5G-Infrastructure-Association
Gaston Crommenlaan 1 PO box: 101
9050 Gent
Belgium
www.5g-ppp.eu

Vision Document for B5G Research

Global Editor: Rui L. Aguiar, Instituto de Telecomunicações

Editors	
Rui Luis Aguiar	Instituto de Telecomunicações/Universidade de Aveiro
Jean Sebanstien- Bedo	Orange
Ana Garcia Armada	Universidad Carlos III de Madrid
Barry Evans	University of Surrey
Alex Galis	University College of London
Holger Karl	University of Paderborn
Contributions by	
Adam Wolisz	TU Berlin
Alister Burr	University of York
Alvaro Palomo	A-CING
André Bourdoux	IMEC
Antonio F. Skarmeta	University of Murcia - Spain
Bengt Ahlgren	Swedish Institute of Computer Science
Carla Fabiana Chiasserini	Politecnico di Torino;
Dirk Kutscher	NEC Eurolabs
Dirk Trossen	Interdigital
Donald Parruca	Paderborn University
Eduard Grasa Grasa	i2cat
Felipe Gil-Castiñeira	University of Vigo - Spain
Hamed Al-Raweshidy	Brunel University
Isabelle Tardy	SINTEF
Isidro Cabello Medina	Asi Soft
Jaap van de Beek	Luleå University of Technology – Sweeden
James Gross	KTH Royal Institute of Technology
Julian Sessena	A-CING
Luis Nero Alves	Instituto de Telecomunicações, Universidade de Aveiro
Mats Eriksson	Arctos Labs Scandinavia AB - Sweeden
Matthias Herlich	Salzburg Research Forschungsgesellschaft
Matti Latva-aho,	University of Oulu
Norbert Niebert	Ericsson
Patrick Svedman	KTH Royal Institute of Technology
Paulo Mendes	Universidade Lusófona
Pedro Martinez-Julia	University of Murcia - Spain
Roberto Llorente Saez	Univ. Politècnica de València
Rute Sofia	Universidade Lusófona
Stefan Valentin	Huawei
Torsten Braun	University Bern

Executive Summary

European R&D Community is considering the challenges that will be appearing on what can be called “beyond-5G” networks, in preparation of research for the period after 2020. The first step for this analysis was taken in a workshop organized by the 5GPPP Infrastructure Association and the Networks 2020 ETP. The workshop had a large audience, and highlighted that the community is still not clear on what will be the expected progresses deployed in the 5G networks. This document, originating as an outcome of the 5GPPP workshop, is the result of an analysis of a wide community involved in a call for contributions to present a snapshot of the European community view on “research for beyond 5G”. More than 30 experts from academia and industry contributed with ideas and concepts.

The documents considers both research in the scope of evolutions of the 5G systems (for the period around 2025) and some alternative/longer term views (with later outcomes, or leading to substantial different design choices). This document reflects on four main system areas: fundamental theory and technology, radio and spectrum management; system design; and alternative concepts.

The result of this exercise can be broken in two different strands: one focused in the evolution of technologies that are already ongoing development for 5G systems, but that will remain research areas in the future (with a “more challenging” requirements and specifications); the other, highlighting technologies that not really considered for deployment today, or that will be essential for addressing problems that are currently non-existing, but will become apparent when 5G systems begin their widespread deployment.

Overall conclusions surround the increasing relevance and impact of notions of native networking “intelligence” and “flexibility” backing up higher performance and capacity in the networks.

We expect that the simpler notion of communication network to become gradually replaced by the notion of communication system, with the added value perceived by such a change. Even at physical layers, links that operate with a high level of intelligence on its surroundings are expected to appear, bringing realizability aspects that need to be addressed. The added trend towards software defined systems, and the added control intelligence expected in these systems, brings effective complexity management into play at multiple levels, with multiple potential techniques able to trade-off complexity with response time being identified in this document. A different line of development in the document is the connection with the human, with several aspects of interaction between humans (as beings) and the evolving system. It is thus not surprising that body communications and human-networked system interface.

Some of the areas here identified were perceived as being hard to consider in existing programme approaches. Interestingly, researchers highlighted that several of the technologies here considered had potential deep societal impacts (if successful) or were hard to evaluate considering traditional evaluation methodologies in research programs dedicated to mobile telecommunications. It is recommended that all these aspects, from the technological areas to the relevant of non-standard metrics, are considered when discussing funding associated with research lines for beyond 5G research.

Contents

1 Introduction

Under the context of preparation of research for the period after 2020, Europe is considering the challenges that will be appearing on what can be called "beyond-5G" networks. The first step for this discussion was taken on the 29th October, in Paris, in a workshop organized by the 5GPPP Infrastructure Association and the Networks 2020 ETP. The workshop had a large audience, and highlighted that the community is still not clear on what will be the expected progresses deployed in the 5G networks, muddling the definition of a follow-up set of research challenges.

The timelines that seem to be now accepted around the 5G discussion are as follows:

- 5G phase 1—technically introduced 2018/20 using spectrum allocated in WRC 2015 below 6GHz. This is what is referred in this document as "5G".
- 5G phase 2---technically introduced 2025 and maybe lasting to 2030 adding spectrum allocated in WRC 2019 above 6GHz.
- Post 5G phase 2—after 2030 with what might be lead to 6G.

The idea of this document is to input to an EU research programme B5G. This could be 5G phase2 or post 5G phase. The vision that we currently have in this document is **mostly** targeting 5G phase2: industry, in particular, views B5G as 5G phase2. Academia is not so uniform in its view, and depends on the area: some fields, specially new system architectures, may probably be for Post 5G phase 2 (and then again, if these system views reach standardization before, then they will become 5G evolutions). As the three phases above are unlikely to be revolutions and more likely to be evolutionary, in this document B5G is considered as a mixture of 5G phase 2 and post 5G, but with an emphasis on the first.

Thus, this document aims to contribute to clarify potential research areas relevant for B5G networks, in this context, and identify areas that need to have basic research developed in the next few years.

Section 2 makes a brief overview of what 5G (phase 1) scenarios should be. Section 3 addresses fundamental techniques for Tb/s communications, discussing challenges associated with information theory and coding, antenna processing, and in general advanced signal processing to optimize and reach Tbit/s in wireless communications. Section 4 handles Spectrum and Radio management and advances for novel use of the spectrum, including innovative usage scenarios. Section 5 discusses System Design, heavily influenced by the needed evolution of the software network for mobile environments. Section 6 presents Alternative Technologies and Designs, new (or not in standard-track) concepts, including opportunistic networks, low-latency delay, and efficient mobility for high-bitrate and zero-latency services. Finally section 7 presents the major conclusions of this community view.

Across the whole document, the topics are presented in an order to facilitate readability, and in no way the order of topics, sections of technologies, should be perceived as indicating priority or importance.

2 5G Scenarios

This introductory section summarizes expected features of 5G networks as to be deployed by 2020/5. Although work is still on-going, and a large amount of imprecision still exists, there are also a large amount of concepts that are considered to become reality in 5G phase 1 networks. In this section we discuss, in a top-down approach, the 5G views from service, system and component perspectives.

2.1 Service perspective on 5G

Firstly, 5G will ensure user experience continuity in challenging situations such as high mobility (e.g. in trains), very dense or sparsely populated areas, and journeys covered by heterogeneous technologies. There are a couple of user scenarios that can be used to illustrate 5G service features.

Great service in a crowd

Today many users do not expect to have a good user experience of mobile and wireless internet access service when being surrounded by crowds of people. Furthermore, the network load will increase due to increased penetration of high-end devices and challenging services such as mobile computing, this might further degrade the user experience provided by the legacy network. In the 5G network, users will expect good service even in very crowded places, despite the increase in traffic volume.

Best experience follows you

This scenario strives at bringing a similar user experience for end-users on the move as for static users e.g. human users at home or in the office. No matter where or how one is moving e.g. walking in a city, travelling on a train or subway or in a car on the highway, the beyond 2020 end-user is provided a communication that works reliably and provides a high user experience as if the best experience were following the user.

In addition, 5G will be a key enabler for the Internet of Things by providing a platform to connect a massive number of sensors, rendering devices and actuators with stringent energy and transmission constraints.

Ubiquitous things communicating

In a fully connected society, where today’s human centric communication is complemented with machine-type communication, basically anything that profits from being connected will be connected. A majority of the connected machine-type devices will most likely be simple, such as sensors and actuators, for which the main requirements are low energy consumption and low cost. This scenario addresses the communication needs of a massive deployment of ubiquitous machine-type devices. The resulting, widely varying, requirements in several domains e.g. in terms of energy consumption, cost (complexity), transmission power, latency, cannot always be best met by today’s cellular networks.

Furthermore, mission critical services requiring very high reliability, global coverage and/or very low latency, which are up to now handled by specific networks, typically public safety, will become natively supported by the 5G infrastructure.

Super real-time and reliable connections

The reliability and latency in today's communication systems have been designed with the human user in mind. For future wireless systems we envisage the design of new applications based on M2M communication with real-time constraints, enabling new functionalities for traffic safety, traffic efficiency, smart grid, e-health or efficient industrial communications. Such new applications may require much higher reliability and lower latency than today's communication systems: for certain use cases, a maximum E2E latency must be guaranteed with very high reliability, e.g. 99.999%.

These scenarios provide guidance on the performance and functional behavior of 5G, but are of course very technology agnostic.

2.2 System perspective on 5G

in order to meet the expected business and performance requirements, especially in terms of latency and reliability, and to support new business models and scenarios, the various 5G subsystems and interfaces need to be inspired by modern operating system architectures, and software concepts.

In particular, 5G will provide a generic approach to make network platforms much more flexible and universal, in order to cope with heterogeneous environments and requirements. In this context, the concept of Network Slicing will be particularly relevant. Network Slicing is a concept for running multiple logical networks as virtually independent business operations on a common physical infrastructure. The target is to provide network customer specific functionality, without losing the economies of scale of a common infrastructure.

Beyond that, 5G will implement the convergence between fixed and mobile networking services with the associated evolution of core and transport networks. 5G will integrate networking, computing and storage resources into one programmable and unified infrastructure, which can be customized according to the interests of multiple costumers ("vertical industries").

Energy consumption will also be dramatically reduced in the terminal and infrastructure through optimization of the air interface and signaling (to focus power when and where it matters) as well as sleep modes. A reliable power model will also be developed to explore the individual components relevant for power consumption analysis of future architectures. However, we envision that *harvesting energy systems will not be included in the first release of 5G*.

While the diversity of services and the complexity of the infrastructure will apparently increase, 5G is expected to radically cut total cost of ownership (TCO) of the infrastructure, on one hand, and the service creation and deployment times, on the other. Identifying network traffic patterns proactively is necessary to avoid situations that threaten the performance and availability of the network. This can be done by integrating Self-organizing network (SON) features, so as to reconfigure resources to achieve high QoS and also, solve load balancing optimization problems. Hence, service/network management will evolve accordingly with advanced automation including cognitive operations and exploiting Big Data for better QoS and QoE, whatever the prosumer will be (human, machine or thing).

For current mobile systems, the trust model is rather straightforward, involving a subscriber (and their terminal) and two operators (the home and serving networks). Since 5G is aimed at supporting new business models and involves new actors, trust models will change, giving rise to extended requirements in areas such as authentication between various actors, accountability and non-repudiation.

The use of clouds and virtualization emphasizes the dependency on secure software, and leads to other effects on security. Decoupling software and hardware means that telecom software can no longer rely on the specific security attributes of a dedicated telecom hardware platform. When operators host third-party applications in their telecom clouds, executing on the same hardware as native telecom services, there are increased demands on strong isolation properties.

2.3 Key technological components of 5G

5G wireless access will support a heterogeneous set of integrated air interfaces: from evolutions of current access schemes to brand new technologies. Seamless handover between heterogeneous wireless access technologies will be a native feature of 5G, as well as use of simultaneous radio access technologies to increase reliability and availability. The main concepts for RAN architecture with regards to 5G will be Ultra Dense Networks (UDN), Device to Device (D2D) and Moving Networks (MN). They will enable a better experience continuity on the move, under low coverage and in crowded areas. 5G RAN will also require new waveforms; agile access techniques; advanced multi-antenna beam-forming and beam-tracking and MIMO techniques; new radio resource management algorithms, to name just a few.

5G will leverage on the strengths of both optical and wireless technologies. To achieve the expected capacity, coverage, reliability, latency and improvements in energy consumption, the 5G architecture is expected to i) run over a converged optical-wireless-satellite infrastructure for network access, backhauling and fronthauling; ii) leverage flexible intra-system spectrum usage; iii) make optimal utilization of the specific strengths of the different underlying infrastructures (e.g. leverage multicast for satellite or flexible spectrum for optical).

