Valeria Loscri
Luca Chiaraviglio
Anna Maria Vegni *Editors*

The Road towards 6G: Opportunities, Challenges, and Applications

A Comprehensive View of the Enabling Technologies

Springer

The Road Towards 6G: Opportunities, Challenges, and Applications

Valeria Loscri • Luca Chiaraviglio •
Anna Maria Vegni
Editors

The Road Towards 6G: Opportunities, Challenges, and Applications

A Comprehensive View of the Enabling Technologies

Springer

Editors
Valeria Loscri
Inria research centre Lille - Nord Europ
Villeneuve d'Ascq, France

Luca Chiaraviglio
University of Rome Tor Vergata
Rome, Italy

Anna Maria Vegni
Roma Tre University
Rome, Italy

ISBN 978-3-031-42566-0 ISBN 978-3-031-42567-7 (eBook)
https://doi.org/10.1007/978-3-031-42567-7

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2024
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

Preface

Welcome to a journey from 5G to 6G systems. This book has the primary objective to introduce its reader to the main technologies that will mark the transition from 5G to 6G systems, with a focus on the main reasons that lead to this transformation. Without claiming to be exhaustive, this book was the opportunity to identify the key points of a technology that advances galloping, and evolves daily to offer us applications and services in everyday life, which were simply unimaginable a few decades ago. In this book, *The Road Towards 6G: Opportunities, Challenges and Applications, A Comprehensive View of the Enabling Technologies*, you will have the opportunity to go through the expertise and vision led by deep thoughts from experts on different representative technical subjects of the 6G networks. A comparative approach with the 5G network will be the way to present the main rationale beyond this rapid evolution from 5G to 6G as well as the open challenges and new perspectives to improve the user's experience. The authors of this book have provided excellent perspectives on different key aspects for the 6G systems, by providing a clear vision of the existing gaps that need to be filled in the next future. The authors are scientists and engineers coming from academia and industry, with expertise on wireless networks, advanced wireless technologies, cyber security, Machine Learning and Artificial Intelligence. The book is covering the different aspects concerning the use-cases and requirements, how the technology is evolving from 5G to 6G with a focus on the Internet of Things (IoT) paradigm, that has been the key technology characterizing the 5G (Chap. 1). An overall vision of the principles of 6G wireless networks is presented in the book (Chap. 2), and a detailed discussion of some representative technology, such as cell-free Massive MIMO, TeraHertz and Free Space optical, that have the unique capability to enable high data rate services will be considered (Chap. 3). Some specific architectures relying on the reprogrammability concept of the material and the interaction of the signals with the material is another core concept in 6G networks (Chap. 4) as well as the Non-Terrestrial Networks and their interaction with the ground systems. Of course, a primary role on 6G systems is played by Machine Learning and Artificial Intelligence, which will enable an unprecedented development of dynamic systems (Chap. 5). Another crucial point that will take the attention of scientists and industrial community is the cyber security evolution and the new open issues in 6G networks. In particular, the massive use of ML/AI will determine an unprecedented evolution in the effectiveness of attacks, which will become more undetectable (Chap. 6). Finally, the EM exposure in 6G is another hot topic (Chap. 7), which we considered paramount to present in our book.

Villeneuve d'Ascq, France	Valeria Loscri
Rome, Italy	Luca Chiaraviglio
Rome, Italy	Anna Maria Vegni

Contents

The "Transitioning" from 5G to 6G 1
Andrea Detti and Michele Nitti
1 5G Technology and Its Evolution Towards 6G 1
    1.1 Introduction 1
    1.2 5G Architecture 2
    1.3 5G Path Towards 6G 6
    1.4 6G Architectural Proposals and Driving Technologies 7
    1.5 Conclusion 11
2 First Glance on Use-Cases and Requirements of 6G Systems 11
    2.1 Introduction 11
    2.2 Proposed 6G Use-Cases 12
    2.3 Preliminary 6G Performance Requirements 14
    2.4 Conclusions 15
3 Internet of Things Technology 16
    3.1 The Internet of Things as Pervasive and Ubiquitous Technology 16
    3.2 The Evolution of IoT Towards the Internet of Everything Paradigm 21
References 25

