can be low cost and less sophisticated than current mobile devices. As more devices are connected to the cellular networks, the networks will be exposed to denial of service (DoS) targeting the limited resources of specific services, much like botnet-driven distributed denial of service attacks in the Internet. Limited resources on cellular networks will include spectrum bandwidth, processing capacity of control functions (e.g., MME), processing capacity of user-plane functions (e.g., PDN gateway), and network bandwidth. Each of those resources may be a target of DoS attacks. It should be noted that service outage due to DoS attacks would pose substantial threats to mission critical services. 5G systems should have mechanisms to identity DoS attacks and to limit or mitigate the impact of such attacks. Potential attack vectors are:

- **Packet injection attacks** – compromised IoT devices are orchestrated to send packets simultaneously to overwhelm the network. Those devices may use radio resources allocated for other devices to inject bogus messages into the network. Such bogus packet injection cannot be effectively countered in the absence of an integrity check at the base station. Mobile-botnet driven DoS attacks may also become a significant threat.
- **External flooding attacks** – a high volume of unsolicited packets from the Internet may exhaust the bandwidth of mobile networks. Those packets may also be used to exhaust the battery of IoT devices. Such attacks may be launched by a botnet consisting of large number of bots (e.g., millions of malware infected devices) connected to the Internet.
- **Radio jamming attacks** – devices emit jamming signals to disrupt communication between a device and a base station or between devices (e.g., D2D). Jamming can be launched against control-plane signaling messages or user-plane data messages. Attackers may employ intelligent jamming in order to jam the radio spectrum persistently. For example, a presentation from this summer's Blackhat Conference[5] describes the use of a consumer UAV performing a jamming attack on stationary and moving targets.

[5] https://www.blackhat.com/docs/us-16/materials/us-16-Melrose-Drone-Attacks-On-Industrial-Wireless-A-New-Front-In-Cyber-Security.pdf

Recommendations

#1: It is recommended that 5G standards be defined in such a way as to enable resource isolation techniques such as network slicing to confine the effects of DoS attacks

Resources for different classes of traffic, services and devices can be isolated based on:

- Message type: Control plane traffic, user plane traffic
- QoS requirement: delay, jitter, bandwidth, priority
- Delivery requirement: guaranteed delivery, best-effort traffic
- Service type: Mission critical services (e.g., public safety), eMBB services, IoT services, vehicular services (e.g., V2X, V2I, V2V)

Network slicing and path diversity should be deployed based on classes of traffic, services, and/or devices to limit the effects of DoS attacks.

#2: It is recommended that 5G networks be able to deauthorize an individual device (or multiple devices) in such a way as the device does not continue to utilize the control plane or media plane resources

Deauthorization of devices essentially requires detection of DoS attacks and identification of devices used for the attacks. For DoS attack detection, the network should have an intrusion detection system that triggers alarms in the occurrence of attacks. Identification of devices used for the attack may require the following capabilities:

- **Packet accountability** – network should be able to identify the source of traffic. In the Internet, source address spoofing has been widely used for DoS attacks, e.g., DNS reflection/amplification attacks, to hide the origin of the traffic or concentrate the impact to the target. Countermeasures employed to mitigate the address spoofing attacks include ingress/egress filtering at routers. More recently, self-certifying network address (e.g., a device public key as a network address) has been proposed for the future Internet. A similar countermeasure needs to be considered for the cellular network to enhance packet/message accountability.
- **Device authentication/identification** – may be used as a way to prevent unauthorized, malicious devices from accessing a network even if those devices have valid subscription and corresponding credentials, e.g., based on (U)SIM credential, to attach to the network.
- **Device integrity** – it is desired for a network entity (e.g., a management entity) to verify device integrity, e.g., based on secure boot, to ensure that a device including hardware and software components are in trusted state.

#3: It is recommended that base stations have the ability to schedule the radio resource for each device in an unpredictable way

In 5G, a base station should schedule the radio resource for each device in an unpredictable way by other devices. This would significantly reduce the risk of jamming attacks targeting a specific set of devices, e.g., mission critical devices.