Network functions are expected to run over a unified operating system in a number of points of presence, especially at the edge of the network for meeting performance targets. As a result, it will heavily rely on emerging technologies such as Software Defined Networking (SDN), Network Functions Virtualization (NFV), Mobile Edge Computing (MEC) and Fog Computing (FC) to achieve the required performance, scalability and agility.

Any research on B5G must consider that the starting point at the time (ci.2020) will roughly reflect the descriptions above. We can always consider later research and development paths which are more or less evolutionary, or that will become relevant for this new communication network.

Section 6.1 will add some more alternative views to how we expect scenarios to change after 2020.

3 Fundamental techniques for Tb/s communications

Evolving wireless networks must be capable of efficiently supporting an increase in peak and average data rates far beyond what is today possible. This section addresses some of the challenges associated with information theory and advanced signal processing, to optimize and reach Tbit/s in wireless communications, including implementation considerations and energy constraints.

It is recognized that some of these topics may undergo evolutions for deployment in 5G networks. However, regardless of the dimension of 5G developments, they will still remain a research issue for further evolutions.

Huge capacity increase could be achieved by having more spectrum, better spectrum efficiency and having a large number of small cells. Although details on spectrum will be discussed in a later section, these concepts set a totally new design paradigm to current macrocellular dominated system design and radically new concepts need to be developed at fundamental level to realize them. We should explore all aforementioned three dimensions, namely increased bandwidth at higher frequencies, improved spectral efficiency and increased number of connected small cells, the last one being a key driving force. Nonsurprisingly, most of these seem evolution of ongoing research techniques.

3.1 Enabling techniques and technologies for higher carrier frequencies

It is an accepted fact within the 5G research and development community that mm-waves communications will constitute an important part of the future heterogeneous network that will form 5G (and beyond 5G) systems. This is consequently a popular field of research worldwide for outdoor communications and certainly indoor ones.

In order to facilitate the development of 5G technology in the mm-wave frequency bands, beamforming possibilities, and spatial multiplexing and space diversity need all be addressed jointly. While the A/D converters and antenna designs will allow some smart combinations, the impact on the costs on user terminals and small base stations will limit practical implementations.

Additionally, stringent demands in terms of energy usage will further demand quicker network discovery or handovers, while leaving most time for data communications. Combining location knowledge to communication need will greatly speed up the process, providing that location information can be retrieved both outdoor and indoor. This is a research topic of great interest for the years to come. While the physical issues have been partly addressed in 5G, the combination of physical and MAC issues need more research in a B5G programme in order to strike a balance between performance, complexity and power of mm-wave communications.

Yet, Tb/s communications will require extremely wide bandwidths that are available at carrier frequencies of 300GHz or higher. At these frequencies, bandwidths of 50GHz or more will probably be available. Exploiting this portion of the spectrum will require multi-disciplinary research effort in several domains:

- semiconductor technology for the RF, analog baseband and digital logic;
- new modulation, demodulation, synchronization methods targeting unprecedented efficiencies of 1pJ/bit;
- agile antenna arrays with tens or hundreds of elements.

It is anticipated that breakthroughs will be needed to achieve these goals. Significant efforts will also be needed in areas such as propagation measurements and channel modeling.

Given the challenges mentioned, it is likely that a phased approach will be needed whereby intermediate solutions featuring bit rates of 100 to 300 Gb/s will be developed and standardized. This fits perfectly on the vision described in the Introduction of a phased evolution for 5G systems.

Yet another aspect will be regulation. At the moment, the regulatory situation at frequencies above 300GHz is unclear and not harmonized on a worldwide basis. Significant efforts will be needed at ITU level and in WRC conferences to reach a suitable consensus. (See next section)

3.2 Realizable Massive MIMO

Massive multiple input – multiple output (MIMO) techniques have emerged as a solution to both capacity and energy efficiency demands. It is well known that the benefits of increasing substantially the number of antennas beyond what is in use today are, among others [Ngo 2013, Marzetta 2010]:

- To improve the data rates and reliability (through multiplexing and diversity gains)
- To decrease the required transmit power.

And this is achieved with very simple pre-coders and decoders. However, to obtain all these benefits the knowledge of channel state information (CSI) at the transmitter and receiver is required. Even assuming a TDD scheme and the adequate calibration, the problem of pilot contamination, that is, uplink pilot sequences from different cells interfering with each other, make obtaining even imperfect CSI very challenging for massive MIMO where the number of channels that need to be estimated is very large.

Therefore, simplified solutions are required where the number of channels is not so high, or CSI is not required.

Antenna Selection [Gao 2013] and Spatial Modulation [Renzo 2011] are two possibilities that, when combined with massive MIMO, allow the reduction of the number of radio-frequency (RF) chains associated with the antennas, with a positive impact on the system complexity and hardware cost, while preserving the system performance at a certain required level.

Non-coherent schemes where CSI is not needed at either the transmitter and receiver side become very interesting [Chowdhury 2014], [Armada 2015]. Current research on non-coherent massive MIMO shows that the same scaling laws in terms of achievable rates and energy efficiency can be obtained as with coherent systems with full CSI. Also the performance in terms of bit error rate is very close, when realistically obtaining CSI is assumed in the coherent counterpart. Today's preliminary results support the idea that it is time for non-coherent communications and their application to massive MIMO may be the key to really implementing these systems. Moreover, dispensing with CSI has some interesting side effects on lowering the signalling overhead and latency in the air interface because of not requiring sending pilots and channel feed-back.

Recent studies show that massive MIMO can be very advantageous both in classical sub-6 GHz frequencies and above 6 GHz where more spectrum is available [Heath 2014]. The application of all above to millimetre-wave frequencies is more challenging, while the potential benefits are also more promising. In this case, ways to combine them with the beamforming required to counteract the high propagation loss must be designed.

3.3 D2D networking and offloading

An enabling technique for efficient wireless transmission is Device-to-Device (D2D) communication. D2D is currently identified as a high-potential technological paradigm where devices in proximity autonomously establish communication links. Exploiting the natural proximity of communicating devices may provide multiple performance benefits. First, D2D *User Equipment (UE*) may enjoy high data rates and low end-to-end delay due to short-range direct communication. Second, communication becomes more efficient if traffic is kept locally, by ensuring that devices in the range of each other communicate directly, instead of having to go over a controller such as the *evolved Node B (eNB)*. In particular, compared to normal downlink/uplink cellular communication, direct communication saves energy and improves radio resource utilization, reaching high bitrates with low power. Third, switching from an infrastructure path to a direct path offloads cellular traffic, alleviating congestion, and thus benefiting other non-D2D UE as well. And, operators are becoming increasingly interested in context-aware applications where a plurality of services can be provided depending on the users location.

An interesting and promising twist on D2D is to reconsider the control structure. So far, D2D work draws a lot of its motivation and ideas from self-organization work. It is interesting to consider how an operator might actually assist in such D2D decisions and orchestrate those actions. This feature opens up a range of opportunities for improving spectrum efficiency and achieving higher data rates, at virtually no cost for the operator. However, it also poses several fundamental challenges:

- Which D2D links should be established, taking into account current content requests?
- How to implement D2D communication: direct links or via a clustering scheme?
- Which radio resources should be allocated, taking interference management into account?
- How to set D2D transmit power levels?
- Which edge node (if any) should control D2D communication, taking the current traffic pattern into account?

Research should be devoted to the study of new strategies that heavily involve D2D communication between users, evolving from the current ideas of D2D development into massive usage of short links.

3.4 Wireless and optical fronthaul and backhaul

In small cell systems, link distances are quite small enabling the increase of center frequencies. Improving the spectral efficiency in small cell systems call for drastically new design principles. As the number of base stations is drastically increasing, and more and more network building is unplanned in nature, there are major challenges in the future for organizing backhaul links to core network. It is clear, that Ethernet or optical fiber connection cannot be always organized. Thus, wireless backhaul links must be used in many cases. In vehicular environment (trains, buses, cars) only wireless solutions are possible.

For wireless backhaul links, two alternatives exist: in-band and out-band transmission. In-band backhauling uses the same frequency band for transmission and reception, i.e., it is a form of relaying. Out-band backhaul links utilize different frequency band for backhaul transmission and reception from mobile terminals. When we talk about Tbps communications, backhaul traffic becomes even more of a challenge than it is today. However, as in many cases wireless backhaul has very slowly changing channel statistics, more complicated waveforms can be deployed as well as stronger channel coding could be used. Of course the requirements for backhaul depend heavily on traffic requirements at fronthaul which cannot be sacrificed due to bad design of backhaul.

On the other hand, optical fronthaul/backhaul for next-generation wireless systems is also challenging due to the specific nature of the signals transmitted. Fronthaul connectivity, nowadays implemented in digitized Radio-over-Fibre (digitized-RoF) [Tanaka 2015], as in ACPI industrial standard [cpri v6.1], are expected to cope with ultra-wideband as in current definition of 5G wireless. Considering the bandwidth and MIMO requirements in Beyond-5G systems, where a massive number of antennas should be connected, a direct Radio-over-Fibre fronthaul has been indicated as a viable solution [Morant 2015].

Radio-over-Fibre transmission on new specialty Fibre, as multicore optical media, opens up the possibility of transmitting different MIMO signals for the different antennas on different cores with the advantages of simultaneously downstream/upstream at the same wavelength, and no need for electrical or optical up/down-conversion stages [Llorente 2015]. As the number of MIMO signals increases, multicore fibre permits to avoid sub-carrier multiplexing (SCM) or dense wavelength-division (DWDM) schemes that could imply chromatic dispersion side effects, rendering the MIMO digital processing algorithms inefficient

Beyond-5 cellular systems require optical backhaul with multi-Terabit/s capacity and specific characteristics of scalability and adaptability to the network traffic patterns. Optical backhaul technology can fulfil these requirements. Conventional modulation schemes on Standard Single-Mode Fibre exhibit a transmission limit of 100 Tbit/s due to non-linearity and power budget limitations. This limit has been reached employing digital coherent receivers [Morioka 2009].

Spatial-division multiplexing (SDM) optical transmission, sometimes further including modal multiplexing, has become an emerging transmission technology capable of dealing Tb/s backhaul transport requirements in B5G. SDM optical transmission technology has demonstrated bitrates exceeding 2 Pbit/s in September 2015, [Puttnam 2015, Soma 2015]. These results indicate SDM optical transmission as a good candidate for Beyond-5G backhaul systems.

4 Spectrum and Radio management

This section addresses visions of spectrum and radio management for next generation systems, considering both near-term 5G (phase 1) and far-term 5G systems (phase 2)

4.1 Spectrum Management

Fixed spectrum access as a property of an operator is the current system. Operators would argue that long term access to this is needed to make their business cases. Regulators to some extent are complicit in this due to economic reasons. **Thus the system will be difficult to change and will evolve gradually**. The fact is that the demand for spectrum for 5G cannot be satisfied by adherence to the status quo.

In 5G phase 1, below 6GHz there will be use of spectrum aggregation across several bands and hopefully more sharing on a geographic and temporal fixed basis. There is room here, for example to geographically reuse on an area basis (e.g. coastal areas to inland rural), and on a temporal basis for overnight use, e.g. airports. There is also mileage in sharing between MS and others (such as FS and FSS) where current technology is adequate to guarantee non interference. Data bases and spectrum sensing can play a part here but both have their limitations.

For 5G phase 2 the use of millimetre waves for small cell high capacity systems will be introduced. The study of bands above 6GHz will be initiated by WRC 2015 and report to WRC 2019 for allocations. Already there are ongoing propagation and channel studies in some of these bands. To secure the required 5G bandwidths of 500MHz to 1GHz, bands such as 10GHz, 25GHz, 28-31GHz, 40-52GHz, 66GHz and 70-80GHz have been suggested. In some of these there are existing systems that may make co existence challenging and equipment costs and availabilities can also be constraints. The studies will be completed by 2019 and although essential as they form part of 5Gphase2 could be argued not to be part of B5G.

Returning to the essential problem which is more flexible spectrum and the move away from fixed licence access. The move here is away from spectrum ownership to one where spectrum is free to all but access to the spectrum is regulated. The idea of Leased Spectrum Access (LSA) promoted by European RSPG, has already been worked on and researched. The idea being to share resources on an agreed basis for the same block of spectrum between operators. It can be foreseen that simpler versions of this scheme between say two operators in a particular area could be adopted in 5Gphase2. The allocation of spectrum holdings, as blocks to groups of operators facilitating dynamic access would be the extension and be particularly useful in rural areas. Such schemes would be applicable in 5G phase 2. They are technologically feasible today but the management schemes require more research.

