5G Requirements and Key Performance Indicators

Toon Norp

Chairman of 3GPP SA1, Senior Business Consultant at TNO, Netherlands
E-mail: toon.norp@tno.nl

Abstract

This paper presents an overview of 5G requirements as specified by 3GPP SA1. The main drivers for 5G were the requirement to provide more capacity and higher data rates and the requirement to support different 'vertical' sectors with ultra-reliable and low latency communication. The paper discusses basic requirements that are new for 5G and provides 5G performance requirements. The paper also discusses a number of vertical sectors that have influenced the 5G requirements work (V2X, mission critical, railway communication) and gives an overview of developments in 3GPP SA1 that will likely influence 5G specifications in the future.

1 Introduction

Work in 3GPP on 5G started in 2015 with a "Feasibility Study on New Services and Markets Technology Enablers" in the 3GPP Working Group SA1, which is responsible within 3GPP for services and feature requirement specification. The study first collected a large number of use cases – 74 in total – that illustrate what new capabilities are required from 5G [1]. The study took inspiration from a number of whitepapers from NGMN [2], the European 5G-PPP [3],

the China IMT2020 project [4], 4G Americas [5], the GSMA [6] and the Japanese standardisation development organisation ARIB [7].

A clear driver for 5G that emerged from all the whitepapers was the enormous growth of mobile data. Figure 1 shows mobile data growth projections from ARIB [7] which indicate that between 2020 and 2025 at the start of 5G mobile data traffic will have grown 1000x from 2010, at the start of 4G. 5G technology will have to be able to support network operators to cope with that data growth. It is clear that a 1000 times growth of data volume cannot lead to a similar growth of costs of energy consumption; 5G will have to be more efficient than earlier generations.

Not only the data volumes are expected to rise further, also the number of devices continues to grow. Expectations for the number of connected devices in the Internet-of-Things vary significantly; but all projections agree that in the 5G era several billions of devices will be connected.

The growth in mobile data volumes is largely related to the popularity of video applications. This trend is expected to continue in 5G. With the advance of technology the applications become ever more bandwidth demanding. Where users used to enjoy watching a funny video of a cat in low resolution; in the future they want to see similar videos in 3D, Virtual Reality, or Ultra High Definition. If also large TV screens are connected wirelessly, then 5G will need to deliver much higher data rates compared to 4G in order to support the demand for video applications.
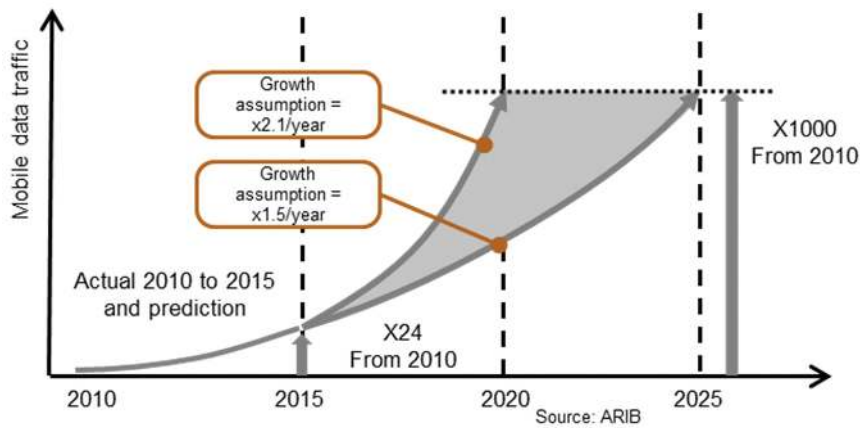

Figure 1 Projected growth of mobile data [7].

The inspiration for 5G was not only in providing a more efficient and higher data rate version of 4G. As is shown in the name of the 3GPP SA1 study – "New Services and Markets Technology Enablers" – 5G is also very much about enabling new types of applications. In a digital society, consumers, governments, corporations and industries will make use of mobile telecommunications to improve all kinds of processes. These so called 'vertical' sectors, often use specific applications, with a diverse set of requirements on mobile telecommunications. Rather than designing specific wireless technology for each of these vertical applications, it is expected that 5G technology is flexible enough to support all kinds of applications, even together on a single network.

