Using layers for policy analysis: 5G Technologies

Introduction

This module builds on Framework: Tech, layers and (un)bundling. You find this module here: https://docs.google.com/document/d/1-9hXabsoL94MeRi3D60r6CpUc62y3oe2dmkH9FSEVbI/edit#

Other parts in this series include Net neutrality (https://docs.google.com/document/d/1CUr-h5WayWRWuUEIKqkqF9UMEuUFw1G9BkcUzIcoQnc/edit) and Smartphone apps (https://docs.google.com/document/d/1Uo6iT3NjA4ONczWag-OipL-BjnLsmjMQH9l__vTDlco/edit).

A companion glossary is available here:

https://docs.google.com/document/d/1fxsbRxBYkSzh0stmc9liXTljqYIpQ7fdAH_rC31-zJI/edit?usp=sharing

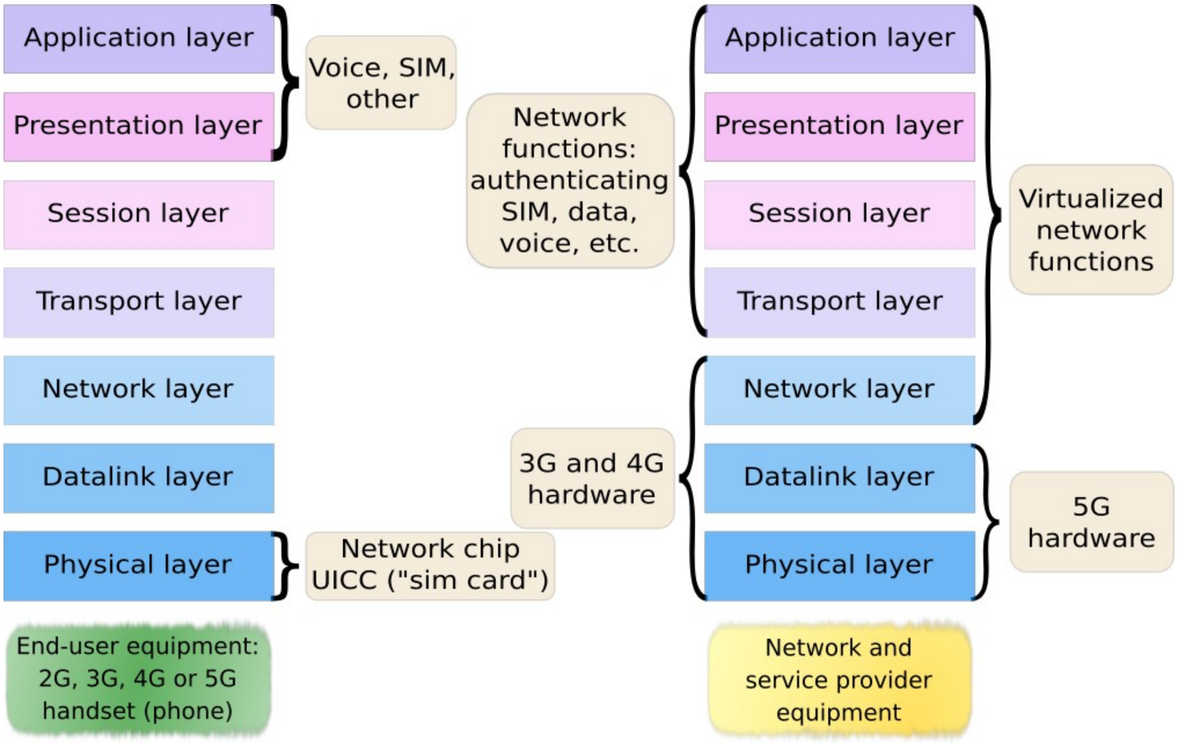
5G technologies

5G is the marketing term for a new set of technologies that conform with the ITU's International Mobile Telecommunications (IMT)[1] requirements for 2020 (IMT-2020). These requirements are established by governments and large companies in ITU-R, which is a subgroup of the ITU that which is responsible for radio communications. While any technology that conforms with the IMT-2020 requirement can call itself 5G capable, the marketing around 5G has essentially restricted it to *mobile network communications*.