Principles of 6G Wireless Networks 27
Qinghe Du, Houbing Song, Xiao Tang, Zixiao Zhao, Chunge Xu,
Yuquan Xiao, and Zhongmin Ma
1 Introduction 27
2 The Gap from 5G to 6G 29
    2.1 Peak Data Rate 30
    2.2 User-Experienced Data Rate 30
    2.3 Spectrum Efficiency 30
    2.4 Area Traffic Capacity and Connection Density 31
    2.5 Mobility 31
    2.6 Latency and Reliability 31
    2.7 Positioning and Sensing Capabilities 32
    2.8 AI Capability 32

2.9 Security, Privacy, and Trustworthiness ... 33
2.10 Sustainability ... 33
3 PHY Layer Design ... 33
3.1 Dig New Signal Spaces ... 34
3.2 Reshape the Propagation Environments ... 35
3.3 Build New Coverage Paradigm ... 36
3.4 Integration of Communications with Sensing ... 37
4 Resources Allocation ... 38
4.1 Distributed AI for Resource Allocation ... 38
4.2 The Aggregation of Heterogeneous Resources ... 40
5 Networking-Related Issues ... 42
5.1 6G Core Network ... 42
5.2 Tightly Coupled Cross-Layer Design with Service Awareness ... 43
5.3 Service-Aware Transmission over Heterogeneous RATs ... 44
5.4 Enabling User Socialization for Cloud Extension ... 45
5.5 Age Control Protocol for Synchronization Services ... 45
References ... 46

6G Wireless Technologies ... 51
Carmen D'Andrea, Josep Miquel Jornet, Arjun Singh, Priyangshu Sen, Zabih Ghassemlooy, Stanislav Zvanovec, Shivani Rajendra Teli, and Asghar Gholami
1 Cell-Free Massive MIMO ... 51
1.1 Network Topology ... 52
1.2 User-Centric Approach and Centralized Processing ... 53
1.3 The Communication Process ... 54
1.4 Performance Assessment ... 55
1.5 Applications of Cell-Free Massive MIMO in 6G ... 57
2 Terahertz Communications for 6G Networks: How Far Are We? ... 58
2.1 Introduction ... 58
2.2 Terahertz Materials, Devices and Testbeds ... 60
2.3 Terahertz Propagation and Channel Modeling ... 65
2.4 Ultrabroadband Communication Techniques ... 70
2.5 Ultradirectional Networking Protocols ... 74
2.6 Policy and Standardization ... 76
2.7 Conclusion ... 77
3 Free Space Optical (FSO) in 6G ... 78
3.1 Introduction ... 78
3.2 Free Space Optical Links ... 80
3.3 Radio over FSO ... 90
3.4 The VLC-OCC Technology for 6G ... 94
3.5 Conclusion ... 101
References ... 101

6G Wireless Architectures 115
Giacomo Oliveri, Francesco Zardi, Arianna Benoni, Marco Salucci, Andrea Massa, Ernestina Cianca, Debashisha Mishra, Evgenii Vinogradov, and Enrico Natalizio
1 EM Static and Reconfigurable Passive Skins Within the Smart ElectroMagnetic Environment Paradigm 115
  1.1 Principles, Opportunities, and Challenges 116
  1.2 Design and Implementations 119
  1.3 Conclusions 123
2 LEO Constellations 124
  2.1 Introduction 124
  2.2 Main Characteristics of Current LEO Constellations 126
  2.3 LEO Satellite Application Scenarios in 5G and beyond 129
  2.4 3GPP Standardization Activities on NTN 131
  2.5 The Issue of Space Sustainability 134
  2.6 Concluding Remarks 134
3 On the Role of Non-terrestrial Networks in 6G 135
  3.1 Non-Terrestrial Radio Access Network and 6G 136
References 149