A common characteristic of vertical applications is that they have higher demands on reliability, availability and coverage. It is a nuisance when consumers cannot make use of mobile telecommunications to watch video-clips or update their social media. However, when payment terminals, trains, or the control of the electricity network become reliant on mobile telecommunications, then an outage of a mobile network has more far reaching consequences. For applications such as e-health, excellent coverage – both indoor and outdoor – is essential. It is not really possible to send a patient home with a hearth-monitor, if that has to come with a warning not to go into cellars or not to go camping.

Another aspect that 5G is improving, to cater for vertical applications, is the end-to-end latency. Many applications use control loops that will not work if the latency of data communication between sensors, controllers and actuators is too high. With current 4G networks, worst case latency in case of congestion can be up to a second. It is clear that this is not good enough if you are relying on mobile telecommunications for example to control an electricity network.

The 74 use cases that resulted from the first phase of the 3GPP SA1 feasibility study were consolidated in four different Technical Reports [8–11], each of which focuses on a different use case category (see Figure 2). The categories enhanced Mobile Broadband, Critical Communications, and massive Machine Type Communications are loosely based on similar categories introduced by the ITU. The main difference is that 3GPP SA1 also defined a category Network Operations. This category does not focus on performance requirements of specific services, but more on operational requirements that 5G networks have to fulfil.
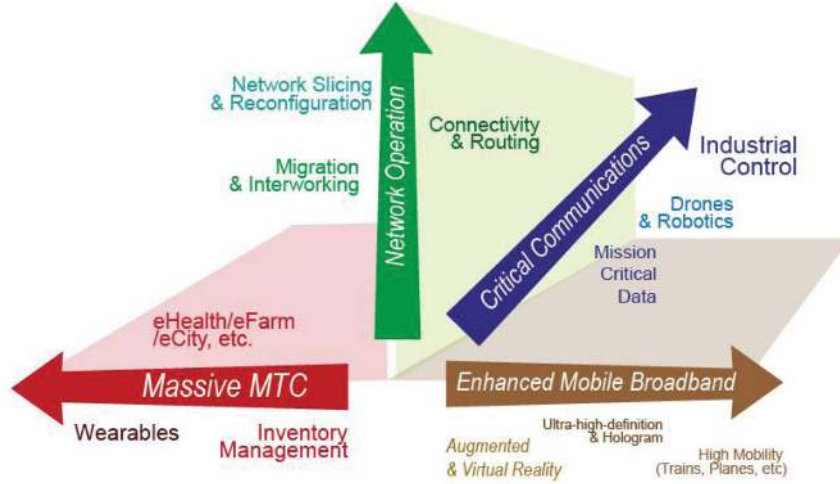
Figure 2 The four use case categories.

The four Technical Reports were subsequently consolidated in a normative 3GPP Technical Specification 22.261 "Service requirements for next generation new services and markets" [12]. This specification forms a basis for the 5G work in other 3GPP Working Groups. Note that not all requirements in [12] will be implemented in 5G Phase 1 (Release 15). 3GPP has concluded that it is unrealistic to implement all functionality for 5G in a single 3GPP release. However, this phased implementation has not yet been taken into account in the definition of requirements in [12].

In the remainder of this paper, first the 5G functional requirements from [12] are introduced in Section 2. Then Section 3 focuses on the 5G performance requirements. Section 4 discusses a number of vertical sectors that had a strong influence on the development of 5G requirements in 3GPP. Section 5 concludes with some future developments in 3GPP SA1 that are aimed at 5G Phase 2 (Release 16).

2 Basic Capabilities

5G is partly an evolution of existing mobile technologies. That implies that a lot of the functionality commonplace in earlier releases is also provided in 5G. The 5G requirement specification [12] only lists new functionality that is added in 5G; for existing functionality [12] refers to requirements specifications for

3G and 4G. The remaining part of Section 2 provides an overview of what functional requirements are specified for 5G.