There is currently a wide, and sometimes confusing debate, about the role that 5G will play in the future of mobile communications and the Internet of Things (IoT), with many alarmed calls for action concerning its possible impact on human rights. In the following paragraphs we will analyse the changes which result from 5G from the OSI perspective, which will enable an assessment of its possible indirect effect on human rights.

Mobile network equipment vendors promise many new capabilities in 5G. One of them is the increasing reliance on *software-defined networking*. This means, essentially, that functions in the network that were previously reliant on physical equipment (layer 1-3) will now be reliant on software defined components (layers 2-5). That in turn means that they will be easier to update and easier to centralize. Physical equipment will become less important, while control over the management software will be more important.

[1] See Annex: Glossary.



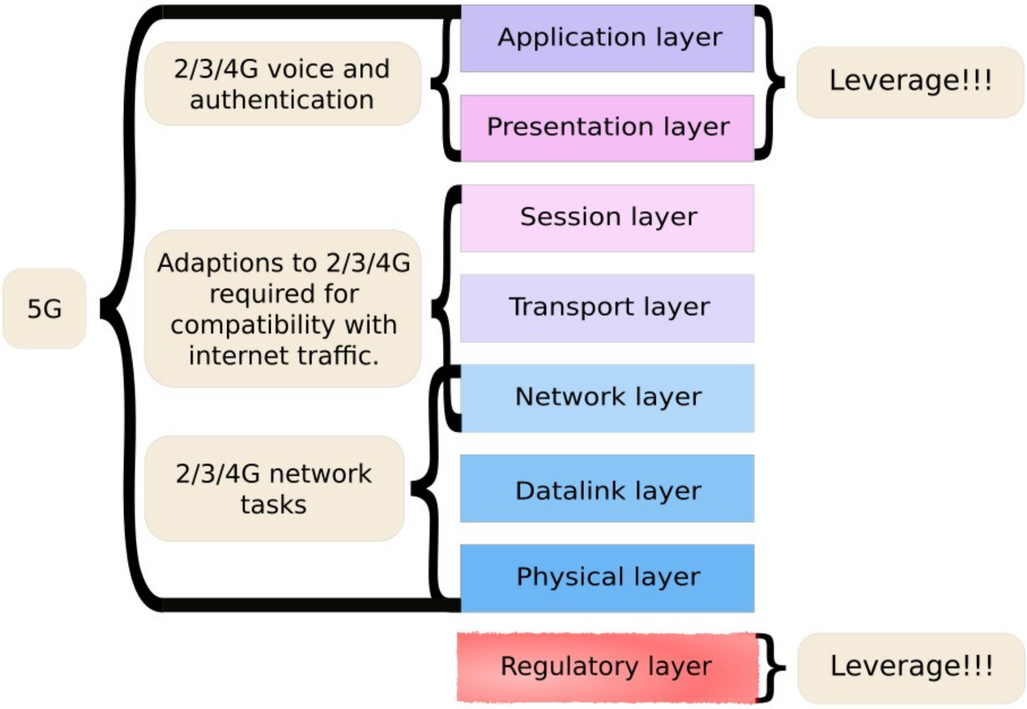
Controversies and marketing around 5G in large part reflect the differences between mobile network technologies and internet technologies since the 1990s:

1. Mobile network technologies are *network-centric*. They sell *an idea of a network*, including its management and all of its capacities and services. In this sense, mobile network technologies have built-in *barriers to entry*. Since the mobile network decides what kind of network services are available on the network, there can be no separate innovation on network layer 3 services. Since mobile networks can only provide services to authenticated users, any mobile network service on OSI layers 4-7 relies on network functions on layers 4-7.

   This is different from internet technologies, which rely on strict layering to allow separate innovation and service development on each layer. A constructive example is the almost total dominance of Cisco in the network layer segment of technologies in the 1990s, which ended in the early 2000s with the market entry of several strong competitors (for instance Juniper).