Another consideration is the use of unlicensed bands.(Note that it is not the bands that are unlicensed but the terminals used therein) Indeed **the longer term view might be to consider all licensed and unlicensed spectrum as one**. Unlicensed bands could be used early in 5G phase1 in hot spots to offload traffic when congestion occurs. The spectrum may be used for instance by mobile operators for offloading of non-delay-sensitive traffic, for residential home networks, for public wireless access in commercial venues, for wireless access in corporate networks, etc.

In both licence-shared and unlicensed spectrum access, it is important to avoid colliding transmissions, which might block the access altogether. In licence-shared access, this is achieved by letting the non-incumbent entity use the spectrum only when the incumbent entity will not use it, within some coarse time-period and area. In unlicensed spectrum access, there are several methods to avoid collisions, such as listen-before-talk and dynamic channel selection.

Potentially, a large number of nearby access points (APs) and devices are passive and silent during use of a licence-shared and unlicensed spectrum. Many of those nodes are owned and controlled by other entities, such as another operator or private persons, and are called **alien nodes** here. By inviting the nearby alien nodes to play a more active role in the communication, **the scenario can be dramatically changed, from a sparse network to an ultra dense network, without the need to deploy more nodes**. The participating nodes should typically use the same RAT, but cooperation over multiple RATs that may be used in the band could also be considered.

Nearby alien nodes could be used for relaying to **extend the coverage** in the band, which might be particularly useful on unlicensed bands with low maximum transmit power or on millimetre wave bands with high propagation loss. The alien nodes could also be used for CoMP to **improve throughput, reliability etc**. In high load situations on an unlicenced band, dynamic coordination of the communication of the different entities could be achieved, for instance synchronized coordinated scheduling. This could **drastically increase the throughput for all entities**, compared to if only coarse mechanisms, such as listen-before-talk, are used. Since the alien nodes are owned by other entities, the cooperation needs to be **voluntary**. However, helping another entity can be rational, even if it improves your own QoS only at a later time.

Many aspects of this topic have been studied in the contexts of for instance

- self-organizing networks (SON),
- ad-hoc networks and
- device-to-device (D2D) communications.

Still, a number of challenges remain before the power of the alien nodes can be unleashed. Firstly, security must be bullet-proof and two-sided; both for the communicating node and for the assisting alien node. An attractive but voluntary model for cooperation needs to be defined. The model could include multiple levels and modes of assistance or cooperation and also some measures for fairness. Standardized, robust and efficient protocols and interfaces for negotiation, cooperation setup, cooperation signalling, data forwarding, cooperation termination etc. are needed. Both wireless and wired interfaces could be considered for the communication with an alien node, perhaps based on the 5G interfaces for inter-node or D2D communication. Perhaps also the alien node behavior needs to be standardized to a greater extent than that is typical for network nodes, which can be challenging for more advanced forms of cooperation. Furthermore, the required air interface mechanisms, such as alien node discovery and synchronization, may need further consideration.
If operator networks will not be able to orchestrate and control alien nodes like the operator-owned nodes, this might have an impact on the system design. New functions, protocols and interfaces might be needed.

The spectrum sharing architectures include the functionality to handle all types of spectrum authorizations. This includes support for the classical licensing methods like dedicated spectrum as well as the new licensing methods like Licensed Shared Access LSA or co-primary sharing. Additionally, license exempt sharing and new sharing models using spectrum pools or mutual renting are covered. The spectrum sharing architecture consists of two layers, which covers technical (network view) as well as regulatory aspects (Administration view). While the technical layer provides all functions to use spectrum resources in a network, the administration layer deals mainly with functions to plan the spectrum use, allocate and assign spectrum licenses and to enforce license conditions [ICT-R]. The METIS project has developed spectrum management on administration level which could be a model for future research. The splitting of the control and signaling from the data plane in 5G architectures opens up more possibilities for adoption of shared spectrum usage.

4.2 Radio Management

The move towards an architecture that splits the control and data planes and uses a macro overlay for the control plane opens up real possibilities of improvements in mobility management and hand over for small cells as well as energy efficiency. The production of an integrated MAC that adapts to everything and is independent of the radio technology is a desirable feature. Terminals will need to be dual chain for the C plus U division and parallel processing will reduce latency. IoT and high rate terminals will be developed separately.

New breakthroughs in the integrated access node and heterogeneous convergence are required to enable the ultra dense radio nodes to work efficiently. The plug-and-play function becomes essential to commercial deployments, in which the available spectral resources should be allocated and the corresponding parameters should be self-organized. Furthermore, the software-defined air interface technologies should be seamlessly integrated into the 5G radio access network (RAN) architectures. The cloud computing based radio access infrastructures would provide on-demand resource processing, delay-aware storage, and high network capacity wherever needed. There are several proposed solutions.

Some advanced technologies, such as the cloud radio access network **(C-RAN**) and ultra small cells based heterogeneous network (HetNet), have been presented as potential 5G solutions. C-RAN has attracted intense research interests from both academia and industry (such as China Mobile, Huawei, Alcatel Lucent, Qualcomm).

In C-RANs, a large number of low-cost remote radio heads (RRHs) are randomly deployed and connect to the base band unit (BBU) pool through the fronthaul links. CRANs have several advantages: First, by moving RRHs closer to the users, a higher system capacity and lower power consumption can be achieved because the signal doesn’t need to propagate a long distance to reach the users; Second, since the baseband processing is centralized at the BBU pool, the cooperative processing techniques to mitigate interferences can be leveraged; Third, by exploiting the resource pooling and statistical multiplexing gain, C-RAN is much more efficient in both energy and cost aspects because it is needless to dimension the computing resource of each traditional BS according to the individual peak load. However, the fronthaul constraints have great impact on worsening performances of C-RAN, and the scale size of RRHs accessing the same BBU pool is limited and could not be too large due to the implementation complexity. Furthermore, many kinds of system architectures have been proposed by different mobile operators, manufactories and even researching institutes to explore potential advantages of C-RANs. Therefore, an unified C-RAN for 5G is still not straightforward.

To increase the capacity of cellular networks in dense areas with high traffic demands, the low power node (LPN) serving for the pure “data-only” service with high capacity is identified as one of key components in **HetNets**. One key advantage of HetNets is to decouple the control plane and user plane. LPNs only have the control plane, while the control channel overhead and cell-specific reference signals of LPNs can be fully shifted to macro base stations (MBSs).

Unfortunately, an underlaid structure that MBSs and LPNs reuse the same spectral resources could lead to severe inter-tier interferences. Hence, it is critical to suppress interferences through advanced signal processing techniques to fully unleash the potential gains of HetNets, such as adopting the advanced coordinated multi-point (CoMP) transmission and reception technique to suppress both intra-tier and inter-tier interferences. It was reported that the average spectral efficiency (SE) performance gain from the uplink CoMP in downtown Dresden field trials was only about 20 percent with non-ideal backhaul in.

To fulfill new breakthroughs anticipated in 5G systems, and overcome the aforementioned challenges in both C-RANs and HetNets, the heterogeneous cloud radio access networks **(H-CRANs)** are fully backward compatible with different kinds of C-RANs and HetNets. The motivation of H-CRANs is to embed the cloud computing technology into HetNets to realize the large-scale cooperative signal processing and networking functionalities, and thus SE and EE performances are substantially improved beyond existing HetNets and C-RANs. In the H-CRAN based 5G system, the control plane and user plane are decoupled. MBSs are mainly used to deliver control signalling of the whole H-CRAN and provide the seamless coverage, while RRHs are used to provide the high speed data transmission in the hot spots with huge data services. The large-scale cooperative signal processing and networking techniques in H-CRANs are threefold. The first comprises advanced spatial signal processing techniques in the physical layer (PHY), including the centralized massive multiple input-multiple-output (MIMO) and the distributed large-scale spatial cooperative processing; the second comprises large-scale cooperative radio resource management and cloudification in the medium access control (MAC) and upper layers; the third comprises intelligent and self-organizing network management in the network layer to support self-configuration, self-optimization, and self-healing in the ultra dense communication scenario with huge number of nodes deployed randomly.

More generically the “Wireless Cloud” is a novel paradigm for future wireless networks in which a dense cluster of very cheap and low-energy wireless nodes collaborate to provide transparent communication services to other terminals. Instead of implementing a protocol stack, the cloud nodes operate using a “network-aware physical layer”, which processes the mixture of superimposed signals they receive, and extracts from it the relevant information to forward towards the destination. It uses wireless physical layer network coding to encapsulate the information to be forwarded, coordinated in such a way that each destination can decode the information it requires. The cloud implements distributed self-organisation to perform the required coordination and ensure nodes are synchronised. Because nodes extract information from whatever signals arrive, they are not limited by interference in the same way as in a conventional multihop relaying network, which results in dramatic improvement in efficiency, and by avoiding the need for retransmissions requested by higher network layers the latency is also greatly improved. This is a guiding persisting trend that will be seen from 5G networks onwards.

4.3 The Rural Challenge

A different view on this problem is related with worldwide coverage issues, the need to avoid the digital divide between urban and rural worlds. Appropriate (low-frequency) spectrum bands must be assigned, able to operate with equipment that must be cheap and energy-efficient. Only through standardization of key aspects is this cost-efficiency possible. But extensive dedicated research targeting wide coverage is needed. While current 5G research does not focus on coverage aspects there appear to be a number of promising technologies and innovations that could be taken as a starting point for dedicated research. B5G research is needed in at least the following areas:

- Radio spectrum. In rural regions with low aggregate capacity demands there is no spectral crunch. However, spectrum must be wisely assigned. Today’s use of spectrum for mobile networks requires high output power when distances become large, along with access to mains power for the base stations. With a balanced link budget terminal batteries are also quickly drained. B5G systems should therefore use low spectrum and research in this direction is needed. In contrast to densely populated urban regions, VHF/UHF broadcast spectrum is not fully used in rural areas. The broadcasting grid (transmitter infrastructure and sites) could also be re-used. Moreover, in the long term, the role of broadcast television in general is unclear: with the currently growing popularity of on-demand video services it should be no surprise if society would abandon traditional TV services altogether.
- New coverage metrics. Coverage of a cellular network is not an obvious and simple measure. While it is typically associated with a spatial notion (coverage relates to a certain geographical location) there are more complex aspects relevant to B5G systems associated with people's mobility patterns. The lack of nuances and in-depth understanding of a coverage measure is clearly noticeable in the historical coverage requirements adopted by the regulators. Invariably, the percentage of coverage required in license conditions relate to the place where people live. There is an important role for engineers and the research community to re-define coverage in terms that are useful and will guide the direction of new radio technologies. In particular, the spatial and temporal patterns with which people move should find a natural place in such measures.
- Large antenna constellations. The revolutionary possibilities that emerge when base stations are equipped with a large number of antennas are promising. This technology, often referred to as massive MIMO holds a widely-recognized promise of high capacity in dense urban networks, but it is likely that there is a benefit to large cells and rural areas as well.

5 System Design

The goal of this chapter is to complement the previous ones by providing B5G system design and architectural design patterns. It includes a proposal for B5G design and architectural objectives, a design pattern for integration and specific design patterns for extreme rural convergence. As mostly centered on evolutions over expected 5G standardization, research breakthroughs to these areas may have the potential to reach industry at early stages. A later section will address more disruptive visions.

5.1 B5G Massive Network Softwarization Design

In the 5G phase 2 view, B5G differentiates itself from 5G by further evolution in radio networks performance and capacity and more specifically in the expected network natively enablers in support for massive flexibility unifying and leveraging network virtualisation, functions virtualisation, programmability, autonomicity and softwarization in the radio network, front /middle / back haul networks, core networks, mobile edge computing and software defined network clouds. Such native flexibility would:

- Support for on demand composition of network functions and capabilities to levels unattainable in 5G phase 1
- Enforce required capability/capacity/security/elasticity/adaptability/ flexibility where and when needed
- Enable Management and Control to be part of the dynamic design of the software architecture
- Enable Services to be executed in one (or more) Slices or virtual networks (i.e. a slice is made of a set of virtual machine)

Network Softwarization is an overall transformation trend for designing, implementing, deploying, managing and maintaining network equipment and/or network components by software programming, exploiting the natures of software such as flexibility and rapidity all along the lifecycle of network equipment / components, for the sake of creating conditions enabling the re-design of network and services architectures, optimizing costs and processes, enabling self-management and bringing added values in network infrastructures. Additional benefits are in enabling global system qualities (e.g. execution qualities, such as usability, modifiability, effectiveness, security and efficiency; evolution qualities, such as testability, maintainability, reusability, extensibility, portability and scalability). Viable architectures for network softwarization must be carefully engineered to achieve suitable trade-offs between flexibility, performance, security, safety and manageability.