Artificial Intelligence and Machine Learning in 6G 155
Shashwat Mishra, S. Muhammad Karam, Chung Shue Chen, and Luca Rose
1 Introduction 155
  1.1 Drivers for 6G Development 157
  1.2 Key Performance Indicators and Related Use Cases 158
2 The Role of AI and ML in 6G 161
  2.1 Decreased Computational Complexity 162
  2.2 Improved Robustness and Versatility 162
  2.3 Leveraging Data Traces 163
  2.4 Knowledge Transfer and Adaptability 163
3 AI Enhanced PHY Layer 164
  3.1 Waveform Design 164
  3.2 Antenna Design 165
  3.3 Channel Prediction 166
  3.4 Source and Channel Coding 166
4 AI Supported MAC Layer 167
  4.1 Distributed Grant Free Access 167
  4.2 Radio Resource Allocation 168
  4.3 Network Optimization 169
  4.4 Security 169
5 ML Enhanced OFDM Systems: A Case Study 170
6 Next Generation Networking Paradigms 170
  6.1 Cell-Free Massive MIMO 170
  6.2 Ultra Massive Machine-Type Communication 171
  6.3 Sub-terahertz Communication 171

6.4 Intelligent Transportation Systems 172
6.5 Joint Communication and Sensing 173
6.6 Semantic Communication 173
6.7 ML-Based Standardization 174
7 Machine Learning Frameworks for 6G 175
7.1 Federated Leaning 175
7.2 Graph Neural Networks 176
7.3 Q-Learning 176
7.4 Adversarial Learning 177
7.5 Auto-encoders and Transformers 178
8 Streamlining the Deployment of ML Models 178
9 Some Useful Tools and Resources 180
References 181

Cyber Security in 6G Networks 185
Valeria Loscri
1 Transitioning of Security Aspects from 5G to 6G 185
1.1 A Glimpse to the Differences Between 5G and 6G Technologies 185
1.2 Security Changes from 5G and 6G 187
2 Cyber-Security Vulnerabilities, New Risks and Threats in 6G 188
2.1 Open Source Software: The New Evil for 6G Security? 188
2.2 The Double Nature of AI/ML in 6G Security 189
3 Physical Layer Security—PLS 190
3.1 PLS Categories 191
4 Discussion 194
References 194

6G EMF Exposure 197
Jack T. Rowley
1 What Is New with 6G from an EMF Point of View 197
2 Frequency 198
2.1 Scientific Evidence Up to 300 GHz 199
2.2 Reviews of 5G mmWave RF-EMF Exposure 200
2.3 Scientific Evidence Above 300 GHz 201
3 Power 202
3.1 Mobile Devices 202
3.2 Internet of Things (IoT) Devices 203
3.3 Mobile Networks 203
3.4 5G Network RF-EMF Exposure Levels 205
4 6G Technology Characteristics 205
4.1 Influence of Beamforming on 6G Network RF-EMF Exposure Levels 206
4.2 Influence of 6G Network Densification on RF-EMF Exposure Levels 206
4.3 Influence of New Use Cases and New Device Types on 6G Exposures 207

5 Technology Independent Exposure Limits 207
  5.1 Exposure Limits Up to 300 GHz 208
  5.2 Controversy About Biological Effects from Low-Level Exposures . 209
  5.3 RF-EMF Protection for Flora and Fauna 210
  5.4 Exposure Limits Above 300 GHz 211
6 Assessing RF-EMF Compliance 212
  6.1 Network Assessment Methods 213
  6.2 Device Assessment Methods 213
7 Information and Misinformation 214
  7.1 COVID-19 and 5G Misinformation 215
8 Preparing for the Future of 6G and RF-EMF 215
9 Conclusion 216
References 217