2. Mobile network technologies in general specify features in the network, *not in the end-user equipment*. The idea is that an authenticated user, a user with a SIM card (which is a mobile operator provided authentication and identification token) can access services in the network. End-users do not provide their own services in a mobile network. It is just not possible.

   This is different from internet technologies, where the idea is that any end-user, including a user that has only a home-computer, can contribute to layer 4-7 services with his or her own innovation. In practice, internet access providers restrict the capacity of residential users to run servers and provide services, but there is nothing in the technical architectures which requires such a restriction.



3. Mobile network technologies rely on *licensed spectrum frequencies*. Governments are able to collect substantial sums from auctioning spectrum to companies who then have to invest in a network capable of transmitting and receiving on the assigned frequencies. Spectrum auctions are a regulatory imposed *barrier to entry*. The way in which the spectrum licenses are structured by the government completely determines the conditions of the market.

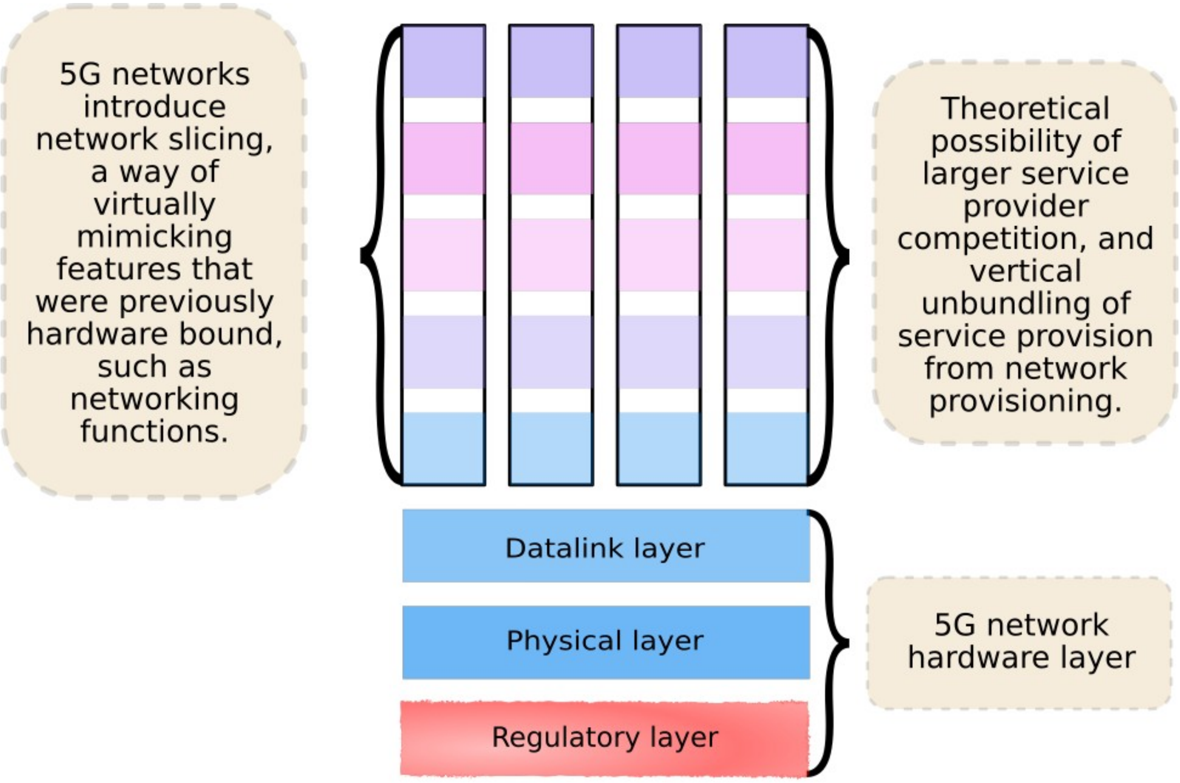
   a. If the spectrum license covers the entire territory of a state, then the only type of operator that can build a network permitted to broadcast is the type of operator that can simultaneously buy a spectrum license and build a nation-covering network. The territory of the spectrum license also determines the territory of the network. In practice it has proven challenging for operators to make mobile networking work across spectrum license borders (and mobile networks are technically more difficult to internetwork, since mobile networks rely strictly on the authentication of a user to an individual network connected to some spectrum territory).