#4: It is recommended that 5G network elements embed DoS detection and mitigation functions into the RAN functions via key security indicators with related dynamic resolution

DoS detection functions would include a set of measureable security indicators. Examples of key security indicators are; detect/identify excessive Attach Requests beyond a certain threshold from an anomaly pattern of devices, and detect/identify an anomaly pattern of devices continuous streaming Uplink data beyond a certain threshold (>x). Key security indicators can also be attributes of the functions that monitor and detect performance/threshold alarms.

The related dynamic resolution will be the mitigation aspect of these functions.

5. Key Management

Communication for the Internet of Things (IoT) is often constrained today to using short-range communication protocols such as Bluetooth-LE, Near Field Communication (NFC), 802.11 WiFi, ZigBee, ZWave, etc. A typical IoT implementation within a home or business environment is a wireless sensor network, where deployed sensors can communicate between themselves using mesh networking capabilities. This communication can take place directly or proxied through a variety of external service gateways.

Typically, the nodes that participate in this meshed architecture are provisioned with cryptographic material that supports confidential, authenticated and integrity protected communications amongst themselves and to/through the gateway(s). The underlying cryptographic material and services required depends on the protocols that are being used (both communication and messaging) and the security objectives of each. For example, ZigBee-based communications require a network key (that is shared across nodes in the network, i.e., symmetric).

In addition to keys required for communication protocols, messaging protocols (e.g,. MQTT, CoAP, DDS) also levy cryptographic algorithms and key material. Although some messaging protocols only support username/password, many provide options for using symmetric keys, key pairs, and certificates to secure communication between devices.

With the introduction of 5G cellular technologies, IoT product developers will be able to redesign their products with broad, direct access to the cloud, with reduced dependencies on local gateways. With the introduction of 5G cellular technologies, IoT product developers will be able to redesign their products with broad, direct access to the cloud and new capabilities for peer to peer communications. This requires flexible key management capabilities that support myriad use cases.

Today's key management methods for 4G/LTE networks are based on symmetric keys. The carrier loads a pre-shared symmetric key in the AuC (network-side) at subscribe time and the USIM (user-side) at manufacture time. The mutual authentication of user and network results in a derived key, Access Security Management Entity (ASME), which is then used to derive additional encryption and integrity keys for the NAS, RRC Signaling and User Plane.

Symmetric keys typically suffer from a manage-ability problem as evidenced by the current processes' lack of flexibility in provisioning keys for the USIM. In order to support more flexible deployment models and the usage characteristics exhibited by IoT implementations, the Subcommittee provides the following recommendations.

Recommendations

#5: It is recommended that industry standard encryption techniques be used to protect data during transport

Proprietary encryption protocols should be excluded from industry standards.

Some of the key aspects that may be considered for encryption and authentication are listed below:

- Integrity of contents of communication or Secure 2-party communication

  - Application Level encryption
- Integrity of transport
  - Ensuring industry standard encryption techniques are utilized to protect data during transport. Avoiding using proprietary encryption protocols and ensuring the message payload encryption and secure encryption key handshaking
- Advanced packet filtering
- Content detection in the presence of encryption

#6: It is recommended that 5G networks provide options for using asymmetric key material to support diverse IoT Use Cases

With 5G and the explosion of "things" connecting to it, we need to consider alternative security provisioning solutions such as device certificates and some form of certificate management solution to provide more scalable means of provisioning trust between devices, networks and application domains.

#7: It is recommended that 5G networks enable privacy protections to guard against using keys and certificates to identify and track consumers

The IoT includes devices that can be bound to a particular person or property. The ability to identify a person or their actions through IoT devices must be protected. Strong protections against insiders within the key management systems and the cellular systems should also be put in place. For example, in the automobile industry connected vehicle design, certificate provisioning includes a pooling function for transaction signing to disallow tracking.