The "Transitioning" from 5G to 6G

Andrea Detti and Michele Nitti

1 5G Technology and Its Evolution Towards 6G

Andrea Detti

Abstract 5G has introduced new paradigms of mobile network design and deployment and delivery of new services. Starting with 3GPP Release 15, 5G has evolved with Releases 16 and 17 and now fully supports all the envisioned service scenarios and, consequently, many vertical sectors and industries. The path to 5G-Advanced has just begun with Release 18. Meanwhile, industry and research have started to think about what 6G will be, and a common vision of 6G providing a cyber-physical continuum experience exists. This chapter outlines the path from 5G to 6G. We present the 5G architecture, discuss its progress, release by release, and finally arrive at 6G, presenting initial architectural views and promising leading technologies.

1.1 Introduction

5G has introduced a new paradigm in the design of mobile networks and the deployment and delivery of new services [14]. Compared to previous generations, there has

---

A. Detti (✉)
University of Rome "Tor Vergata", Rome, Italy
e-mail: andrea.detti@uniroma2.it

M. Nitti
Department of Electrical and Electronic Engineering (DIEE), University of Cagliari, Cagliari, Italy

National Telecommunication Inter University Consortium, Research Unit of Cagliari, Cagliari, Italy
e-mail: michele.nitti@unica.it

V. Loscri et al. (eds.), *The Road towards 6G: Opportunities, Challenges, and Applications*, https://doi.org/10.1007/978-3-031-42567-7_1

been a shift from a horizontal service delivery model, whereby all customers use the same services, to a vertical model, whereby services are customized for users and industries. The need to specialize services for different customers with heterogeneous requirements requires a deployment flexibility typical of software/cloud environments. Indeed, the 5G architecture is a blend of Cloud, Software Defined Network (SDN), and Network Function Virtualization (NFV) technologies.

Although 5G is not yet 100% operational, 5G standardization is still in progress, and the 5G business for the industrial sector has not yet taken off, researchers and companies are already wondering what 6G will be, considering that following the G+ rule every ten years, 6G is scheduled for 2030. Accordingly, in the following sections, we will first provide an overview of the 5G architecture and its progress release by release and then present some initial 6G architectural ideas and promising driving technologies.

1.2 5G Architecture

Figure 1 shows the architecture of the 5G network [4] whose primary goal is to provide mobile terminals with data pipes called PDU sessions that interconnect them with external data networks, such as the Internet or the IP Multimedia Subsystem (IMS) for phone calls [15]. From a mobile terminal to an external network, a PDU session consists of (i) a radio bearer on the air interface between the mobile terminal and its 5G base station (gNB) and (ii) a sequence of network tunnels, usually based on the GPRS Tunnelling Protocol (GTP), starting from the gNB and traversing one or more User Plane Functions (UPFs) of the core network until they reach the external network. The most challenging task of the mobile network is to maintain this end-to-end connectivity continuously, regardless of

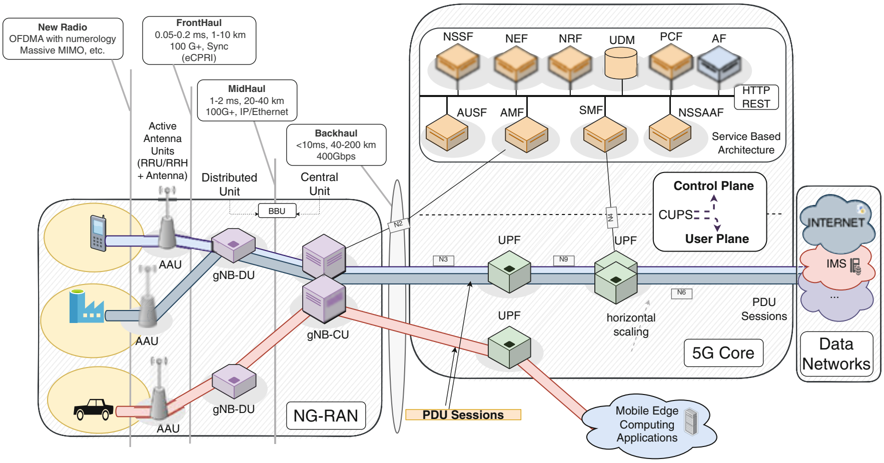


Fig. 1 5G architecture

terminal mobility, and through efficient mechanisms to improve network capacity, user data rates, energy efficiency, quality of service, etc.