This is different from internet technologies, since internet technologies do not, in general, rely on licensed spectrum (WLAN, for instance, operates in unlicensed bands), and also provide strong features for inter-networking networks. Internet networks can be built in smaller chunks or pieces (for instance, in a residential building, a university or a municipality) which are then connected by larger networks (which may be national backhaul, regional backhaul or transcontinental backhaul).[2] Because authentication and identification of users are strictly higher-layer functions, any issues of authentication and identification can be decoupled from the inter-networking of networks as such.

b. If the spectrum license does not cover the entire territory of a state, which is the case in India for mobile technologies in general and will be the case for 5G licenses in Germany and the UK, the architecture of mobile network technologies and authentication still makes it difficult to inter-connect networks. This technical limitation of mobile network technologies is at the heart of all discussions on *roaming*, which, broadly speaking, is the process of allowing an authenticated user from one network to use services on a network different to the one they are authenticated to.

Mobile network technologies have historically been less flexible than internet technologies, and they are more dependent on huge corporate entities. The discussions and concerns around 5G reflect this.

[2] See Annex: Glossary.


5G networks introduce network slicing, a way of virtually mimicking features that were previously hardware bound, such as networking functions.
Theoretical possibility of larger service provider competition, and vertical unbundling of service provision from network provisioning.
Datalink layer
Physical layer
Regulatory layer
5G network hardware layer

That said, mobile network technologies have gradually adapted to their inadaptability. The process to make internet traffic compatible with mobile networks started already with 3G technologies, but 4G and 5G have shifted even closer towards compatibility.

Network slicing is envisaged by network equipment vendors and network operators to enable *specialized services*, for instance providing a connection to a consumer which is *only* dedicated to television services, or only dedicated to high-quality gaming broadband, or only dedicated to connected vehicles, banking services, etc. In jurisdictions such as India or the EU, where net neutrality legislation is in place,[3] it is not clear that this would respect the guaranteed service quality promised to internet consumers by their governments.

Example G.1: An autonomous car relies on V2X (vehicle-to-anything) communication which requires low latency[4] but not necessarily a high throughput.[5] A video streaming service watched while the car is in motion requires a high throughput and is susceptible to latency. The same physical network is able to deliver both services on network slices that optimise its use.

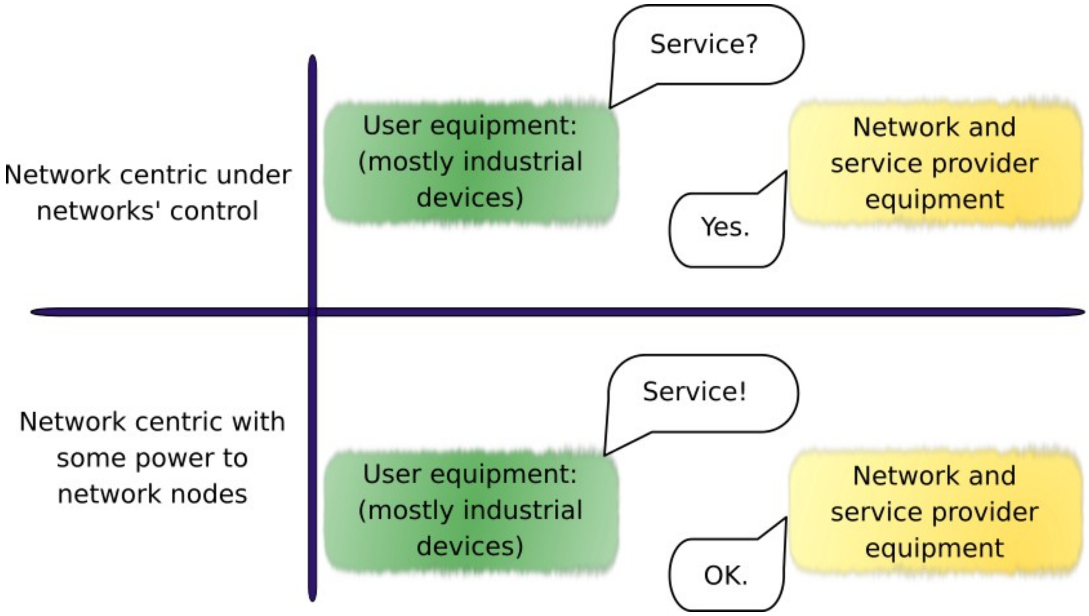
[3] For deeper reflections about net neutrality, see Using layers for policy analysis: Net neutrality.
[4] See Annex: Glossary.
[5] See Annex: Glossary.