2.1 Network Slicing

With the concept of network slicing (see Figure 3), operators can customise their network for different applications and customers. Slices can differ in functionality (e.g., priority, policy control, and security), in performance requirements (e.g., latency, availability, reliability and data rates), or they can serve only specific users (e.g., Public Safety users, corporate customers, or industrial users). A network slice can provide the functionality of a complete network, including radio access network and core network functions. One network can support one or several network slices.

A 5G user equipment can provide assistance information to enable the network to select one or more network slices. The network operator controls what slices should be available to a specific device and associated subscription.

Given the multitude of use cases for new verticals and services, a specific network may support only a subset of the vertical industries and services. However, this should not prevent an end-user from accessing all new services and capabilities. Therefore, the 5G system shall enable users to obtain services from more than one network simultaneously on an on-demand basis.
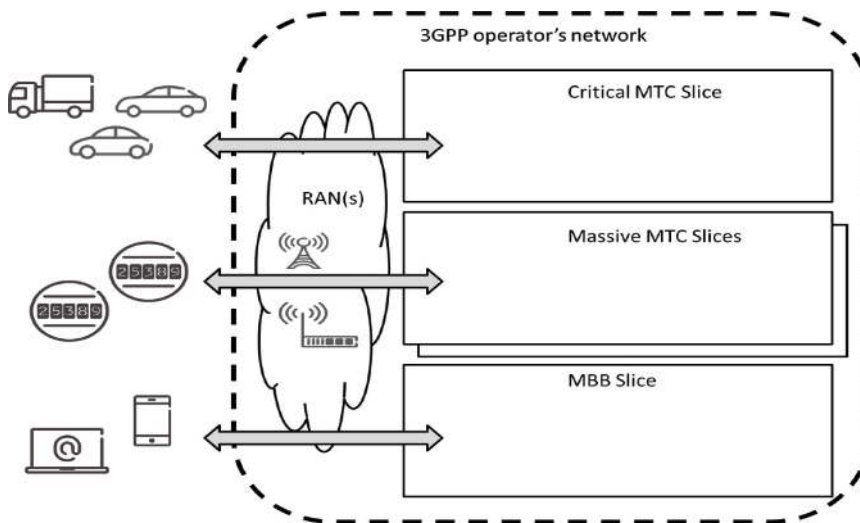
Figure 3   The concept of slicing [1].

2.2 Efficiency

5G will support diverse devices and services with different performance (e.g., high throughput, low latency and massive connection densities) and data traffic models (e.g., IP data traffic, non-IP data traffic, and short data bursts). In order to do this efficiently, 5G needs to be optimized for these different requirements.

For Internet of Things based applications, optimizations are needed to handle very large numbers of devices. Configuration, deployment, and use of IoT devices may benefit from optimizations such as bulk provisioning, resource efficient access, and optimization for device originated data transfer.

Sensors send data packages ranging in size from a small status update to streaming video. Smart phones similarly generate widely varying amounts of data. Where 4G was designed mostly with large amounts of data in mind; 5G will also have to efficiently support short data bursts without the need for lengthy signalling procedures before and after sending a small amount of data.

Cloud applications like cloud robotics perform computation in the network rather than in a device. This requires low end-to-end latencies and high data rates. The 5G system optimizes the user plane resource efficiency for such scenarios by locating operator or third party provided applications in a service hosting environment close to the end user. Video-based services (e.g., live streaming, Virtual Reality) and personal data storage applications have generated a massive growth in mobile broadband traffic. In-network content caching, provided by the operator, third party or both, can improve user experience, reduce backhaul resource usage and utilize radio resource efficiently for such applications. These optimization efforts also contribute to achieving higher reliability.

Energy efficiency is a critical issue in 5G. It is clear that a 1000 times increase of mobile data traffic from 4G to 5G cannot imply a similar increase of energy usage. Mobile operators are already one of the most significant users of electricity in many countries. For devices, energy efficiency translates directly into battery standby time. Small form factor devices also typically have a small battery and this not only puts constrains on general power usage but also implies limitations on both the maximum peak power and continuous current drain.