As shown in Fig. 1 the architecture of the 5G network consists of two sections, namely the New Generation Radio Access Network (NG-RAN) and the 5G-Core network (5GC). A PDU session flows through these two network sections by passing across different network functions. In the following parts, we shed light on NG-RAN and 5G-Core.

1.2.1 NG-RAN

The NG-RAN comprises 5G base stations, and each base station is referred to as g-NodeB (gNB). These base stations are radio equipment designed to serve cells within a specific area, enabling users to access 5G services [1].

The 5G radio interface, called New Radio (NR), is based on OFDMA (Orthogonal Frequency-Division Multiple Access) and can be considered an evolution of 4G [2]. Compared with 4G, 5G NR can achieve a higher bit-rate, lower latency, and operate with wider channel bandwidth (e.g., 400 MHz) and in a broader frequency region, including mmWaves (26–60 GHz). The 5G NR interface is extremely flexible in adapting to service needs. Compared to 4G, 5G can customize the communication services in several aspects, ranging from physical layer configurations, such as the spacing between OFDMA subcarriers (aka numerology), to upper layer configurations, such as the time elapsing between data and related acknowledges.

A gNB can be a single device that integrates the antenna and the electronic/computational equipment that implements the functionality of the protocol stack. However, gNB is usually split into two or three units that implement different layers of the radio interface protocol stack. The protocol stack layers can be classified into transport, user, and control plane protocols. From the bottom up, the transport protocols are as follows.

- Physical layer (PHY) handles OFDMA radio transmissions and receptions of I/Q symbols, including the management of multi antennas, channel coding, and synchronization;
- Medium Access Control layer (MAC) manages the radio resource allocation (scheduling) to transmit MAC frames, channel priorities, data multiplexing, information reporting, and hybrid automatic repeat request (H-ARQ).
- Radio Link Control (RLC) improves the reliability of communication service offered by the MAC layer according to the need of the different data flows.
- Packet Data Convergence Protocol (PDCP) implements security and header compression.
- Service Data Adaptation Protocol (SDAP) manages the quality of service (QoS) among different flows belonging to the same PDU session.

The transport protocols deliver higher layer information over the radio interface, namely user-plane data, i.e., IP packets of user applications, and control-plane data, i.e., Radio Resource Control (RRC) signaling messages.

Figure 1 shows a gNB split into three units: the Active Antenna Unit (AAU), the gNB Distributed Unit (gNB-DU), and the gNB Central Unit (gNB-CU). This functional split results from different trade-offs, including the feasibility of deployment, the simplification of management, and the potential optimization capability.

An AAU is an active antenna connected with an optical fiber to the gNB-DU through a network section called *fronthaul*, which usually implements the CPRI or eCPRI protocols.[1] An AAU is made of a passive antenna combined with a device called the Remote Radio Unit (RRU) or Remote Radio Head (RRH). The RRU implements low-level Physical (PHY) layer functions of the air interface, converting digital symbols from the gNB-DU to analog signals for the antenna and vice versa. Integrating the RRU with its passive antenna simplifies implementation, especially when using massive MIMO (Multiple-Input Multiple-Output) antennas with many ports. However, the RRU and the passive antenna can still be different devices connected by a coaxial cable per port, whether the MIMO technology is not so dense (e.g., only 4/8/16 antenna ports).