Example G.2: Public safety personnel require guaranteed connectivity with high bandwidth when responding to an incident. Network slicing allows guaranteed secured bandwidth to be purchased, for example via a bandwidth-on-demand application as and when needed. In turn, on-demand functionality provides that the option to request priority has proliferated (e.g. fireman, security guards, hospital staff and others have this access).

However, network slicing could perhaps be used to vertically unbundle mobile networks (see picture), in a manner similar to what has been done in fixed networks. One entity would operate the spectrum license and the physical network, and a large number of entities could each operate network slices that all function as an internet connection. This would require network slicing technologies to support separate means of authentication to each slice, rather than having users authenticate to a centralized network and then being allocated a slice. Unfortunately, 5G technologies have not yet incorporated slice-specific authentication.

Mobile network standardization is still very *network centric*. This may imply that service innovation and product development in the slice is difficult. For instance, network slice operators could not do independent innovation on services or technologies. They would have to get approval for their enhancements from the network equipment vendors and network operators. Since the equipment vendors and network operators can be assumed to be very large entities which control network operations on a sizeable territory, these entities can already be assumed to be unwilling to make upgrades (which would be costly) - especially on behalf of only one network slice.

Requirements from industrial use-cases are pushing more control functions to user equipment, or at least the user equipment intended to be used in the particular slice or use-case. This is a break from the previous philosophy of mobile network technologies, which consists of user equipment asking for network permission to perform activities (authentication). Some proposed 5G features appear instead to make it possible for user equipment to demand that the network complies with user equipment demands (e.g. with respect to location data or quality of service).


Network centric under networks' control
Service?
User equipment: (mostly industrial devices)
Network and service provider equipment
Yes.
Network centric with some power to network nodes
Service!
User equipment: (mostly industrial devices)
Network and service provider equipment
OK.

Example G.3: In some industrial applications, like vehicle-to-anything communications, it is crucial that an application can have exactly the right through-put at exactly the right latency. The mobile equipment might want to be able to make specific demands on the mobile network, which the network cannot refuse. In this case, it is up to the network to provide the demanded functionality, rather than allocate to the equipment the resources it has available according to some best-effort principle.

These features are only envisaged for specialized environments, where mobile network equipment vendors are probably pressed to accept demands of specialized customers to expand their customer base. In the long-term, they could lead to a generally more permissive environment for all user-equipment, including consumer equipment, which at least mimics the freedom of consumer equipment in internet networks.

Electronic, or software-based, SIM[6] are another proposed consequence of industrial use-cases. SIM is usually the data provided to the application which authenticates the identity of a network user (note the distinction between data and application in this sentence!).