Network Softwarization is a set of software techniques, methods and processes applicable to a heterogeneous assembly of component networks or an enhanced version of an existing grouping of network components that is operated as a single network. Network Softwarization provides abstractions and programmability that are utilized to deliver extreme flexibility in networks to support variety of applications and services, to accelerate service deployment and to facilitate infrastructure self-management.

Network Softwarization enables the following high-level network native capabilities, their extensions, their trade-offs, unification and integration:

- Network Virtualization (NV) / Network Slices (NS), which enables virtualization of network resources,

- Network Function Virtualization (NFV), which permits virtualization of software-based network functions. Instead of installing and managing dedicated hardware devices for these networking and servicing functions, they are instead realized as software components and deployed on commodity or special hardware infrastructures.
- Network Programmability empowers the fast, flexible, and dynamic deployment of new network and management services executed as groups of virtual machines in the data plane, control plane, management plane as service plane in all segments of the network. Programmable networks are networks that allow the functionality of some of their network elements to be programmable dynamically. These networks aim to provide easy introduction of new network services by adding dynamic programmability to network devices such as routers, switches, and applications servers. Dynamic programming refers to executable code that is injected into the execution environments of network elements in order to create the new functionality at run time. The basic approach is to enable trusted third parties (end users, operators, and service providers) to inject application-specific services (in the form of code) into the network. Applications may utilize this network support in terms of optimized network resources and, as such, they are becoming network aware. As such the behaviour of network resources can be customized and changed through a standardized programming interface for network control, management and servicing functionality. The key question is: how to exploit this potential flexibility for the benefit of both the operator and the end user without jeopardizing the integrity of the network. The answer lies in the promising potential that emerges with the advent of programmable networks in the following aspects: Rapid deployment of large number of new services and applications; Customization of existing service features; Scalability and operational cost reduction in network and service management; Independence of network equipment manufacturer; Information network and service integration; Diversification of services and business opportunities.
- Software-Defined Networking (SDN), which allows network control to be separated from the forwarding plane and allows for a flexible management of the network resources which facilitates the design, delivery and operation of network services in a dynamic and scalable manner.
- Software-defined Network Clouds - Cloudification of networking and servicing functions, which enables ubiquitous network access to a shared services and shared pool of configurable computing, connectivity and storage resources. It provide users and providers with various capabilities to process and store their data and services in data centers. It relies on sharing of resources to achieve coherence and economies of scale, similar to a utility (like the electricity grid) over a network. It uses virtualization concepts such as abstraction, pooling, and automation to all of the connectivity, compute and storage to achieve network services. It could take also the kind of mobile edge computing architecture where cloud-computing capabilities and an IT service environment are available at the edge of the mobile network or fog architecture that uses one or a collaborative multitude of end-user clients or near-user edge devices to execute a substantial amount of services (rather than in cloud data centers), communication (rather than routed over the internet backbone), and control, configuration, measurement and management.

Although many of these features are to be expected in 5G networks, B5G specific network softwarization has following additional key characteristics and requirements:

I) Harmonization of SDN and NFV - Coordination of the current SDN and NFV technologies for realizing B5G mobile network is required.

II) B5G Extensions to the current SDN and NFV - B5G network needs extreme flexibility in supporting various applications and services with largely deferent requirements. Therefore, 5G specific extensions to the current SDN and NFV, especially pursuing even further and deeper agile software programmability is required. For example, SDN data plane could be enhanced to support deep programmability and NFV MEC needs light-weight management for extreme edge network functions, especially in the area of access network and user equipment (UE) or user devices (UD).

III) Considerations for applicability of softwarization - Considering the trade-off between programmability and performance is required. Especially in B5G context, it is important to respect the performance improvement in wireless technologies. Therefore, it is necessary to clearly define the area and criteria for applicability of softwarization in the infrastructure.

IV) Application Driven B5G network softwarization - B5G mobile network is indispensable communication infrastructure for various applications and services such as IoT/M2M and content delivery. Rapid emergence of applications and services enabled in 5G mobile network must be considered in designing and developing the infrastructure.

V) B5G network softwarization energy characteristics - The architecture design, resulting implementation and operation of B5G network softwarization are recommended to minimize their environmental impact, such as explicit energy closed control loops that optimizes energy consumptions and stabilization of the local smart grids at the smart city level.

VI) B5G network softwarization management characteristics - The architecture design, resulting implementation and operation of B5G network softwarization are recommended to included uniform and light-weight in-network self-organization, deeper autonomy, and autonomicity as its basic enabling concepts and abstractions applicable to all components of 5Gnetwork.

VII) B5G network softwarization economic characteristics - The architecture design, resulting implementation and operation of B5G network softwarization are recommended to consider social and economic issues to reduce with 50% the components and systems lifecycle costs in order for them to be deployable and sustainable, to facilitate appropriate return for all actors involved in the networking and servicing ecosystem and to reduce their barriers to entry.

Fundamental research for massive and reliable distributed software architectures is required to be pursued at this moment, launching the fundamental grounds for a research area where mobile networks are not usually delving.

5.2 Integrated System Design

The world of communications and networking is suffering an enormous and disruptive shift from the currently established and strongly enforced static models to extremely dynamic models. This has impacted all dimensions of technology: network devices and end-nodes are more and more represented by mobile devices, SDN is enabling network control to move from a passive model to a reactive model, smart objects are gaining attention and enforcing their own requirements, and even the atomic elements of the network are evolving to abstract objects represented in content and information.

Such evolution will continue with the unceasing incorporation of new technologies and interfacing elements, such as Augmented Reality (AR) and Virtual Reality (VR), which are encouraging people to participate, and finally be integrated within the network. However, this trend is not limited to persons and their devices. The whole environment is more and more integrated into the network, promoting the revolution of the Internet of Things (IoT), which constantly exploits network architectures to their limits and thus exposes new research challenges that should be addressed to facilitate the delivery of the enormous benefits associated with such integration.

One of the most important and difficult to resolve challenges exposed by IoT is the management of the huge amount of information that it exposes. Its extension is present in all planes of the network that supports the deployment of IoT systems and solutions. The data plane is enriched with information from the context and environment of every object that participates in communications. The control plane has to deal with the huge number of objects that will be present in the network. The dynamicity, ubiquity, and extension of IoT networks are translated into increased complexity in the identification and discovery processes.

All those requirements and challenges have in common the need to increase the level of integration between network intrinsic elements (devices, infrastructure, and protocols) and services to build whole systems that cooperate and deliver valuable benefits to the final users and organizations. Such view is difficult, even not feasible, to achieve today, since network architectures do not consider dynamic service requirements and/or their specific contexts and environments.

The proliferation of smart objects connected to the Internet, resulting from the raise of IoT, has increased the benefits provided by the information they manage. This has led the consolidation of new network scenarios beyond typical client-server or even person-service communications in order to involve objects by themselves, which is emphasized in Machine-to-Machine (M2M) communications.

Spreading M2M communications among a huge number of objects imposes enormous challenges in the fields of network control, management, security, and privacy. Moreover, a new network model is required to make them feasible and widespread (standardized) while ensuring their typical operations are efficient and they can work on different and changing environments. This implies the definition and integration of new network elements that provide the necessary network functions to fulfill the requirements exposed by M2M, including the necessary evolution ability to overcome future changes in the communication model.

In addition, new communications models should consider the coexistence of underlying infrastructures with different networking and internetworking technologies (heterogeneous networks) that should cooperate in order to offer whole network services for the mentioned "objects". The network object concept abstracts and generalizes the concept of communication party and network endpoint in order to allow communications among elements connected to such heterogeneous networks. This also implies little changes to the concept of service, which shifts from a simple instance of a function to the efficient accomplishment of a whole operation, regardless of its complexity.

5.2.1 Baseline Networking and Internetworking Architectures

Regardless of new or future research challenges, the network and system design beyond 5G will be highly influenced by the wide spread (globalization) of current network functionality that is bounded within the network architecture proposals found today. They are very different in nature and final objectives. Most of them are focusing in specific problems so it is hard to find complete solutions to cover the requirements exposed by IoT and M2M communications. Below we summarize the network aspects derived from the two most outstanding, and research-focused network architectures.

5.2.1.1 HIMALIS

First we highlight the HIMALIS architecture [Kafle 2010], which has the objective of providing a solution for the inclusion of network objects from different natures (including IoT elements), emphasizing mobile nodes that work on top of heterogeneous networks. It achieves such goal by defining a complete set of network elements that start from the currently deployed in the network.

The design of the HIMALIS network architecture follows the identifier/locator separation design principles described in [rfc6227]. Thus, it includes a new naming scheme for generating host names and identifiers (IDs), which are decoupled from low layer network locators. Host IDs are used as anchor values to allow changes in lower layers without affecting transport or application layers. This way, the architecture provides efficient and scalable support for mobility, multi-homing, and security functions [Martinez-Julia 2012].

Host name resolution to host IDs and network locators is a key aspect of the architecture, and it is provided by the Host Name Registry (HNR). Following the separation of concerns design principle and with the objective of scalability, there is a different HNR instance on each administrative or network domain and each instance is in charge of the mapping entries of its domain.

As hosts have to know which HNR they have to contact, they have to query the Domain Name Registry (DNR) that will resolve the locator of the HNR provided the name of the target host. The DNR infrastructure is hierarchically organized like the current DNS and it only stores the names/locators assigned to HNRs but not the global host names of all nodes. Therefore, the DNR database does not grow as fast as the number of hosts, so the retrieval process for DNR records is kept fast and scalable.

Putting these elements together on the network, as shown in Figure 1, the HIMALIS view of the network is mainly composed of edge (or access) networks, gateways, a global transit network, and the unified logical control network. The edge networks provide network access to various end systems or hosts, which are able to use different access technologies and network layer protocols as well as different numbering spaces for locators (called local locators), which are routable only in the respective edge networks. For example, some edge networks may use prefix aggregately locators, like the current IP addresses, while others may use geographical coordinates given by GPS as locators.

The global transit network, which would be the collection of service providers' backbone networks, will use a single network layer protocol and single numbering space for global locators that are used to forward packets among edge networks. The gateways located on the border between edge and transit networks translate local locator and network layer protocol into the global locator and network layer protocol for communication expanding across the border.

Finally, in the latest iterations of the architecture design, the aforementioned elements are complemented by two other elements: the Local Name Server (LNS) and the Authentication Agent/Registrant (AAR). These are introduced to locally store host information (closer to the nodes and gateways) and also enforcing security in network access [Kafle 2012].

5.2.1.2 MOFI

The architecture proposed by MOFI [mofi] is completely centered around mobile devices, considering mobility the default behavior of the Future Internet (FI). It is primarily built with three functional blocks: 1) host ID and network LOC, 2) global ID-based communication with local LOC-based delivery in the data plane, and 3) search-based distributed mobility control in the control plane. Joining these components, the architecture has complete separation of control and data planes and integrated mobility and multi-homing support.

The first functional block is the Host Identifier and Network Locator (HINLO) [Jung 2011a] and it achieves the separation of host identifier from network locator, an essential requirement for the FI. In the data plane, the packet delivery is accomplished by Global ID-based communication and Local LOC-based delivery (GIC-LLD) which separates the delivery mechanism of access network (AN) and backbone network. For mobility control, MOFI uses the Search-based Distributed Mobility Control (SDMC) [Jung 2011b] in the control plane, which meets with the FI requirements of separation of control and data delivery functions and the distribution of mobility control.

MOFI assumes the host-based communication that is already adopted in the current Internet. HINLO is used to support ID/LOC split, in which independent host ID (HID) is employed and separated from locator (LOC).

The mobility control is performed by the Search-based Distributed Mobility Control (SDMC) functional block, which provides search-based and distributed signaling operations based on the distributed mobility management presented in [Chan 2011]. It naturally implies the separation of the control plane from the data plane. This search-based mobility control may induce a certain amount of delay for LOC query signaling. However, such an initial signaling delay may be negligible in the session setup phase. In case of handover, the LOC binding will be updated directly between the two concerned ARs for route optimization.

The Global ID Communications and Local LOC Delivery (GIC-LLD) functional block provides the possibility for each host to have a global HID with several LOCs used for packet delivery in each network. Each LOC is used locally in the networks, without any assumption on global uniqueness. In GIC-LLD, the packet delivery operations have two stages, i.e. access networks and backbone. On it, communication between two hosts is performed with HIDs, whereas LOCs are used for packet delivery inside the Access Network (AN) and through the Internet backbone. The SDMC control operation is used to find appropriate LOC for each HID in each network.

5.2.1.3 Efficient Mobility Through Dynamic Multi-Homing

From the initial analysis of the general challenges for network and system architectures beyond 5G, together with the specific objectives of the mentioned network architectures, we can highlight the necessity of efficient mobility. This aspect should be present in all dimensions of network research because all network objects, which are the generalization of network nodes, will be dynamic and, at least, non-static by default.