2.3 Diverse Mobility Management

The flexible nature of 5G will support different mobility management methods that minimize signalling overhead and optimize access for user equipment with different mobility management needs. Devices may be;

- stationary during their entire usable life (e.g., sensors embedded in infrastructure),
- stationary during active periods, but nomadic between activations (e.g., fixed access),
- mobile within a constrained and well-defined space (e.g., in a factory), or
- fully mobile.

Furthermore, different applications have varying requirements for the network to hide the effects of mobility. Applications such as voice telephony rely on the network to ensure seamless mobility. Applications such as video streaming on the other hand have application layer functionality (e.g. buffering) to handle service delivery interruptions during mobility. These applications will still require the network to minimize the interruption time.

Because of the much more distributed nature of 5G networks, mobility also has an impact in the network. With IP traffic offload or service hosting close to the network edge, mobility of a device also implies that the anchor node in the network may need to be updated. Internet peering and service hosting will have to follow the device when it is travelling across the network coverage area.

2.4 Multiple Access Technologies

The 5G system will support multiple 3GPP access technologies; next to one or more 5G New Radio (NR) variants, it will also include 4G radio interface technology (E-UTRA). Furthermore, 5G will support various non-3GPP access technologies.

Interoperability and integration among the various access technologies is important. 5G will select the most appropriate 3GPP or non-3GPP access technology for a service, taking into account e.g., service, traffic characteristics, radio characteristics, and the speed at which a device is moving. A single device can simultaneously use multiple access technologies, adding or dropping the various access connections as and when appropriate.

To support coverage even at sea or in remote areas and to improve availability in disaster scenarios, 5G shall also be able to provide services via satellite access. Service continuity is required between land based 5G access and satellite based access owned by the same operator or based on an agreement between the operators.

5G will also support fixed broadband access. A fixed access residential gateway can be a relay device, forwarding 5G connectivity to other end user

devices within the premises. Alternatively, one or more home base stations are connected to the fixed access, which can then provide 5G coverage within the home.

2.5 Priority, QoS and Policy Control

The 5G network will support many commercial services and regulatory services (e.g., Public Safety communication) that need priority treatment. Some of these services share common QoS characteristics such as latency and packet loss rate, but may have different priority requirements. Mobile telephony and voice based services for Public Safety share common QoS characteristics, yet may have different priority requirements. The 5G network will have to be able to decouple the priority of a particular communication from the associated QoS characteristics such as latency and reliability. The traffic prioritisation may be enforced by adjusting resource utilization or pre-empting lower priority traffic.

As 5G is expected to operate in a heterogeneous environment with multiple access technologies, multiple types of devices, etc., it should support a harmonised QoS and policy framework that applies to multiple accesses. Furthermore, 5G QoS needs to be end-to-end (including radio access, back-haul, core network, and network to network interconnect) to achieve the 5G user experience (e.g., ultra-low latency).

2.6 Connectivity Models

5G will support different connectivity models. Next to direct network connections, 5G will also support indirect network connections (see Figure 4). With indirect network connections, a remote device can connect to the network via a relay device. Indirect network connections can be used to connect wearables via a mobile phone, but can also be used to improve indoor coverage, connecting e.g. printers or consumer electronic devices. The relay device can access the network using 3GPP or non-3GPP access technologies (e.g., WLAN access, fixed broadband access). Also for the connection between remote device and relay device 3GPP or non-3GPP radio technologies can be used.

When a remote device attempts to establish an indirect network connection, there can be multiple relay devices in proximity to choose from. A discovery and selection mechanism needs to be supported to select an optimal relay device for the remote device.
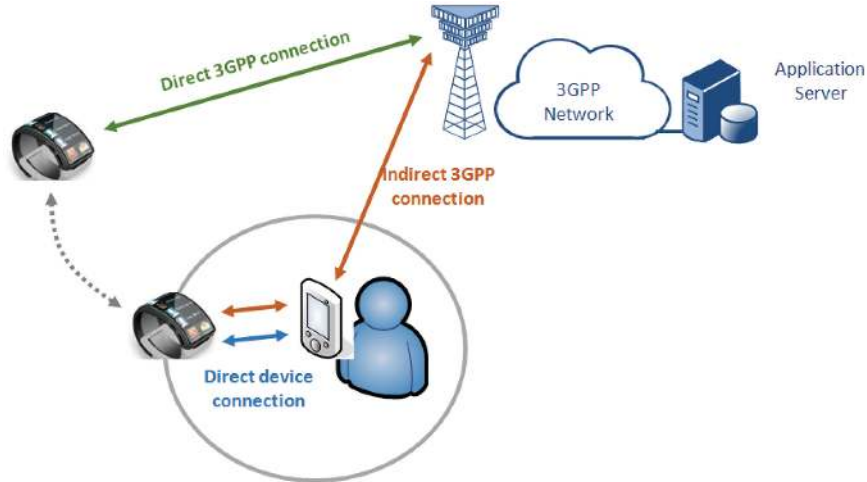
Figure 4 Connectivity modes for devices [8].