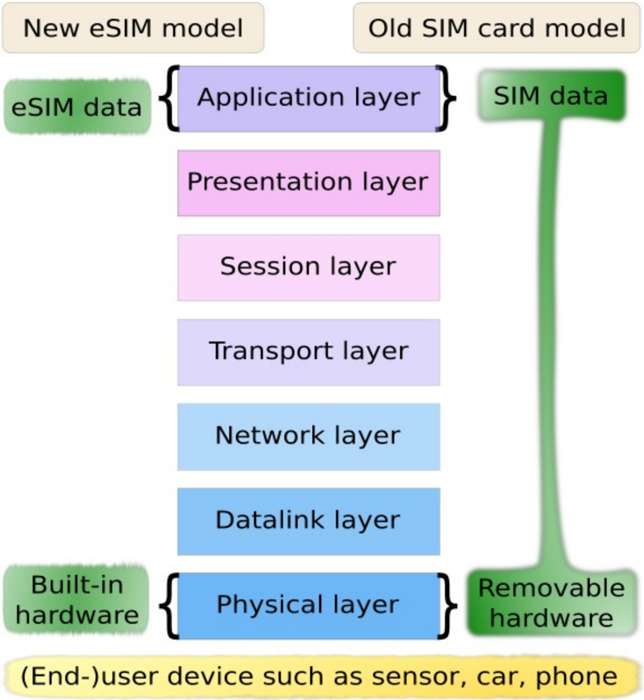
Most end-consumers will know that a mobile network connection is tied to a SIM card, a small, plastic card with a chip (layer 1) containing the data (layer 7) that allows the mobile network to authenticate whoever has the SIM card.

For small devices, like a sensor, this system is not very space efficient. This is increasingly true also for slimmed down consumer phones. Of course, SIM cards are getting smaller too. But an eSIM could rely on one permanently installed hardware module. Software-based SIM could rely entirely on the regular chipset.

For large quantities of devices, like a factory full of sensors, it is impractical to insert a SIM card into a small slot on each and every sensor. Having a built-in mechanism to store subscriber identity in each device makes deployment easier.

While the idea of software-based SIM is not new - it has been around for almost two decades - it is faced with two challenges. The first is that it is not obvious how to get authentication data to the device. The device would have no obvious way of authenticating to the network, since it is the SIM that contains the data used to perform the authentication. The second is that mobile network operators view the SIM card as a key way of managing their customers' identities. Software SIM would tie the subscriber's identity to the subscriber's device, rather than to the operator.

[6] Subscriber identity module.



Mobile network operators have historically been the largest buyers of mobile network equipment, so their preferences for identity management have been given priority by equipment vendors. With increased perceived utility for machine-to-machine communications in industrial environments this dynamic is changing.

Example G.4: One of the most popular reforms ever undertaken by the European Union is the reduction of roaming charges between EU member states through price regulation. If software SIMs had been available to end-consumers before the roaming regulation, it may have been possible for consumers to get temporary network access at local charges in the country of destination since the consumers' devices could have fitted multiple SIM. Consumers could then have avoided roaming charges by signing local contracts.

Much of the politics around 5G still centers on spectrum licensing.[7] This regulatory barrier to entry defines the market, who is active on the market, and what those entities feel empowered to do. Spectrum licensing is motivated by a scarcity argument: because there is only a limited amount of frequencies, and transmitting data on the same frequency can cause interference, some form of transmission management is assumed to be necessary.

Mobile network technologies solve this problem by having licensed spectrum. Motivated by the industrial use-cases, many governments are trying to create new spectrum licensing models. While not all of them are strictly 5G related, the 5G spectrum auctions are also bringing in new ideas - for instance the possibility of having local network operators.

[7] A comprehensive introduction to spectrum licensing can be found here: https://www.internetsociety.org/resources/2018/unleashing-community-networks-innovative-licensing-approaches/

Example G.5: The Ugandan,[8] Mozambican and South African Communications Commissions, among others, have worked on dynamic spectrum sharing[9] in 2018 and 2019. While 5G is not a target technology for spectrum sharing, other technologies are. Spectrum sharing opens up the possibility for having several, concurrent users of the same spectrum, with a primary licensee having priority in using it.