However, the problem derived from mobility can be generalized into dynamic multi-homing. This is because mobility can be modelled as a special kind of multi-homing, in which the object that moves from a network to another is using the multi-homing capabilities of the network architecture to achieve a smooth handover among those networks. In fact, future network objects will be connected to multiple networks at the same time, gaining and loosing links in a dynamic way. Some of those events will be due to mobility but others will be due to availability.

In summary, modeling and implementing mobility and multi-homing together in the network provides a huge benefit to the objects connected to current and future networks while keeping the simplicity of the architecture that supports them.

5.2.1.4 Cohesive and Integrated Network for Service Provisioning

In a world wide network, a user, which can be a person, machine, software, or device and which is generally represented as a network object, will use the available infrastructure to store or retrieve information, send or receive messages, establish a live (multimedia) communication, etc. Such kind of operations should be enabled regardless of the specific technologies, protocols, and even architectures deployed in such infrastructure. This incarnates the nature of modeling network architectures for the provision of services, therefore ensuring they are open to evolution and that they are capable of instantiating any kind of communication required by users.

As mentioned above, current network research challenges and architectures are incorporating more and more elements from such view. At the beginning, they came initially from the inability of current network infrastructures to overcome the problems clearly raised by current requirements, such as mobility and security. However, every revolution proposed by the research community regarding the architecture of the network has incorporated other elements, such as the specific workload and qualities of communications, the representation of real lives into the digital sphere, or even the quick adoption of new technologies demanded by network stakeholders (users, organizations, and operators).

Despite the fact that some architectures include user requirements into their designs, considering the intrinsic qualities of every new function and capability that is added to the network, we can easily identify that most of them are focused in the simplification of network operations in order to improve the efficient and effective exploitation of its resources. However, little is done in current architectures to reflect the specific behavior or requirements of (any kind of) users into the network. Therefore, this would be a key challenge for future research in network architectures.

5.2.2 Slicing architectures: Multi-tenancy as a first-order property

Conventionally, networking infrastructure has been operated under a single-tenant model: whoever owns the infrastructure decides how it is run. This model has been overcome in the general IT world with the introduction of virtualization and cloud computing, and it has come under criticism in the networking work as well. Rather than a single-tenant model, networks today are expected to provide isolated resources and individual control to separate tenants.

Clearly, such a model should integrate tightly with cloud computing tenant models applied in data centres. But it should not end there. Also obvious is the requirement to separate out networking resources, encompassing fixed and wireless networks, on a large scale. Typical examples are the dynamic, on-the-fly creation of mobile virtual network operators.

But the real challenge here lies in the intimate integration of networking and cloud multi-tenancy. It should be able to provide isolated and dedicated resources both on networking and IT processing / storage layer, with the ability to dynamically install software at required places. This is partially supported by SDN and NFV concepts, but not at the deep level of integration.

For example, today's NFV concepts would not be up to the task of dynamically rolling out new CloudRAN functionality for a single tenant. We expect the slicing concept to become a part– in some form, at least - of 5G, but further work will be needed.

This slicing concept entails advances not only in network virtualization and network function virtualization, it also requires dynamic installation of RF functionality (e.g., in the form of downloadable DSP codes) as well as dynamic installation of new functionality in the packet processing pipeline of switches and routers, going beyond today's capabilities of SDN. These capabilities will take slicing architectures to a new level.

5.3 System Design for Extreme-Rural Coverage

A critical performance aspect of 5G networks will be widespread coverage. However, in 5G networks this objective will be not fully accomplished, even resorting to satellite technologies.

5.3.1 Introduction: 5G is becoming an urban system

With the ongoing accelerating development of 5G systems, its most important services and use cases are broadly discussed, as presented in Section 2. New B5G capabilities include for instance radically higher bandwidths and lower latencies. Basic thinking in industry focuses on capacity as the most crucial issue, technically supported by proposals based on more spectrum and dense deployment of base stations.

At the same time, ITU estimates that 3G currently covers 69% of the world's population [itu] and hence a staggering two billion people still have no coverage. On a global scale, and widely recognized today, Internet connectivity is still urban in nature. The digital divide is widening and is arguably of much larger concern than a local tenfold capacity increase in dense urban locations.

Recent initiatives by Google (www.google.com/loon) and Facebook (www.internet.org) capitalize on this idea, both initiatives adopting a revolutionary aerial network topology.

Beyond these initiatives, today's traditional stakeholders, vendors, operators, service providers do not address the issue with much emphasis. The recent white paper by NGMN on 5G [ngmn] sketches many scenarios and use cases for the urban citizens of the coming decades, but few pages are dedicated to the rural coverage. Most worrying is that the "Coverage Everywhere"-regime in the document is restricted to a "service area", those places where operators will choose to roll out their networks. Target data rates are "indicative, depending upon the 5G technology evolution to support these figures economically".

Rural regions suffer the same fate in the recently closed European FP7 flagship project METIS [metis] where the larger vendors and operators joined forces. Here, rural 5G coverage is found under Test Case 7 ("Blind Spots"), in a scenario with 100 users/vehicles per squared kilometre. Disappointedly associated with a disclaimer again: "high data rate coverage is expected at every location of the service area even in remote rural areas".

In short, it appears unlikely that 5G systems will address these issues seriously and to their full potential. Terrestrial 5G solutions under development do not show a convincing evolution path to connect rural areas to the Internet in the coming decade.

5.3.2 Rural regions: why the public sector will become involved

Mainstream telecom business to a large extent ignores scenarios for rural coverage and the silence should not come as a surprise. Not only is it well known that the building cost of a network is related to the area to be covered, but more importantly, as recognized in [Kovacs 2014], "the real barrier to rural deployment is the lack of potential revenue per square mile. The revenue potential for a wireless carrier in a major urban center is $248,000 per square mile of service. By contrast, in the least densely populated areas, the potential revenue per square mile drops as low as $262 per square mile." There simply is no good commercial reason today to connect the un-connected part of the world's population, and hence little substantial and necessary research is carried out.

Yet, there are compelling societal arguments for improved B5G rural coverage: counteract urbanization, improve public safety, develop e-health and entertainment services that make these regions attractive places to reside. In a near future, regulations with respect to rural coverage may well come back to influence research, standardization and network operation. The public sector may become involved in the development of networks and we cannot even preclude new unlicensed types of operation, where users, villages, society or companies would contribute to the network, on voluntary basis.

While spectrum scarcity is not in itself a problem yet [smith] there appears to be a growing need for regulatory and societal 5G-initiatives to achieve progress. Like water, electricity and sewage systems, mobile broadband, too, is rapidly become a basic human need and comes with associated structural transformations for society and policy makers.

Beyond-5G systems will evolve in, and must be developed for, this societal context.

Regulation of coverage is not new. Initial cellular generations were rolled out under tight regulatory requirements. Although these have been relaxed in 3G and 4G, recent 3GPP standardization for mission-critical communication in LTE-A (proximity-based services and network-assisted D2D [Lin 2014]) are influenced by new societal and regulatory demands for improved public safety (not really the high revenue regime many an operator would wish to address).

5.3.3 Potential B5G technologies and the need for research

As discussed above, there are several technologies that need to be developed for cost-efficiency systems for rural environments. We discussed areas where extensive dedicated research is needed for rural coverage: radio spectrum management, coverage metrics, and large antenna constellations. However some other aspects are needed to be investigated:

While current 5G research does not focus on coverage aspects there appear to be a number of promising technologies and innovations that could be taken as a starting point for dedicated research. B5G research is needed in at least the following additional areas:

- Cheap energy-efficient relay nodes. To achieve wide area coverage it can no longer be assumed that every base station is connected to fiber backhaul. The notion of fronthaul and backhaul may not be clear beyond 5G. New autonomous relay nodes could bridge long distance connections in areas where a fiber backhaul is not available. If a power grid is not available either, battery powered operation must be considered. Cheap, energy-efficient relay stations with intelligent sleeping modes could well be an enabler for rural coverage. Battery-operated network elements need a wake-up functionality that allows stations to sleep during time of no traffic. Similar features exist in today's technologies such as IEEE

802.15.4, originally designed for Internet of things or industrial applications. Extensive further research is needed to adopt this concept to the rural B5G networks.

- Transmission technology. Rural areas, where improved coverage is desired typically operate in a low-capacity regime, with very few simultaneous connections. Modulation and coding techniques used in 4G/5G have been designed without such explicit assumption and could well be improved to address the low-power uplink.
- Lean design. Arguably most importantly, a ubiquitous B5G system must be leanly designed. Control signaling overhead, scheduling and other transmission protocols must change into new mechanisms, where no unnecessary transmission is performed: energy should be used efficiently for long-range communications and tailored to the low-capacity, noise-limited regime. Radio access mechanisms including broadcast channels, synchronization channels, random-access channels must be re-assessed with such a lean approach.

While internet access is close to be considered a human right by the United Nations, 5G is emphatically becoming an urban system. We expect that in a near future the public sector in many countries will show a tighter engagement and that extreme rural coverage will become a relevant technological challenge.

6 Alternative Technologies and Designs

The goal of this chapter is to complement the previous ones by looking at a broader range of requirements and scenarios as well as technologies, architectures and design patterns. The hope here is to identify topics that challenge current approaches, business relationships, typical scenarios, or even "common wisdom" by taking an alternative, hopefully creative look at the problem space.

Some of the ideas outlined below in this sense continue current research but reconsider known topics from a new point of view, try to apply known results or approaches in a different manner, or rethink current generally accepted procedures. Other ideas below definitely classify as "blue sky" research with a long-term application horizon; but these ideas could provide entirely new applications than today and provide paradigm-shifting possibilities. Some of these ideas may be simple to include in evolutions for 5G phase 2, but some of them would lead to an internal redesign of the 5G system.

After some considerations about new requirements and scenarios (Section 6.1), the chapter has three main points: it looks at new architectural patterns (Section 6.2), basic techniques and new communication paradigms that complement these patterns (Section 6.3), and lastly looks at some ideas for future communication technologies (Section 6.4).

6.1 The missing requirements and scenarios

Before delving into some concrete ideas, it seems advisable to briefly consider any potentially unusual scenarios and requirements for a post-5G system. Clearly, today's common figures or merit all apply: performance, efficiency (in spectrum, energy, time, …), dependability (as an umbrella term spanning over security, availability, reliability, maintainability, etc.), flexibility, sustainability, scalability, and so on. From this perspective alone, post-5G systems will not be substantially different than 5G or even 3G and 4G systems. The target numbers for these figures of merit are likely to become more challenging, but it seems implausible to assume that an entirely new type of metric can be discovered that has been overlooked so far.

Also, the type of scenarios will, at a first glance, not change substantially. Section 2.1 seems to cover most of our technology imagination at this stage. The scenarios will become even denser, the application requirements will become more stringent, the scales even bigger, the number of devices higher and the data rates even bigger. The type of applications might shift somewhat, with even more consideration given to IoT and M2M applications as already the case in 4G and 5G research.

Nonetheless, in the discussions leading up to this section, some shift in emphasis was notable.

> It is no longer just the mere increase in some arbitrary performance number. Rather, metrics that are less easily quantifiable come to the front: flexibility, dependability, ease of use, even the integration of humans in new ways: new ways of exploiting the human body as well as human behaviour, visible or invisible to the individual.

In a way, some of the work here could be characterized as the reappearance of the human, and the society itself, on the research scene after the heavy IoT and M2M focus of recent years, that will lead the development of the upcoming 5G network.

This is also evident in the need to consider both technological metrics as well as service metrics jointly. Again, this is an old trend but likely to become more and more important:

neither does it make sense to look only at a network in isolation nor is it useful to ignore the capabilities and limitations of the actual network when deploying or managing applications.

6.2 Architecture patterns

Over the course of its existence, both the Internet as well as the mobile communication systems, both practically deployed and in research, have suffered from an overabundance rather than a lack of architectures. Often, these architectures have been little more than bug fixes to earlier iterations but have nonetheless been pragmatically necessary. In recent years, however, a research trend started to emerge to consider the fundamental *patterns* behind communication architectures: what are the underlying principles, the core abstractions, the absolute invariants in a networking architecture?

This work has produced some initial results, but it still harbours many unresolved questions. The hope here is that we concrete architectures can be derived from these basic patterns and that these architecture excel in many ways: not only mere performance but, more importantly, manageability, maintainability, low complexity, and simple applicability to new challenges.

The following sections outline three trends for such pattern – reclusiveness, and information-centricity. Also, an old idea – converged architectures – reappears to address new challenges.

6.2.1 Recursive architectures

In computer science, recursion is a basic principle: We build a complex functionality out of simpler instances of the same functionality. Recursion has turned out to be a powerful concept and has, for example, paved the way to current trends like virtualization: a machine architecture is built of out a simpler version of itself.