2.7 Network Capability Exposure and Context Awareness

Network capability exposure enables third party providers to interact with an operator network. With the advent of 5G, new network capabilities need to be exposed to the third party (e.g., to allow the third party to customize a dedicated network slice or to allow the third party to manage an application in a service hosting environment).

If network conditions can be provided to applications through network capability exposure, the applications can adjust resource usage to suit the network. At times when resources are scarce the application can be frugal with its resource usage. When resources are plenty the applications can compensate and use extra resources. Operators may provide incentives for applications to make their resource usage more network friendly, thus sharing the advantage of reduced network investments with the third party providers.

Applications may also provide the network with context information. For example, radio resource management can be optimised if the network is informed about application characteristics (e.g. expected traffic over time). Other characteristics of the device such as mobility, speed, battery status can be used to optimize allocation of functionality and content in the network.

2.8 Flexible Broadcast/Multicast Service

The proliferation of video services, ad-hoc multicast/broadcast streams, software delivery over wireless, group communications and broadcast/multicast IoT applications have created a need for a flexible broadcast/multicast service. Such a flexible broadcast/multicast service should allow flexible and dynamic allocation of radio resources between unicast and multicast services within a network, but also the deployment of stand-alone broadcast networks. It should be possible to stream multicast/broadcast content efficiently over wide geographic areas as well as target the distribution of content to very specific geographic areas spanning only a limited number of base stations.

3 5G Performance Requirements

Performance requirements highly depend on traffic scenarios. In an indoor hotspot scenario (e.g. for an office), the focus is on providing high data rates and high capacity. In a rural scenario, the focus is more on providing coverage. The data rates for a rural scenario will be lower, but 5G should ensure that a minimum data rate is available also in urban and rural macro scenarios. Table 1, shows the different performance requirements for the basic scenarios from indoor, dense urban, urban macro to rural macro. There are also scenarios for specific situations, such as broadband access in a crowd, broadcast, and connectivity in high speed trains, vehicles and airplanes.

Network access also needs to be supported in more extreme scenarios, with long range coverage, or in low end market scenarios, where access to power and backhaul facilities are not a given. Very large cell coverage areas of more than 100 km radius shall be supported with 1 Mbps downlink at cell edge. For constrained circumstances, and even larger areas, 5G shall be able to support a minimum user experience with 100 kbps, end-to-end latency of 50 ms, and a lower availability of 95%.

For vertical applications, other performance requirements are more important than data rates. For industry applications, the end-to-end latency is crucial. Motion control will not work if the time it takes to send information from a sensor to a controller is too long. Reliability – the percentage of packets successfully delivered within the time constraint – and communication service availability – the percentage of time the end-to-end communication service is delivered according to an agreed QoS – are crucial requirements for many industrial applications. Table 2 shows performance indicators for a number of vertical scenarios with low latency and high reliability requirements.