#8: It is recommended that 5G standards development consider alternative trust models that enable flexibility in establishing trust models across heterogeneous devices, access technologies, network domains and communication modes

Trust establishment is rigidly defined in 4G networks. 5G will require greater flexibility in being able to establish trust across heterogeneous devices, access technologies, network domains, and communication modes (e.g. human to device, device to device, device to infrastructure). Today's trust model solutions, e.g., PKI, that support mobile devices, web browsers and other applications may introduce difficult-to-scale trust management problems (via Trust Anchors) for 5G paradigms. Alternative, non-hierarchical, distributed trust models and technologies should be considered for 5G to maximize deployment model flexibility. Designing 5G capabilities such that relevant trust models can be "plugged and played" depending on the environment and use cases is ideal.

Greater communications flexibility is also needed. Device to device communications also calls for a means for devices to perform more dynamic peer-to-peer authentication with less dependency on the infrastructure. Such use cases require us to look beyond today's pre-provisioned symmetric key based solutions, and look toward other technologies such as PKI or Identity Based Encryption (IBE).

The use of alternative trust models also requires the ability to support flexible trust management features. IoT devices serviced by different telecommunication providers will have to interoperate. Some of these devices will establish trust relationships dynamically (e.g., a wearable health-monitoring device entering a hospital environment). This requires mechanisms for updating the trust stores that determine whether peer or infrastructure identities are trusted or not.

A relevant and timely example of alternative trust model considerations for 5G can be found in the recent efforts of the European Union's Horizon 2020 5G PPP 5G-ENSURE project[6]. Per the 5G-ENSURE website,

[6] http://www.5gensure.eu/project-vision

their vision statement is as follows: "*The 5G-ENSURE project brings to the 5G PPP a consortium of telco and network operators, IT providers and cyber security experts addressing priorities for security and resilience in 5G networks…..5G-ENSURE will define a shared and agreed 5G Security Roadmap with various 5G stakeholders. The outcome will be a trustworthy 5G system offering reliable security services to customers with a "zero perceived" downtime for service provision*."

From a trust model perspective, the 5G-ENSURE project has examined high level technology enablers and published a preliminary roadmap document (dated March 2016[7]. That document has defined several aspects of trust, all of which should be considered for 5G deployments[8]:

- Trust between automated systems (e.g. through advanced certificate and token based methods): that is M2Mt;
- Trust between human stakeholders holding responsibilities for different parts of 5G networks, between user and network operators and between users of the network (U2Ut);
- Trust that a human stakeholder has towards a system (U2Mt);
- Trust that an automated system (machine) has in users that it interacts with, such as whether it believes the user is who they claim to be (M2Ut).

For its first release targeted for September 2016, 5G-ENSURE envisions defining a trust model ontology to enable the consistent encoding of the assets, threats, and controls in 5G systems. This will then be used for modeling the system and ensuring the system is designed to mitigate threats as they relate to the complex and dynamic nature of trust across 5G system providers, users, and automated systems.

While the 5G-ENSURE effort is still early in its development and its scope is focused on the EU, the security topics being matured and documented are clearly relevant to the FCC's communicated areas of interest. The TAC's perspective of the EU 5G efforts like 5G-ENSURE is that such efforts may serve as useful technical solution references for US based 5G security standards activities. The intent of highlighting the EU activities is not to imply a desire to influence them, but rather to learn from their progress on addressing common technical challenges such as 5G trust model development.

As a supporting element for the alternative trust models recommendation stated above, it is suggested that both the FCC and the TAC regularly monitor future 5G-ENSURE progress for potential reuse for US focused 5G security recommendations. The first 5G-ENSURE security and privacy enablers release is scheduled for September 2016, Release 2 is scheduled for August 2017.

#9: It is recommended that 5G networks support new secure enrollment processes that allow entities other than carriers to provision enrollment certificates to devices

Flexibility will be key when it comes to provisioning. For example, homeowners may need a simple but secure means of linking their smart home devices together into one home network. Provisioning will also need to be very scalable and adaptable to different network configurations, due to large numbers of devices interconnecting and forming collaborative networks. Solutions will also need to facilitate and streamline transfer of ownership when devices are bought and sold in secondary markets.

Flexible generation capabilities are also needed. Some IoT products will generate their own key material and initiate certificate signing requests. Other devices may be provisioned with centrally-generated key pairs and associated certificates. The ability for the infrastructure to handle both models will be important.