Oddly enough, this principle has not been embraced in networking. While from the very beginning of networking it was clear that a network has to be subdivided into simpler subsystems, these subsystems were very different and functionally very specialized (compare the ISO/OSI 7-layer stack or the Internet 5-layer stack).

Historically, there have only been a limited number of attempts to develop a recursive network architecture (e.g., [Day 2008]). More recently, these ideas have been taken up by the RINA [Trouva 2011] project, which attempts to build a networking stack solely on the abstraction of interprocess communication. What is commonly perceived as a layer simply turns out to the reapplication of IPC at a different scope, and these additional "layers" can be added at will without introducing significant overhead. Similarly, the PURSUIT project (from the context of information-centric networking) seeks abstraction that can be re-applied at different scopes, with different responsibilities. An example from the NFV domain is UNIFY, which also encompasses various recursive features it its "Universal Node" concept.

A primary, if yet unfulfilled goal in this research area is to find core patterns that cannot be further simplified – in the sense that perfection is achieved when nothing can be taken away any more. The advantages of such radically simplified architectures compared to today's jungle of complexity are obvious, yet much work is still to be done here. For example, issues in how to design control loops, security models, the optimal separation between invariant vs. variant parts of an architecture, choosing optimal parameters over a set of recursion levels, the best way to deal with resource isolation, how to enable operators, deployers, and programmers to leverage the powerful flexibility of such recursive architectures are all issues that still have be resolved. Exploring this model in depth is a challenging research task that has the potential to dramatically simplify how we design, develop and compose the structure and building blocks of distributed applications, operating systems and networks; achieving a tighter integration between these three areas and ultimately obtaining a simpler, more predictable, performing and secure distributed computing infrastructure.

These open issues notwithstanding, it is very plausible to assume that already the existing architectures like RINA would form an excellent platform upon which to extend such experiments.

6.2.2 Information-centric architectures

Currently there are major research efforts to devise an architecture for information-centric networking (ICN) either on top of the current Internet, next to it or even replacing some of its basic tasks. Examples are Data-Oriented Network Architecture (DONA), Content Centric Networking (CCN), Named Data Networking (NDN), Network of Information (NetInf), or Pursuit. Most proposals provide an overlay architecture supporting content distribution and retrieval. Current work in this area encompasses routing and name resolution, security (anonymity, privacy, access control), multipath forwarding particularly in hybrid access scenarios, and positioning ICN as a generic platform for 5G applications of different kind, e.g., also for Internet of Things from an Information distribution perspective.

6.2.2.1 Pervasive data sharing as next step

While all these are necessary tasks, it appears unlikely: i) that this will suffice to address all the research challenges in the context of ICN; ii) that it will be enough to put ICN firmly on the roadmap for actual 5G systems; and iii) that it will suffice to develop ICN further towards *pervasive data sharing (PDS)* systems. This research need is based on the expectation that B5G networks will be capable of connecting everything from people, things, content, knowledge, and information, independently of the service/data location, and network availability. Consequently, PDS is expected to support connected sensors, connected vehicles, smart meters and smart home gadgets way beyond our current experience with personal device (e.g. smartphone) connectivity. This will impact the value chains, roles and relationships between the players; it will also open new innovation needs addressing the necessary self organization of such systems – clearly, manual management will fail. An example for such ICN-oriented self organization is the integration of opportunistic communication (compare also the sections on D2D communication later on).

Current ICN proposals do not provide any guideline suitable to devise a B5G architecture able to support pervasive data sharing: focus is on ICN for current mobile applications. Hence, the research agenda B5G should start by investigating a set of networking paradigms to be considered in the design of B5G PDS systems able to support the expected heterogeneity of applications in different sectors of our society, such as health care, education, industry, and government agencies. A critical aspect is the whole challenge of data security, with aspects as legal intercept (whether advisable or not), which could here be rethought to come up with a “flow record” analogy to mollify regulators, or aspects like automatic deletion of records intended to be kept only for a limited time.

In addition, the considerably larger scale of such PDS will require more sophisticated ideas about request routing and scalability. Some first ideas exist here (routing on hyperbolic metric spaces), but these are only in their infancy and are far from being worked out.

6.2.2.2 From named function networking to object-oriented networks

In ICN research so far, data has been predominately been considered as a static collection of bits. But this is far from the way data is typically treated in IT today: Even a simple web page today is a highly dynamic entity, computed on the fly upon a user request by executing functions chosen by an application developer for this purpose. Such a model of dynamically generated data is difficult to incorporate in most ICN proposals that exist today. Only recently some first attempts have been made here, known under the term *named function networking*.

This development is fortuitous as it coincides with the pervasive introduction of software into the networking core: the switch from fixed functionality to virtualized, easy to deploy network functions (network function virtualization, NFV). NFV shines in that it allows to dynamically deploy new functions and chains of functions at will at short delays. But it is, as the name correctly says, only *functions* – it has no intimate dependence on the data these functions work upon. While this limitation may be acceptable in many common NFV scenarios, the next step suggests itself when conceptually combining NFV with ICN: instead of only having functions manipulating packets in a flow, we should think about methods manipulating objects. In a sense, this repeats the step from functional/procedural programming to object-oriented programming, where we can finally talk about classes of objects (obtained from a suitable ICN data model) defining methods executed by an NFV-based machinery. Clearly, there is considerable conceptual, architectural, protocol and implementation work necessary here, but the advantages could be considerable.

6.2.2.3 Relationship to recursive architecture patterns

In principle, ICN architectures can – and have been – developed and designed just like another layer in a conventional layering model or even as sitting merely in the application layer on top of a conventional IP stack (in that sense, conventional peer-to-peer networks already are an information-centric network).

In fact, some ICN architectures – notably PURSUIT and its follow-on work – have been designed with reclusiveness in mind. This highlights the feasibility of building even complex systems in a recursive manner. In this sense, recursively structured ICN architectures are more focused than the more generic, interprocess-communication-oriented recursive approaches like RINA.

In follow-up work, it remains to be determined which of the more specialized concepts can or should be integrated in a generic system underpinning and what should be left to a specific architecture. Such insights could serve as blueprints for functional separation in a host of networking architectures (think data centre networking, for example).

6.2.3 Converged architectures

The section 5.2.2 on sliced architectures has already highlighted on imperative for B5G systems: the need for close integration of different system aspects and functionalities. This need has already been considered in the past under the designation of “converged architectures”; a designation with a wide range of interpretations ranging from the convergence of cable and data networks, phone and data networks, or different data network (rather: link layer) technologies. From a more fundamental perspective, we perceive a need to consider at least two levels of convergence:

- **Technological convergence**: Individual techniques should be investigated regarding their applicability across different technological domains. An example would be coding techniques, which have been researched fairly independently in the radio context and the data distribution context. A tighter integration of such work seems promising. Another example would be the integration of processing functionality with data distribution functionality (NFV meets ICN) as outlined above.
- **Conceptual convergence**: Basic architectural patterns have been developed (reclusiveness, information-centricity – see above) but have predominantly been applied to a single or a few domains (e.g., core networks). While these patterns do claim universal applicability, this is so far an untested hypothesis. We require research that test this hypothesis and finds out whether these patterns can be applied to a wide range of different technological domains (e.g., radio access or optical backhaul) as well as different scenarios.

We postulate that the mechanisms for better convergence of such research trends employed so far (cross-project activities for projects that are themselves specialized and are already very large, e.g., the 5GPPP cross-project activities) are not well amenable to this goal as such work is not part of the mandate of the projects but is required "on top".

In a nutshell, funding research for "cross-X" (cross-domain, cross-technology, …) technologies and methodologies is needed, specially because such critical fundamental research items would today frequently have a very hard time meeting typical selection criteria for European research funding.

6.3 Basic techniques and paradigms

The previous section has concentrated on big-picture patterns: how to organize a communication (and storage and processing) system on a large scale. This section zooms in on individual techniques and paradigms that promise considerable advances and new ideas.

6.3.1 Cognitive self-organization

Network operators are currently spending between 15 and 20% of their operational costs on operating and optimizing their networks. The problem is exacerbated as "smart things/objects" are expected to become active participants in networks and autonomously manage their communications, to harvest the energy they need, and to configure themselves when facing contextual changes. Such objects will be able to show "intelligent/cognitive" behaviour when faced with unpredictable situations, e.g., intermittent connectivity. Based on such contextual cognition, these objects should be able to coordinate and manage their behaviour amongst themselves, possibly with the aid of a coordination framework that partially rests outside themselves.

Cognitive management of IoT and M2M networks include self-configuration, self-optimization, and self-healing features. These features (which have been heralded in research in the last years, albeit with very limited results) could be implemented through M2M gateways that would be present at the border of a B5G network; they could also be implemented in a truly distributed fashion without having to rely on dedicated boxes.

Unlike previous research efforts in this direction (e.g., mobile ad hoc networks, wireless sensor networks), we claim that a new quality is possible by integration advances in in data analytics, machine learning, and artificial intelligence. Both the coordination quality as well as the ability to handle large amounts of data should increase considerably. For example, knowledge about the roaming of networked mobile objects will be beneficial (see below). Possible inspiration to the challenges in this area could come from human and social cognition research.

A fundamental challenge is hence to assist the network in self-organizing to deploy, optimize, and coordinate a large number of devices in dense environments, evolving from the simple approaches realized for current (claimed) cognitive technologies into widespread large information processing at fast rates.

6.3.2 Anticipation

Related to cognitive approaches, but a separate functionality, is the idea of *anticipation*. It refers to a system's ability to predict external events that will impact its ability to deliver the desired quality of service/experience. Such external impacts can be user behaviour (movement, starting or stopping downloads, etc.), changes in the environment (radio channel variation), or even faults – many options exist here.

Since no crystal ball is perfect, it is well understood in this research direction that all anticipation has only limited accuracy for a limited time horizon. The challenge here is to derive operational regions for both the prediction techniques themselves, as well as for proper anticipatory actions to take. The difference to conventional prediction techniques (deployed, e.g., in radio channel prediction) is the idea to work over longer time horizons with potentially large errors and to incorporate other sources of information (e.g., user behavior, weather, traffic news).

Two basic examples are roaming behavior supporting D2D communications and application behavior in video streaming.

On the former, anticipation techniques are used **to better assist individual and collective roaming based on anticipated user behaviour**, providing better self-organization. Current self-organization in the context of clustering and networking operations stems mostly from bio-inspired environments, or from synthetic roaming models that attempt to mimic individual and collective human movement behaviour. In the future, it is envisioned that self-organization modelling shall become human-centric and integrate models derived from social sciences, such as social psychology (e.g. group formation theories; emotional contagion and social bonding/sharing modelling). Such theories can assist in anticipating and better estimating how to perform resource and mobility management derived from knowledge on group formation theories, and based on device sensing and learning.

The later example for anticipation is video streaming. The core idea is relatively simple: Suppose a user streams a video over a wireless link. The user's behaviour (e.g., stopping the video, fast-forwarding) might be anticipated to some degree based on past experience; similarly, user mobility might be anticipated (at least, over the range of some tens of seconds). Based on user mobility, some ideas about radio quality is available (again, rough estimates over some tens of seconds is sufficient). Based on this information, the prebuffering of video streams can be fine-tuned to improve video playout (no stalls, improving QoE) and balance it against the amount of radio resources invested (no needless download taxes the radio network).

A core part of anticipation is anticipation of user activities. Two different approaches are considered to learn individual activity: participatory sensing and opportunistic sensing. With participatory sensing, the user consciously opts to meet an application request out of personal interest. With opportunistic sensing, the user may not be aware of active applications. Instead a user's device (e.g., smartphone) is used whenever its state (e.g., body location) matches the requirements of an application. The **non-intrusive approach of opportunistic sensing** is known to more easily support large-scale applications, such as the ones based on an accurate understanding of group dynamics (e.g. spread of viruses). Aspects that can easily be integrated concern the adequate definition of context similarity (space, co-presence, inter-contact times; activities similarities in terms of roaming - speed, characteristics); social (embeddines in groups); relational (identification of basic relational types based on dynamics and physical aspects of recent interactions). Furthermore, such concepts have to take into consideration the need to ensure privacy and anonymity at all times, even for the cases when observations shall take into consideration social interaction. Hence, data tracking at an individual and local level is essential. All data, including identifiers of the devices, are to be made anonymous.

Other prediction techniques well known from other fields – e.g., channel prediction – of course can be leveraged for anticipatory systems as well. Clearly, research is needed to reconsider the time horizon and accuracy these techniques work on – unlike previous work, longer but less accurate predictions may seem suitable here.