Table 1 5G performance requirements for high data rate and traffic density scenarios [12]

Scenario	Experienced Data Rate (Down-link)	Experienced Data Rate (Uplink)	Area Traffic Capacity (Down-link)	Area Traffic Capacity (Uplink)	Overall User Density	UE Speed
Indoor hotspot	1 Gbps	500 Mbps	15 Tbps/km^2	2 Tbps/km^2	250 000/km^2	Pedestrians
Dense urban	300 Mbps	50 Mbps	750 Gbps/km^2	125 Gbps/km^2	25 000/km^2	Pedestrians and users in vehicles (up to 60 km/h)
Urban macro	50 Mbps	25 Mbps	100 Gbps/km^2	50 Gbps/km^2	10 000/km^2	Pedestrians and users in vehicles (up to 120 km/h
Rural macro	50 Mbps	25 Mbps	1 Gbps/km^2	500 Mbps/km^2	100/km^2	Pedestrians and users in vehicles (up to 120 km/h
Broadband in a crowd	25 Mbps	50 Mbps	3,75 Tbps/km^2	7,5 Tbps/km^2	500 000/km^2	Pedestrians
Broadcast-like services	Maximum 200 Mbps (TV channel)	Modest (e.g., 500 kbps per user)	N/A	N/A	15 TV channels of 20 Mbps	Stationary to in vehicles (up to 500 km/h)
High-speed train	50 Mbps	25 Mbps	15 Gbps/train	7,5 Gbps/train	1000/train	Users in trains (up to 500 km/h)
High-speed vehicle	50 Mbps	25 Mbps	100 Gbps/km^2	50 Gbps/km^2	4000/km^2	Users in vehicles (up to 250 km/h)
Airplanes connectivity	15 Mbps	7,5 Mbps	1,2 Gbps/plane	600 Mbps/plane	400/plane	Users in airplanes (up to 1000 km/h)

Table 2 5G performance requirements for low latency and high reliability scenarios [12]

Scenario	End-to-End Latency	Communication Service Availability	Reliability	User Experienced Data Rate	Connection Density	Service Area Dimension
Discrete automation – motion control	1 ms	99,9999%	99,9999%	1 Mbps to 10 Mbps	100 000/km^2	100 × 100 × 30 m
Process automation – remote control	50 ms	99,9999%	99,9999%	1 Mbps to 100 Mbps	1000/km^2	300 × 300 × 50 m
Process automation – monitoring	50 ms	99,9%	99,9%	1 Mbps	10 000/km^2	300 × 300 × 50
Electricity distribution – medium voltage	25 ms	99,9%	99,9%	10 Mbps	1000/km^2	100 km along power line
Electricity distribution – high voltage	5 ms	99,9999%	99,9999%	10 Mbps	1000/km^2	200 km along power line
Intelligent transport – infrastructure backhaul	10 ms	99,9999%	99,9999%	10 Mbps	1000/km^2	2 km along a road

Low latency requirements come with specific service area dimensions as low latency communication is only possible when source and destination are nearby. The speed of light in optical communication – approximately 1 ms per 200 km – makes that network transport is not negligible, certainly not when latency caused by routers, switches and servers is taken into account. For the very low latency requirements of motion control, Table 2 specifies a service area dimension of a factory (100×100 m). This implies communication from a sensor in the factory to a controller that is also located within the factory; a much more distributed network topology than what is commonplace in 4G networks.

4 Vertical Applications

The support for vertical applications is one of the main goals of 5G. However, 3GPP has seen the introduction of vertical applications already in the 4G era. Though these early vertical applications can be supported in 4G, they have played an important role in the definition of 5G as well.

The first vertical in 3GPP SA1 was Mission Critical Communication. Targeted mainly at the Public Safety community, services like Mission Critical Push to Talk [13], Mission Critical Video [14] and Mission Critical Data [15] have been defined. For mission critical services, it is very important to ensure that communication remains available even in congestion situations. It cannot be that a fireman cannot communicate with other firemen because spectators are uploading videos of a fire. This implies that prioritization of mission critical services has to be provided. Recently, the mission critical community has explicitly indicated that requirement specifications for mission critical services shall also apply for 5G.

V2X (Vehicle to Anything) communication is a vertical application that has influenced 5G requirements from the start. The first V2X requirements are intended to be supported already by LTE. However, at the same time as the 5G requirements specification were completed, 3GPP SA1 also completed a 3GPP Technical Specification on enhancements of 3GPP support for V2X scenarios [16]. These functional and performance requirements in [16] are part of the overall 5G requirements.