---

[7] http://www.5gensure.eu/sites/default/files/Deliverables/5G-ENSURE_D3.1-5G-PPPSecurityEnablersTechnicalRoadmap_early_vision.pdf

[8] http://www.5gensure.eu/sites/default/files/Deliverables/5G-ENSURE_D3.1-5G-PPPSecurityEnablersTechnicalRoadmap_early_vision.pdf, Section 5

Some devices may also require multiple types of identities. Flexibility in supporting multiple types of identities when the use cases warrant such support could aid end users in securing their devices (e.g., optional support for signature, encryption, key encipherment certificates). This is especially useful for some IoT protocols that allow multiple profiles to be used, all hosted on a single node.

Support for ownership changeover is also important. Many consumer IoT devices will be integrated directly into a home or a vehicle. This means that the devices will change hands over time, e.g., when a home or vehicle is sold. The ability to bind and unbind the device to a new network and a new identity quickly and easily is important. The ability to bind and unbind a batch of certificates is also important, e.g., a home being sold and the need to rekey all IoT devices in that home. Non-repudiation assurances of the cryptographic keys and the key provisioning designs are crucial for a variety of 5G-enabled IoT use cases.

#10: It is recommended that 5G networks support robust methods for identifying and responding to misbehavior

Depending on their deployment environment, IoT device theft and other compromises may be common. Flexible methods for reporting device compromise and quickly cutting off authentication abilities for devices must be provided. Some devices will simply require an image update to restore to a non-compromised state, which means that the keys bound to a device would need to be revoked and then re-issued. The ability to efficiently perform this re-issuance online should also be explored for IoT devices that do not require higher levels of assurance.

#11: It is recommended that 5G networks support multiple devices that operate at multiple levels of sensitivity/assurance

Not all IoT products require the same levels of security assurance. Some IoT devices (e.g., connected vehicles, other Cyber Physical Systems) require stringent security controls and any keys or certificates issued to those devices must go through a robust identity vetting process. Other consumer devices may require less stringent identity vetting, and could even include self-service capabilities. Security models for identity provisioning should offer flexible options for the levels of identity assurance (identity setup and vetting) prior to certificate issuance, and ideally include a level of assurance attribute embedded in the certificate. System owners can then use that attribute to make access control decisions.

Differentiating between IoT devices can also be supported by embedding attributes of the device within a certificate. Within the IoT, the ability to differentiate between different types or classes of devices will be important within and across industries. In addition to understanding the levels of assurance provided by a particular certificate and its host device or application, it would also be useful to provide system owners with the ability to embed additional attribute information within identity certificates. For example, the ability to cryptographically bind the identity of an emergency services vehicle within a certificate used for authentication would be useful in allowing transportation infrastructures to make appropriate decisions (note - this is the current design of the IEEE 1609.2 certificate format).

The threat environment for different types of IoT devices will also drive the need to support flexibility in the lifetimes configured for keys and certificates used within the IoT. Just as with different levels of assurance associated with IoT device types, different types of IoT devices will have different product and usability lifetimes. Generally, higher assurance certificates (i.e., more rigorous vetting process) will be given shorter lifetimes, however privacy impact assessments performed by IoT device vendors and system owners may also drive the need for shorter lifetimes (e.g., automobiles).

6. Identity Management

Identity Management (IdM) is a broad administrative area that deals with identifying individuals, entities, or in general "principals" (e.g., humans, services, communication endpoints, or devices) in a system (e.g., a country, a network, a compute cloud, or an enterprise). Their established identity is typically the basis to accomplish further security goals, such as policy-based access control decisions to resources within that

system (e.g., granting access to licensed spectrum for communication based on the proper authentication of a post-paid IMSI) or recording of actions mapped to their actors to establish a non-repudiable transaction history (e.g., through blockchain-based transaction integrity preservation). The term “identity” is the relation each entity bears just to itself, while the term “identifier” is a name that labels the identity of a unique entity.