Example G.6: The Mexican regulator, the Federal Telecommunications Institute, amended in 2015 its frequency plan to set aside spectrum "*for social use*". To qualify for a social use license, applicants must serve small-scale networks or communities located in a priority zone. The first licenses were awarded in 2016. As a world-first, this allowed for the installation of a community mobile phone network in the states of Oaxaca, Chiapas, Veracruz, Guerrero and Puebla.[10]

Example G.7: Both the German regulator, Bundesnetzagentur, and the British regulator, Ofcom, have announced in 2018 and 2019 that they would try to issue geographically local licenses for 5G technologies. With a geographically local license, a large industrial plant owner could own and operate its own network, or communities could make their own local networks.[11]

Challenges of 5G technologies from a human rights perspective.

First, it may be helpful for a reader to keep in mind that the changes introduced by 5G are not immediately beneficial for end-consumers or human rights defenders. Many changes introduced by 5G are motivated by industrial use-cases but have the potential to serve end-consumers further down the line - especially as the technology becomes more user-centric and competition-friendly. But in the end, mobile technologies are very much defined by the regulatory landscape. The previous organisation of the mobile network markets implies that most countries possess strong mobile operators who in their capacity of being very large, and crucial for connectivity, wield a lot of power.

Nevertheless, mobile network technology development raises the following challenges.

Security and privacy > Law enforcement agencies have been heavily involved in mobile network standardization since the 1990s through special government working groups in ETSI[12] and 3GPP[13] to ensure that use-cases such as lawful intercept and mission critical systems are met. Requirements codified in technical specifications by law enforcement agencies can end up being crucial for a mobile network operator or mobile network equipment vendor who wants to have legal access to a market. Traditionally, this has raised human rights concerns as law enforcement agencies request real-time (warrantless) access to operator data streams[14] or request weaker encryption, which violates international standards on data protection and privacy..

[8] https://www.article19.org/resources/uganda-analysis-of-draft-tv-white-space-guidelines-2018/

[9] See Annex: Glossary.

[10] https://www.article19.org/resources/malaysia-submission-to-mcmc-ahead-of-wrc-19/

[11] Ibid.

[12] European Technical Standards Institute, a global technical standards organization that has traditionally dealt closely with mobile network technologies.

[13] 3rd Generation Partnership Project. The 3GPP was originally an off-shoot from ETSI, but is now a stand-alone cooperation between seven regional standards bodies that also involves the private sector parties interested in mobile network technologies.

[14] ETSI TS 102

The industrial and IoT use-cases are changing the equation for mobile network equipment providers. While for person-to-person communications, having weaker security to facilitate lawful intercept was sustainable, in industrial or machine-to-machine communications settings this is not the case. Therefore 5G technologies are proposed to include even stronger mutual authentication of the network and the device (i.e., the network must authenticate to the device to the same extent that the device must authenticate to the network), stronger encryption methods which make it more difficult to impersonate a network (known from the media as Stingray operations or IMSI-catchers), or end-to-end encrypted communications to ensure that data streams are not being tampered with.

One of 5G's main functions will be to make possible the generation, storage and sharing of vast tonnes of data on individuals, objects, devices and the environment through the IoT. These data flow and sharing have to be performed and managed in accordance with data protection and privacy standards.

Law enforcement agencies have expressed strong concerns that 5G technologies could render traditional wiretapping methods far more complicated or even redundant.

Proposals for dealing with the situation range from trying to influence the international bodies responsible for establishing the relevant technical standards; passing new laws (at both national and regional level) to enforce police demands, and ensuring a broader discussion amongst major surveillance powers such as the USA, Australia, Canada and the EU.
The risk is that, as 5G technologies are more centralised than 3G and 4G ones, the establishment of the wrong standard or the adoption of the wrong law could expose users to a higher degree of control and intrusion by the provider.

The incorporation of privacy-by-design and human rights due diligence practices into the development and deployment of 5G infrastructure, network services, and IoT devices, can help to avoid or mitigate the potential negative impacts of 5G technologies.

Digital divide > As a next-generation network, 5G is a high-end technology and cannot be assumed to remedy the digital divide. Investments in new, expensive technologies happen first in densely populated and profitable areas, such as cities or travel hubs. This is especially the case for technologies that rely on licensed spectrum, where the upfront costs for even being allowed to start building a network are high.