6.3.3 Participation and proactivity

A classic approach in networking research chooses to be deliberately blind to the characteristics of the data it transports, choosing not to require any particular information from the application that sends the data. There are good reasons for this approach both from information-theoretic, legal/regulatory (net neutrality) and pragmatic (the failure of fine-grained QoS signalling over decades) points of view.

Nonetheless, this self-induced blindness has been reconsidered from various directions recently. Some examples discussed above are information-centric network, traffic-pattern induced D2D communication, or anticipatory video streaming. Another, extremely compelling example can be found in data centre networking where an application framework like Hadoop informs a network controller (in the SDN sense) about its incipient actions, allowing the SDN controller to optimize traffic and prepare the network for any possibly impending large-scale transfers. Similar ideas are currently investigated in the context of NFV research. Similar ideas also exist in the optimization of radio backhaul networks.

But these are only individual dots on a design space are not yet connected into a coherent concept: what type of information is actually necessary from an application to benefit a network or, more generally, an IT infrastructure? What is the right level of granularity to impose? How to express semantics of traffic patterns? Clearly, the old ideas of Quality of Service signalled have to be reconsidered here, being too fine-grained to be actually useful. But the examples above show that there is indeed value in understanding impending application actions and being proactive in preparing a network and an IT system.

6.3.4 Few-hop networking

In current networks, data typically takes a large number of hops. This happens nowadays mostly in the fixed network; the radio systems are today still predominantly one hop. Some research efforts ongoing and proposed here will lead to an increase in the number of hops data has to travel (e.g., D2D introduces additional hops in the wireless link). While there may be good reasons to do so, each hop also introduces additional delay, complexity and overhead.

As an alternative approach, one might consider building networking systems that are optimized towards as few hops as possible. From a wireless perspective, this could mean bridging larger distances with more antennas, creating longer hops (with rather conventional techniques of heterogeneous functionality between simple end devices and complex base stations). In the fixed access and backbone infrastructure, it could mean leveraging modern features of optical networks like direct lightpath routing at very short timescales and with acceptable configuration overhead. These technological options exist and are now also possible to support by SDN-based control structures.

In this sense, few-hop networking is an idea that challenges current major trends in the research community, but which is particularly valuable just because of that. It promises a significantly reduced overall system.

6.3.5 Conflict resolution

With the increased dynamics currently foreseen for networks and IT systems alike and with their tighter intertwining, with the increased strive towards multi-tenancy and fine-grained control, we will unavoidably run into conflicts between different types of control loops. This will become much more challenging to deal with than today's relatively simple conflicts. Today, two data flows conflict over data rate on a bottleneck link (easily handled by congestion control); tomorrow there will be conflicts between entire virtual networks. These conflicts will continue between different applications running in SDN controllers (one control application decides to forward a data stream, another one forbids it); they will also appear in functions applied to a data stream applied by NFV systems.

To consider in more detail one example from self-organized networks discussed above: Here, multiple cognitive functions may have conflicts of interests as may try to perform actions that are incompatible. Avoiding this requires an orchestration function at the network level for decision-making and enforcement of the priorities of the functions. For instance, while a function aims to shutdown a cell or M2M gateway to save energy and minimize interference, a preventive troubleshooting function may detect and treat this as a case of cell outage. From a system point of view, this is undesirable.

In summary, conflicts will change from a microscopic event to a large-scale issue for networks. There is currently no consistent approach visible, barely even an idea how to formulate a theory to express and deal with such conflicts. Some initial work (tussle spaces) exists, but it lacks application and further in-depth investigation. This has the potential to become a serious issue if not tackled early on. This is an area that will be clearly associated with the expect challenges coming from the Massive softwarization of the networks (Section 5.1)

6.4 Specific technologies

From architectural patterns and basic building blocks, we now switch to individual technologies that are on the horizon but for which additional research is still needed.

6.4.1 Visible Light Communication

One possible alternative technology is optical wireless communications (OWC), which includes infrared, visible and ultraviolet sub-bands. Compared to radio communication, OWCs offer superior features such as ultra-high bandwidth (THz), freedom from electromagnetic interference, a high degree of spatial confinement bringing virtually unlimited frequency reuse, cost effectiveness with no licensing fee, and potentially inherent physical security. With plenty of resources, spectral efficiency is less of an issue.

OWC systems in the visible band (390-700 nm) are commonly referred to as *visible light communication (VLC)*, which takes full advantage of visible light emitting diodes (LEDs) for the dual purpose of illumination and data communications at very high speeds. VLC can provide solutions for a number of applications including wireless local, personal, and body area networks (WLAN, WPAN, and WBANs), indoor localization and navigation (where current GPS is not available), vehicular networks, underground and underwater networks, offering a range of data rates from a few Mbps to 10 Gbps. The complementary *free space optical (FSO)* systems operate at near-infrared frequencies (750-1600 nm) and offer a cost-effective, protocol-transparent link with high data rates (10 Gbps per wavelength) and provide a potential solution for the backhaul bottleneck over a short to long range up to a few km in urban areas and longer for rural areas.

VLC systems are currently attracting attention due to growing use of LEDs for general lighting in a multitude of applications. LEDs are solid-state devices similar to IR LEDs, which make them suitable not only for lighting, but also for data communications, a characteristic unique to LEDs with unimaginable implications. Current research trends have demonstrated the feasibility to achieve high data rate communications (links with data rates above 3 Gbit/s have been demonstrated) exploring this dual LED functionality. Another interesting characteristic of VLC systems is their spatial confinement. When used in indoor scenarios, the communications range is limited by the room size, since no radiation crosses the walls. Spatial confinement may also be explored on multi-user scenarios, where different light sources carry different data.

Promising scenarios that require further research are:

- Cellular OWC: Optical access points could offer localised high data rate in a heterogeneous network setup, leveraging the low levels of interference characteristic to this medium. Localization functionality could be integrated.
- Optical hotspots: Specifically in an indoor scenario, hot spots could be directly attached to an optical fibre that provides the data stream to be sent out directly. This evolves older radio-over-fibre concepts to VLC-over-fibre. In such scenarios, research is necessary how to deal with the widely varying uplink vs. downlink data rates.
- Smartphone short-range communication: Optical communication could provide a highly directional (pointing at the smartphone's camera), short-range, possibly even secure communication channel to provide an alternative to today's RFID technology. Unlike these technologies, very high data rates (tens of Gbit/s) are plausible here and would allow very different usage scenarios (e.g., movie download to smartphone after checkout in very short time).

6.4.2 Dynamic radio infrastructures

Currently, our radio infrastructures are relatively static: It takes a long time to deploy a new cellular base station, installing a WLAN access point also is time consuming, and concepts like D2D only start to promise some more capacity on demand. Some ideas to introduce base stations more flexibly have been considered: balloons or unmanned aircraft to carry base stations, but that work was mostly dominated for a coverage and cost perspective.

What about if we used the advances made in unmanned aircraft? The ideas and successful prototypes to deliver cargo by drones have shown that it is indeed feasible to get drones to a desirable spot in the air very quickly. Such drone-deployed aircraft could be means to provide capacity very quickly, very flexibly to almost arbitrary points; they could even follow the need for capacity (think large-scale events, demonstrations, etc.). These capacity-providing drones could mesh nicely with conventional infrastructure and dynamically, on-the-fly patch it where needed. This would allow a very different dimensioning of our networks, no longer providing for peak load with the fixed infrastructure but only for (perhaps) median load. Taking these dimensioning aspects into account, even energy efficiency of such a solution could be quite acceptable, even more so if these capacity drones could be shared with cargo delivery drones.

While it is perhaps a strange idea, it is something that could lead to unexpected results down the road.

6.4.3 Emergency communication

Related to dynamic infrastructure is the need to support emergency communication. Clearly, dynamic setups like drone-based base stations can be quickly deployed in emergency situations, providing either coverage (for large-scale failure of an infrastructure in case of, say, an earthquake) or capacity (in case of events with massive amounts of damage to people, say, mass panic events like during the LoveParade). Moreover, such drones might be suitable to assist in location services, being able to quickly change their location and provide dynamic triangulation, something that is not possible with other techniques.

But even without such flying base stations, our radio networks should be able to efficiently support emergency use cases and survive themselves in such situations. In previous mass panic events, a common detriment was the failure of the cellular infrastructure under overload (compare again LoveParade disaster). The network is not able to switch to different operational modes where other priorities exist: limit to text messages, maybe focus on location or presence services. Similarly for earthquake scenarios: how to switch to a mere sustainability mode but still carry actually vital information?

We currently have no concepts how to design networks, handsets and protocols that can switch between such operational modes and concentrate on different tasks in extreme situations. While this might not be anything that provides short-term revenue, it could indeed safe lives.

6.4.4 Cellular radar

Huge research effort in the last years has gone into the “Internet of Things” and pervasive sensing. The common idea is to flood our environment with a large number of devices doing all kinds of sensing tasks. These devices could be randomly deployed, deployed in some form of organized way (e.g., roadside sensing) or be part of common smartphone platforms; these devices are then all connected together to exchange data over various, complex, mutually incompatible communication protocols. While there may be many good reasons to pursue such an approach, it appears that an promising sensing opportunity has been overlooked.

This opportunity is to use our existing cellular infrastructure not only as a communication but also as a sensing platform. Adequately equipped with sufficient bandwidth, our collection of base stations can be thought of as a huge distributed radar array. This *cellular radar* has the potential to deal with a lot of the sensing tasks that, in today’s visions, still require many additional devices which effectively pollute the environment.

While cellular radar is so far not well investigated, it seems an extremely promising concept. It does away with many issues of current IoT concepts and could again much simplify our overall IT architecture. Clearly, there are huge research problems to be solved, the actual radio/radar questions being only one of those; privacy and safety concerns are also very dominant questions here.

6.4.5 Self-powered devices and systems

Energy management is a challenge for the upcoming B5G systems. Stringent techniques for reducing energy expected in wireless communications are being addressed. However, several other possibilities exist, albeit still in its infancy. Most notably, energy harvesting (systems that extract their energy from the electromagnetic environment), and self-powered devices (systems which are able to generate their own energy from the mechanical environment, such as human walking), are two promising technologies for low-power devices.

Nevertheless, the expected power levels extracted from these technologies for practical cases are such that very strong limits will exist on the wireless power budget of such devices. This will have impact at system level, raising research questions on how to design (e.g.) effective sensing networks built from self-powered devices.

6.4.6 Device based applications & services support in Beyond 5G Networking

Nowadays, users are starting to demand services that are satisfied through the combination of their different electronic devices. The Chromecast device from Google is a successful example, where a relatively simple dongle that is attached to a television receives video streams from the Internet, but the control is performed using apps installed in a smartphone. Arguably most importantly, a ubiquitous B5G system must be leanly designed. Control signaling overhead, scheduling and other transmission protocols must change into new mechanisms, where no unnecessary transmission is performed: energy should be used efficiently for long-range communications and tailored to the low-capacity, noise-limited regime. Radio access mechanisms including broadcast channels, synchronization channels, random-access channels must be re-assessed with such a lean approach.

This is the ubiquitous (or pervasive) computing paradigm, but other concepts such as the ATAWAD (Any Time, Any Where, Any Device) also are related to this scenario. Users will not only want devices that collaborate to provide an enhanced functionality, they will also want to use any of those devices to have access to their information.

Research in the ubiquitous computing field has provided protocols and mechanisms to coordinate and distribute tasks and information between devices. The IoT (Internet of Things) offers global connectivity to any object. Nevertheless, further research is necessary in order to combine both approaches and enable complex scenarios and rich applications. For example, in a pervasive telepresence system, a user will want his movements to be perceived by sensors installed in the room where he is located and replicated remotely, and also to capture what is happening remotely and represent that information locally. This application should work with any kind of setup or configuration, but the scenario is really complex, and requires protocols that coordinate and distribute information among the sensors, devices and actuators in each room, synchronize the events, reserve dedicated bandwidth and guarantee low latency.

The SDN (Software Defined Networking) approach will empower 5G core and access networks because it allows to create advanced applications that make decisions about the management of traffic. SDN protocols should be extended to manage also user devices, a particularly challenging task as this will naturally be a low-trust multi-tenant environment. This way, it will be possible to create centralized applications that will orchestrate and manage hundreds or thousands of devices to provide complex services (e.g. telepresence), and at the same time guaranteeing a perfect synchronization, and the required latency and bandwidth from end to end.

6.4.7 Human-machine (network) interface redesign

Taking some of the ideas in emergency communication and cellular radar at step further, an opportunity for a redesign of our human-machine interface presents itself. This interface has evolved considerably over the last decades, with smartphone interaction today being the dominant form for a big part of the population. In fact, we already witness a human-network interface, mediated by computing devices. We have concepts like wearable computing (from smart watch to smart jewellery and smart jackets) obviously on the horizon, but these concepts still have the same basic distribution of functionality: somewhere on your person, you have a device that acts as local processing hub (today, your smartphone), aggregating the local interactions of a person. This hub then communicates with the environment (cellular, WLAN, …). Note that the challenges raised by section 6.4.5 on self-powered systems may lead to different system structures, but fundamentally such device-hub-network structure may well remain.