In the context of ICT networks there is a plethora of different types of identities involved, at every layer of the stack, in every segment of the architecture. For example, individual chips or IoT devices might need to be identified as hardware trust anchors, IP endpoints, cloud service instances, network services, virtualized network function instances, radio base stations, mobile devices or their subscribers, and administrators, to name just a few. All of those identities need to be defined, provisioned, maintained, validated, revoked, etc., in short *managed*. The term Identity Management captures the entire life cycle of this management task. The term Identity Access Management (IdAM) captures the additional dimension of using these identities for controlling access in a given context, and making and enforcing the access control decisions. Establishing identities requires a naming scheme to uniquely identify principals through identifiers. In most situations identifiers themselves are only primitive identities and need to be augmented through the association with authentication or identification procedures through which it can be proven that a principal is rightfully using an identifier. Authentication can be accomplished by a large number of different schemes and protocols, often employing secret information and cryptographic techniques. In fact, the concepts of identification and authentication are so tightly coupled in systems and in people’s minds that when we talk about identity management systems a large portion of implemented functionality deals with authentication, rather than identifiers. However, their conceptual separation is essential in environments where trust relationships have to be gained instead of assumed to exist beforehand. The separation also implies that identities consist of two parts; the identifier and the credentials the principal uses for the above mentioned proof.

Identities for end user-associated communication endpoints (e.g., smart phones, IoT devices, or over the top applications) are of primary importance in our industry, because they are control points for monetization, customer relationship management, and policy enforcement. At a coarse level we can distinguish between identities used for (over the top) applications/services and identities used for network access control (communications access) decisions. The latter is especially important when access is granted to licensed spectrum, but cannot be neglected for unlicensed spectrum, especially in scenarios, where the end-to-end communication between device and backend service traverses both types of networking technologies. The lines between application and network identities are not always so crisp, as in many scenarios identities are used for multiple purposes, e.g., an identified/authenticated smart phone is granted access to the spectrum for communication purposes based on its authenticated IMSI, but can also use voice, SMS, and general Internet connectivity services. Additional applications and services over the top (e.g., end-to-end application services, such as webmail, social networking services, or banking access) may employ additional, often service-specific identification schemes, at times linking them to their underlying communication identities (e.g., offered through GBA, deriving session credentials from SIM credentials; or GSMA Mobile Connect, matching users to their mobile phones). Many applications do not benefit from establishing their own, custom identity management but rather link to and reuse existing identity providers, including federated identity providers, lowering the barrier of adoption by new users, because users can now reuse their existing credentials (e.g., Facebook or Google credentials) for thousands of Internet applications. In turn, those application providers do not need to worry about managing credentials and have instant access to hundreds of millions of users.

The International Mobile Subscriber Identity (IMSI) and its associated credential, the secret Authentication Key (Ki) are used to identify and authenticate subscribers on mobile telephony and to a lesser extent on IoT devices. They are typically stored in a Subscriber Identity Module (SIM), an application on an integrated circuit chip, together with its integrated circuit card identifier (ICCID), and network operator-specific data. This circuit is called the Universal Integrated Circuit Card (UICC) and can be viewed as a physical smart card. UICC smart cards have undergone stepwise form factor reductions in four generations since their introduction in 1991 from Full-size, via Mini-SIM, Micro-SIM, to Nano-SIM in early 2012. Their external interface and interaction model remained the same though, allowing a subscriber identity to be moved between devices by moving the SIM chip that stored the related identifiers and credentials, often fondly referred to as “plastic roaming”. GSMA developed an embedded version of the UICC, called eUICC, initially for IoT devices, but most recently also for consumer devices. The primary differences being that the eUICC is soldered directly

onto the circuit board of the device, and that the identities and credentials (aka profiles) are provisioned over the air. The recently published Remote SIM Provisioning (RSP) architecture for consumer devices consist of a single service, the Subscription Manager Data Preparation+ (SM-DP+) for enabling the secure download, enablement, disablement and deletion of profiles on the eUICC. This architecture will be adopted to handle also IoT applications that today are served by a separate architecture for remote profile provisioning and management. Such over the air provisioning proves to be highly disruptive to the traditional business models of operators, but is an increasingly important trend for both IoT and consumer devices.