With governments becoming increasingly experimental around spectrum licensing, there is hope that 5G networks can be more quickly and cheaply deployed by more actors. However, for rural areas, low frequency spectrum bands, in the 400-800 MHz ranges, stand the best chance of covering large areas without harmful interference. In this frequency ranges, spectrum planning needs to be more precise as harmful interference could occur if multiple mobile networks attempt to broadcast in the same bands. It is likely that other technologies, such as dynamic spectrum sharing technologies, will be more conducive for getting connectivity to these areas.

Accountability > 5G technologies may lead to a significant increase in the number of entities managing or operating pieces of the network, providing digitally-enabled services, producing digitally-enabled products, and storing and/or processing data. How liability is allocated will affect the effective protection of end users rights, and will also influence how companies are allowed to design, build and deploy technologies.

Centralisation > The entire point of virtualizing higher layers in the network is to provide more economic (and technical) power to those parts of the networks that are not virtualized. This lends itself to considerable centralization. In the explanations above, we wrote that network slicing *could* be used to vertically unbundle mobile networks, similar to what has been done in fixed networks. However, this is not likely to happen in practice and in either case it is not likely to happen for end-consumer oriented internet services. The mobile network operators, who will own and operate the network, own the spectrum licenses and provide end-consumer access services, imagine themselves as central pillars of all digital service provision - guaranteeing everything from user identity to specialized connections to controlling which competitors can use their networks. With a risk of lower diversity among service providers, there is also a growing risk of regulatory capture, human rights infringements which for economic reasons are more difficult to address, and higher market entry barriers for companies that may want to experiment with more human rights friendly business models.

Net neutrality and non-discrimination > The concept of 5G network slices is sometimes presented to the public as a functionality with built-in network discrimination. This may very well be incompatible with existing current rules on net neutrality. In theory, making a commercial offer which contradicts existing rules should not be a viable option for mobile network (slice) operators, but human rights defenders may have to be diligent in their monitoring of supervisory authorities.

Human-centrism > Some of the proposed changes to identity management in a mobile network could have great long-term impacts on users empowerment and their rights and freedoms. Shifting control over identities from the network to the user's own device puts more control in the hands of the user. In technical terms, one could argue that this shifts the power to manage user identities from mobile network operators and mobile network equipment vendors to smartphone manufacturers and mobile operating system vendors. It is a decentralization of identity management - within some limits. For instance, a mobile operating system vendor does not for technical reasons have to exercise any control over how the device-user authenticates to a network, but in practise it will have to provide all the interfaces for the user that enables such authentication.

Of course, a relatively larger local power over identity may also disrupt many of the control systems for which operator-provided identity play a large part today: one-time-passwords, two-factor authentication and password recovery features often rely on the mobile operator having a strong control over user identities. Another side-effect could be larger mobility of consumers between mobile network (slice) operators. This would create serious disruption to many mobile network operators business models.

Self-evaluation questions

1. Who owns mobile network spectrum licenses in your jurisdiction?
2. Are there mobile virtual network operators in your jurisdiction?
3. Could a consequence of increasingly “internet” properties of mobile networks be that end-consumers are able to run their own servers over mobile broadband connections?
4. What is the major difference to the OSI model in the case of 5G?
5. Imagine you are a mobile network operator. What is the advantage of:
    a. A network-centric architecture?
    b. Authentication?

Proposed answers

1. This information is likely available with your local telecommunications market regulator. The number of licenses and their proprietors may not correspond with the mobile rental providers that you are familiar with as a consumer.
2. Similar to answer in 1.
3. In principle yes, in practice it will depend on mobile network operators’ terms of service.
4. It makes it possible to do greater vertical unbundling. It also disconnects consumer billing identities from operator-controlled hardware. It integrates more internet-like features into the mobile phone network, while maintaining the strong, centralized tradition of mobile technologies.
5. Some examples:
    a. Ability to make stronger security guarantees, being less reliant on third-parties for quality of service guarantees, etc.
    b. Being able to charge the right customer for the right amount of service consumption, customizing offers per consumer, etc.