Railway communications are the latest vertical application that introduced 5G Phase 1 requirements. European operators are looking for a replacement of their GSM-R networks in a timeframe up to 2030. Also in other countries, notably Korea, cellular network based technology for railway communication

is planned. Use cases for railway communications, together with potential requirements are identified in [17]. Most of the resulting requirements find their way in updates of the mission critical requirements [13–15].

5 Outlook to Future Requirements

After the completion of Release 15 requirements specifications in June 2017, 3GPP SA1 has been working on further enhancements of 5G in the Release 16 time frame. A number of studies and work items that are intended to be incorporated in 5G Phase 2 are listed below:

- *LAN support in 5G* aims to support 5G LAN-type services over the 5G system. In this context, 5G LAN-type services allow a restricted set of devices to communicate amongst each other using Ethernet style data transport.
- *Communication for automation in vertical domains* identifies key performance indicators for various vertical use cases. Communication for vertical sectors may take place in separate, privately owned networks.
- *Using satellite access in 5G* is a study on how to integrate satellite communication with land based 5G networks. An example is how to support network selection when multiple land based mobile networks share a common satellite based access.
- *5G message service for massive IoT* aims to specify a light weight message service that can be used between (groups of) devices or between devices and application servers.
- *Positioning use cases* identifies new use cases, their scope and environment of use along with the related key performance indicators. Requirements can be achieved with a combination of 3GPP and non-3GPP positioning technologies.
- *Enhancements to IMS for new real time communication services* identifies a number of use cases (e.g. Augmented and Virtual Reality telepresence), where IMS and/or mission critical specifications need to be enhanced for new 5G real time communication services.
- *Layer for centric identifiers and authentication* aims to enable network operators to become identity providers. Use cases include how to use the new user identifier within the 3GPP system e.g. to provide customized services and how to provide this identifier to external parties to enable authentication for systems and services outside 3GPP.

References

[1] 3GPP TR 22.891, "Study on New Services and Markets Technology Enablers"

[2] NMGN 5G Whitepaper v1.0

[3] European Commission's 5G PPP "5G Vision", Feb. 2015. www.5g-ppp.eu.

[4] China IMT2020 (5G) Promotion Group white paper "5G Concept" February 2015.

[5] 4G Americas' Recommendations on 5G Requirements and Solutions, October 2014.

[6] GSMA, "Understanding 5G: perspectives on future technological advancements in mobile", Dec., 2014.

[7] ARIB 2020 and Beyond Ad Hoc Group White Paper, October 2014.

[8] 3GPP TR 22.861, "Feasibility Study on New Services and Markets Technology Enablers for Massive Internet of Things; Stage 1".

[9] 3GPP TR 22.862, "Feasibility Study on New Services and Markets Technology Enablers – Critical Communications; Stage 1"

[10] 3GPP TR 22.863, "Feasibility Study on New Services and Markets Technology Enablers – Enhanced Mobile Broadband; Stage 1"

[11] 3GPP TR 22.864, "Feasibility Study on New Services and Markets Technology Enablers – Network Operation; Stage 1"

[12] 3GPP TS 22.261, "Service requirements for next generation new services and markets"

[13] 3GPP TR 22.179, "Mission Critical Push to Talk (MCPTT); Stage 1"

[14] 3GPP TR 22.281, "Mission Critical Video services"

[15] 3GPP TR 22.282, "Mission Critical Data services"

[16] 3GPP TR 22.186, "Service requirements for enhanced V2X scenarios"

[17] 3GPP TR 22.889, "Study on Future Railway Mobile Communication System"

Biography

Toon Norp is a Senior Business Consultant at TNO. Toon Norp joined TNO (former KPN Research) in 1991, where he has since been working on network aspects of mobile communications. Toon advises European operators on strategy, and architecture related to mobile core network, M2M/IoT, and 5G. He has been involved in standardisation of mobile networks for more than 20 years, and is the chairman of the 3GPP SA1 service aspects working group. Toon is member of the 5G-PPP association, a joint initiative between the European ICT industry and the European Commission to support the research and development of 5G infrastructures. Toon holds a Master's Degree in Electrical Engineering from the Eindhoven University of Technology, The Netherlands.