The industry has also been experimenting with Soft SIMs (or virtual SIMs). This is a selection of software applications and data that perform all the functionality of a SIM card but does not reside in any kind of secure data storage. Instead it would be stored in the memory and processor of the communications device, i.e., there would be no SIM hardware.

A subset of IoT devices is expected to be directly connected to the 3GPP network via the emerging standards of narrowband-IoT (NB-IoT). But there are many more local and wide area networking technologies (e.g., Sigfox, Bluetooth LE, LoRa on unlicensed bands) employed and being developed. Forecasts predict that non-3GPP connected devices will soon significantly outnumber the 3GPP-connected devices. These devices are connected to non-3GPP capillary networks and can be integrated into the 3GPP and IP network infrastructure via gateways. While some technologies use pre-shared key systems for identification and access control (e.g., Wifi), others employ raw public keys or public keys certified in a certification hierarchy, either private or public. A study group in 3GPP is investigating to what extent these technologies could and should be adopted in the emerging 5G security architecture.

Recommendations

#12: It is recommended that the 5G network that provides access to a device be able to uniquely identify, authenticate and authorize each individual device that accesses the network either directly or indirectly (e.g., via a gateway, virtual network)

UEs are the subscriber entry points into the 5G network and are perhaps the weakest element on the architecture as the MNO has little control over its security parameters. UEs can be the gateway for various security vulnerabilities into the 5G service. On the network side, we still have to think about issues such as Rogue eNodeBs or Eavesdropping/Man in the middle attack.

Attackers can take advantage of a known weakness wherein the user identity transference occurs unencrypted, in clear text between the UE and the eNodeB, during the initial attach procedure. This allows an eavesdropper to track the user cell-location or launch a man in the middle attack by user IMSI impersonation and relay of user messages.

#13: It is recommended that an equipment or subscriber identity that is transported across networks and presented to a terminating device be authenticated and authorized

Phone number spoofing has become a significant problem with the proliferation of VoIP networks. It is used to violate regulatory rules, such as those related to robocalling, and even to evade law enforcement when committing a crime, such as SWATing. The IETF has been working on solutions to provide authentication of an originating phone number in the STIR working group[9]. 5G networks should ensure that phone numbers and any other identifying information that is transported across networks and presented to a device can be authenticated and authorized.

#14: It is recommended that UE be able to authenticate the network before attaching

[9] https://datatracker.ietf.org/wg/stir/charter/

5G networks must have the ability for devices to reliably authenticate the network it is communicating with. 3GPP specifies access security in TS 33.203 which includes authentication related mechanisms and traffic protection between the UE and core networks. Strong encryption in the attach phase and UE authentication to the eNodeB will deter both rogue elements and man in the middle attacks. Adopting PKI with the public key of the MNO being stored in the USIM allowing the UE to encrypt privacy related information such as the IMSI transmitted to the eNodeB will enable confidentiality. Encryption should be implemented between the UE and eNodeB to thwart attackers leveraging IMSI paging and location identification vulnerabilities thus protecting subscriber privacy and security.

Future Considerations

- Homomorphic Encryption, allowing operations on encrypted data
    - Homomorphic encryption is a form of encryption that allows computations to be carried out on ciphertext, thus generating an encrypted result which, when decrypted, matches the result of operations performed on the plaintext.
- Private Information Retrieval (PIR)
    - Private information retrieval (PIR) protocol allows a user to retrieve an item from a source without revealing which item is retrieved. PIR is a weaker version of 1-out-of-n oblivious transfer, where it is also required that the user should not get information about other items in possession of the source.

#15: It is recommended that Soft SIMs deploy rigorous cybersecurity measures that can protect against attacks aimed at software applications

Traditional SIMs have the benefit of combined hardware and software security. SIMs stored as a software application will be attractive to hackers. To protect identity and credentials in Soft SIMs would require more extensive security measures than exist today.

7. Isolation Mechanisms

In 5G to achieve authorized access to the base station may require different trusted access mechanisms to SIMs - especially for off-loads. A method may be required not only to identify a network user, but also their location, mobility tracking and data usage attribution.

An ideal approach would leverage network slicing combined with cognitive computing in each base station, local storage in each base station, local networking in every base station plus random-number based encryption coding and recoding during transmission that can only be decoded by the intended recipients.