What if such a hub were not necessary? What if, instead of processing a lot data locally, this processing could be offloaded towards a, say, cellular network? This does require a fairly efficient long-range communication, but the particular data rates could vary tremendously so that the overall energy consumption need not necessarily be a problem.

And taking that idea yet another step further: what if cellular radar actually worked to a high degree and gesture recognition were doable entirely on base stations/access points? This would be unlikely to be suitable for high data rates, but it would create a control channel of unknown characteristics and unknown usage – offering the opportunity for entirely new applications. A simple idea could be to recognize emergency gestures (known e.g. from maritime signalling) via a cellular system and initiate rescue procedures accordingly.

6.4.8 In-body, on-body, from-body communication

The notion of wearable computing and “smart accessories” is well embedded in current IT/networking visions. So far, the communication techniques foreseen here are fairly straightforward versions of well-established techniques (e.g., IEEE 802.15.1 or .4, etc.). But thinking these ideas through shows still untapped research potential.

First, it is foreseeable that processing, sensing and actuating devices will become into the human body in the context of medical procedures. Programmable insulin pumps or pacemakers already exist, smart prosthetics are around the corner. This raises a need for *in-body communication*: should the insulin pump or the prosthetic talk to the pacemaker? Is it sufficient to rely on chemical/molecular communication taking place anyway in the human body (raise and fall of glucose levels in the blood stream, e.g.), should that be ameliorated by additional molecular communication streams? Or are alternative communication techniques preferable, like ultrasound? Similarly, we will need communication schemes to talk to those in-body devices in an efficient and reliable manner; security here also takes on a much more personal note than it has so far.

Likely, these in-body devices would be accompanied and enhanced by on-body devices, attached to the skin or close-by (think smart watches). What are suitable communication technologies for these devices? Can we use the conductivity of the human skin to transmit the required amounts of data, at required reliability? Would that outperform existing body-area network technologies? Or is transdermal optical communication the medium of choice here, exploiting natural translucency of skin and flesh?

And lastly, we need to communicate information from (and to) the body. Clearly, existing technology could be the contender here, but it needs to work at very high efficiency. Perhaps a radical redesign is required here as well: consider the human body itself as one very big antenna!

7 Conclusions

This document provides a snapshot of possible and probable research ideas for B5G systems. Given the nature of the aim, this is not a finished document, but one that will change as technology evolves. The document moves from ideas that will naturally evolve from the expected technology developments inside 5G, to more radical approaches for communication networks, and to more radical ideas of B5G as complex systems involving all aspects of human communication.

As mentioned earlier, the post 5G phase 2 era is from 2030 onwards when many of the concepts discussed above will have been imbedded into the 5G architecture. In this period we expect that capacity issues will have been solved and this will no longer be a major driver. Coverage, reliability and QoE as well as energy may have been solved but will probably still remain a major challenge. For this era we do not really have a vision and perhaps looking to the longer term this is where the research should start. What for instance will be the applications and user requirements in this period? What are the longer term technologies that will play a part? Quantum computing and communications, terahertz technology, massive parallel processing, novel materials and nanotechnology and what is beyond IP? All of these technologies and more will almost certainly mean that Beyond 5G networks (in the 6G path) will be completely different from those deployed till 2030. This means looking 20 years ahead and will constitute a different programme of research which this document is not addressing.

The overall views addressed inside this document reflect more native networking "intelligent" and "flexible" ideas – spanning all technology levels. We see challenges in bringing more intelligence in Massive MIMO, and how to realize it; in bringing more intelligence to spectrum management, and how to optimize its usage; in bringing more intelligence to networking, with software networks and network softwarization designed with embedded intelligent dynamicity; in bringing more intelligence and flexibility to forecasting system actions for anticipatory behavior. We see more flexibility on the ability to use D2D links if available; on the coverage of both urban and rural environments with the same technologies; on dynamic multi-tenancy networking; on the reflection of human desires on networking action.

This positions the needs for funding on a new programme: the technological underpinning of such a programme should always **reflect the development of realizable, real-time, technologies and methodologies able to provide more intelligent actions and react more flexibly** natively to the environment requirements.

In a more concrete enumeration several important messages have been identified at different levels, some of them to be expected from the current perception of the state of the art:

- Terabit communication will require research spanning from physical devices to communication models, to integration wireless-fiber, to advanced HetNets processing
- MIMO will remain a challenge from the perspective of Massive MIMO realization.
- Spectrum management will need to engage in radical research, with regulatory impact
- Technologies to reduce the rural/urban divide is an area that will become critical after the first 5G deployments.
- B5G will require software systems with levels of multi-tenancy, scalability and flexibility hard to fanthom at this moment, all areas where no fundamental and realizable theories exist. Conflict management (amongst others) is an area where we see current knowledge being stretched to its limits in this new world.

- Furthermore, B5G systems will be inherently integrated mobile (at multiple levels) environments, adding more challenges to reliable system design. Architecture patterns may well be a strategy to simplify the complex problem ahead.
- Research on the interaction of human aspects (both from the user and from the designer perspectives) will be essential for the future evolutions of 5G systems. Anticipation, proactivity, participation are features we expect to see in this future.
- Several “novel” technologies merit research opportunities to assess their true potential: optical wireless, dynamic radio systems, cellular radar, self-powered systems, device-level softwarization (and slicing), and body-level interactions.

There is ample ground for research to drive our current systems forward, to make them fit for the challenges ahead, and to also open the door for ground-breaking, scenario-changing technologies.

In the past, we have led our research efforts overwhelmingly from the basis of scenarios that we could easily imagine. But we can only imagine scenarios on the basis of *current* technology, those getting stuck in incremental refinement processes – even this document has a fair share of such research directions. Perhaps it is time again to embrace more strongly opportunities for weird ideas, things that we are not sure if (or when) they will work and where its need is not easy to (immediately) perceive.

But one thing is very clear: our future research will be increasingly intertwined with human behaviour and policy rulings. We recommend that the B5G research community takes self-initiated, change-oriented and anticipatory steps to support the engagement of the wide public sector, by investigating relevant new technologies that may span and impact society in its widest sense and signal the consequences for society and business change from the very inception of concepts at the research stage.

8 References

[Ngo 2013] H.Q. Ngo, E.G. Larsson, and T.L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," IEEE Trans. Commun., vol. 61, pp. 1436–1449, Apr. 2013.

[Marzetta 2010] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," IEEE Trans. Wireless Commun., vol. 9, pp. 3590–3600, Nov. 2010.

[Gao 2013] Xiang Gao, Ove Edfors, Jianan Liu, Fredrik Tufvesson, "Antenna selection in measured massive MIMO channels using convex optimization", IEEE Globecom 2013 Workshop on Emerging Technologies for LTE-Advanced and Beyond-4G, Atlanta, Georgia, USA, December 2013

[Renzo 2011] M. Di Renzo, H. Haas, and P. M. Grant, "Spatial modulation for multiple-antenna wireless systems: a survey," IEEE Communications Magazine, vol. 49, no. 12, pp. 182–191, December 2011

[Chowdhury 2014] M. Chowdhury, A. Manolakos, A.J. Goldsmith, "CSI is not needed for Optimal Scaling in Multiuser Massive SIMO Systems," Proc. of ISIT. Honolulu, July 2014.

[Armada 2015] A. García Armada, L. Hanzo, "A Non-Coherent Multi-User Large Scale SIMO System Relying on M-ary DPSK", Proc. of ICC, London, June 2015

[Heath 2014] R. W. Heath Jr, "Comparing Massive MIMO and mmWave MIMO", IEEE Communication Theory Workshop, May 2014

[Tanaka 2015] K. Tanaka, A. Agata, «Next-generation optical access networks for C-RAN», en Optical Fibre Communications Conference and Exhibition (OFC), 2015, 2015, pp. 1-3.

[cpri v6.1] "Common Public Radio Interface (CPRI), Specification V6.1 (2014-07-01)", Available online (http://www.cpri.info/downloads/CPRI_v_6_1_2014-07-01.pdf). Gathered on 2015-10-05.

[Morant 2015] M. Morant, A. Macho, R. Llorente, "Optical fronthaul of LTE-advanced MIMO by spatial multiplexing in multicore Fibre", en Optical Fibre Communications Conference and Exhibition (OFC), 2015, 2015, pp. 1-3.

[Llorente 2015] R. Llorente, M. Morant, A. Macho, D. Garcia-Rodriguez, J. L. Corral, "Demonstration of a spatially multiplexed multicore fibre-based next-generation radio-access cellular network", en 2015 17th International Conference on Transparent Optical Networks (ICTON), 2015, pp. 1-4.

[Morioka 2009] T. Morioka, "New generation optical infrastructure technologies: EXAT initiative: towards 2020 and beyond," OECC 2009

[Puttnam 2015] B. J. Puttnam et al., "2.15 Pb/s Transmission Using a 22 Core Homogeneous SingleMode Multi-Core Fibre and Wideband Optical Comb", PDP.3.1, ECOC 2015

[Soma 2015] D. Soma et al., "2.05 Peta-bit/s Super-Nyquist-WDM SDM Transmission Using 9.8-km 6-mode 19-core Fibre in Full C band", PDP.3.2, ECOC 2015

[Kafle 2010] V. P. Kafle and M. Inoue, "HIMALIS: Heterogeneity inclusion and mobility adaptation through locator id separation in new generation network," IEICE Transactions on Communications, vol. E93-B, no. 3, pp. 478–489, 2010.

[rfc6227] T. Li, "Design Goals for Scalable Internet Routing," 2011, http://www.ietf.org/rfc/rfc6227.txt.

[Martinez-Julia 2012] P. Martinez-Julia, A. F. Gomez-Skarmeta, V. P. Kaf l e, and M. Inoue, "Secure and robust framework for id/locator mapping system," IEICE Transactions on Information and Systems, vol. E95-D, No. 1, pp. 108-116, 2012.

[Kafle 2012] V. Kafle, R. Li, D. Inoue, and H. Harai, "An integrated security scheme for id/locator split architecture of future network," in Proceedings of the IEEE International Conference on Communications (ICC) 2012. Washington, DC, USA: IEEE, 2012, pp. 5866–5871.

[mofi] "Mobile Oriented Future Internet (MOFI)," 2011. http://www.mofi.re.kr.

[Jung 2011a] H. JUNG and S. J. KOH, "Hinlo: An id/loc split scheme for mobile oriented future internet," in Proceedings of the Future Network and Mobile Summit 2011, pp. 1–8, International Information Management Corporation, 2011.

[Jung 2011b] H. Jung, M. Gohar, J.-I. Kim, and S. J. Koh, "Distributed mobility control in proxy mobile ipv6 networks," IEICE Transactions on Communications, vol. E94-B, No. 8, pp. 2216–2224, 2011.

[Chan 2011] H. Chan et al., "Problem statement for distributed and dynamic mobility management," internet-draft, IETF, 2011.

[itu] International Telecommunication Union (ITU), "ICT Facts & Figures – the world in 2015", 2015. Available from http://www.itu.int/en/ITU-D/Statistics/Pages/facts/

[ngmn] NGMN Alliance, "5G White Paper", Editors: Rachid El Hattachi, Javan Erfanian, 17 February 2015. Available from http://www.ngmn.org

[metis] METIS, "Requirement analysis and design approaches for 5G air interface", Deliverable D1.1, Doc nr ICT-317669-METIS/D2.1, 29 April 2013, available from https://www.metis2020.com/documents/

[Kovacs 2014] A.-M. Kovacs, "Regulation in Financial Translation: Will the Incentive Auction Increase Mobile-Broadband Competition in Rural America?", May 2014.

[smith] D. Smith, "The Truth about Spectrum Deployment in Rural America", 16 March 2015. Available from http://mobilefuture.org/resources/

[Lin 2014] X. Lin, J.G. Andrews, A. Ghosh and R. Ratasuk, "An Overview of 3GPP Device-to-Device Proximity Services", IEEE Communications Magazine, April 2014, pp. 40-48.

[Day 2008] J. Day, Patterns in network architecture: a return to fundamentals. Prentice Hall, 2008

[Trouva 2011] E. Trouva, E. Grasa, J. Day, I. Matta, L. T. Chitkushev, S. Bunch, M. Ponce de Leon, P. Phelan, and X. Hesselbach-Serra. 2011. In Proceedings of the 9th IFIP TC 6 international conference on Wired/wireless internet communications (WWIC'11).