Taking user plane security as an example, some applications may not want to rely on security provided by the network, but may rather use end-to-end application level security. Underlying network-terminated security would not provide a higher degree of security to the applications, but may have an impact on delay or resources on the terminal. Other applications, however, may want to rely on user plane security supported by the network, and may even need user plane integrity protection in addition to encryption.

The energy cost of encrypting one bit is one or two orders of magnitude less than transmitting one bit. However, for the most constrained, battery dependent devices with a long target life time, there may be a need to consider even more lightweight solutions, as every micro joule consumed could be of importance.

Recommendations

#16: It is recommended that 5G standards be defined in such a way as to enable resource isolation techniques such as network slicing to enable different levels of security among different resources

Network slicing can be an important tool to handle the very diverse requirements of different applications and user groups. By having a properly implemented, high-assurance isolation mechanism to support slicing, it is possible to confine the impact of security requirements to single slices, rather than the whole network. The cost of high assurance and certification can therefore be concentrated onto an infrastructure virtualization/isolation layer.

Another option worth considering is simply putting the responsibility in the endpoints, i.e., in connected devices or organization data centers. Data security is an example of a service that could be handled this way. Besides the isolation/slicing itself, many other examples of network-enabled security as a service can be attractive to multiple user groups, including network enforced security policies, authentication, key management and data security services.

#17: It is recommended that there be access to the control plane and media plane at the base station to enable security monitoring of traffic

Anomaly detection will be an important tool to identify potential attacks. The closer the monitoring is to the source, i.e., base station, the greater the opportunity is to limit the attack to a smaller part of the network.

Acronym List

Acronym	Meaning
3GPP	3rd Generation Partnership Project
2G	2nd Generation mobile network
3G	3rd Generation mobile network
4G	4th Generation mobile network
5G	5th Generation mobile network
API	Application Program Interface
ASME	Access Security Management Entity
AuC	Authentication Center
CN	Core Network
CoAP	Constrained Applications Protocol
CPS	Cyber Physical Systems
D2D	Device to Device
DDS	Data Distribution Service
DNS	Domain Name System
DoS	Denial of Service
eMBB	Enhanced Mobile Broadband
eNodeB	evolved Node B
EU	European Union
eUICC	Embedded Universal Integrated Circuit Card
FCC	Federal Communications Commission
GBA	Generic Bootstrapping Architecture
GSMA	GSM Association
HSS	Home Subscriber Server
IBE	Identity Based Encryption
ICCID	Integrated Circuit Card Identifier
ICT	Information and Communications Technology
IdAM	Identity Access Management
IdM	Identity Management
IEEE	Institute of Electrical and Electronics Engineers
IIC	Industrial Internet Consortium
IMSI	International Mobile Subscriber Identity
IoT	Internet of Things
IP	Internet Protocol
IPsec	Internet Protocol Security
Ki	Authentication Key

LoRa	Long Range
LTE	Long-Term Evolution
M2Mt	Machine to Machine trust
M2Ut	Machine to User trust
MQTT	Message Queuing Telemetry Transport
MME	Mobility Management Entity
MNO	Mobile Network Operator
NAS	Non-Access Stratum
NB-IoT	Narrowband Internet of Things
PDN	Packet Data Network
PIR	Private Information Retrieval
PKI	Public Key Infrastructure
RAT	Radio Access Technologies
RRC	Radio Resource Control
RSP	Remote SIM Provisioning
SDO	Standards Development Organization
SIM	Subscriber Identity Module
SM-DP	Subscription Manager – Data Preparation
SMS	Short Message Service
STIR	Secure Telephone Identity Revisited
TAC	Technological Advisory Council
UAV	Unmanned Aerial Vehicle
U2Mt	User to Machine trust
U2Ut	User to User trust
UE	User Equipment
UICC	Universal Integrated Circuit Card
US	United States
USIM	Universal Subscriber Identity Module
V2I	Vehicle to Infrastructure
V2V	Vehicle to Vehicle
V2X	Vehicle to Everything
VoIP	Voice over Internet Protocol
WG	Working Group