A gNB-DU implements the high-PHY, MAC, and RLC protocols of the air interface. On one side, it is connected with the RRU via a CPRI/eCPRI fronthaul optical link that is extremely demanding in terms of bandwidth and jitter. In fact, this link transports baseband time samples of waveforms (CPRI) or symbols (eCPRI) that the AAU transmits/receives over the air. Due to these high-performance requirements, the gNB-DU is usually placed close to the AAU, such as at the base of the antenna pole.

On the other side, a gNB-DU exchanges PDCP packets with a gNB-CU through an IP network, called *midhaul*, consisting of links that can be wired or radio. The supporting radio technology may be proprietary, or 5G itself offers standard support for midhaul/backhaul radio links in a configuration called Integrated Access And Backhaul (IAB). The communication requirements of midhaul links are slightly more tolerant in terms of jitter, and it is possible to move the gNB-CU away from the gNB-DU, such as to a central data center.

A gNB-CU manages the last transport protocols of the radio interface, namely PDCP, and SDAP, and implements the RRC control-plane functionality. Toward the 5G Core, the gNB-CU is connected through an IP network, named *backhaul*, to one or more UPFs and Access and Mobility Management Functions (AMFs). The advantages of having a centralized gNB-CU are many, including (i) the ability to be virtualized in a data center, making the overall RAN more future-proof and less costly, and (ii) the ability to simply implement coordinated control mechanisms of different cells, e.g., to mitigate interference, better manage load balancing and handovers by having a multi-cell view of resources in use. Note that the gNB-DU and the gNB-CU can be integrated into a single device called a Baseband Unit (BBU).

---

[1] http://www.cpri.info/.

1.2.2 5G-Core

With respect to the previous generation, the architecture of the 5G Core Network is innovative and unique in many ways. There are no more nodes communicating with each other through point-to-point protocols but Network Functions (NFs) that expose services to interact with each other. Having functions rather than nodes paves the way to possible cost-effective softwarization and virtualization of them, making 5G Core a cloud-native architecture.

Following the SDN paradigm, the 5G Core has a clear Control and User Plane Separation (CUPS). In the user plane, there are one or more User Plane Functions (UPFs), mainly carrying out packet forwarding between the different tunnels forming the PDU sessions. All other network functions belong to the control plane and compose a service-based architecture (SBA) because related functions expose “services” by using HTTP/2 REST APIs described through OpenAPI.[2]

The ability to softwarize and virtualize network functions has enabled one of the most innovative 5G features, called *network slicing*. Network slices are 5G virtual networks that support different classes of applications with particular characteristics and requirements. Different network slices usually use different instances of NFs, with different configurations and transport resources. By leveraging virtualization, the same 5G network infrastructure can support multiple slices, each dedicated to specific classes of customers/applications, thus offering greater flexibility, efficient use of network resources, and increased opportunities for enterprises, which in turn can use differentiated services with specific speed, latency, and reliability requirements.

Let us conclude this section by briefly describing the main roles of 5G NFs.

- The User Plane Function (UPF) handles the forwarding of data belonging to PDU session tunnels (tunnel switching) and related control services, such as anchoring for handover, Quality of Service (QoS), and traffic policy enforcement.
- The Session Management Function (SMF) is the control part of a PDU session. It configures tunnel switching on UPF, allocates IP addresses, and configures traffic steering (e.g., towards a third party or an edge cloud).
- The Access and Mobility Management Function (AMF) handles all 5G Core signaling from and going to user equipment (UE) and NG-RAN.
- The authentication server function (AUSF) is the inner component of the core network that supports authentication for 3GPP and non-3GPP accesses.
- The unified data management (UDM) function can be considered a repository for UE-related information, such as credentials, identifiers, AMF details, and SMF assignments for the current session. The underlying idea of the UDM is to create, wherever possible, a central database for UE configuration information so that NFs can be designed as *stateless* services, improving architectural agility and scale-out.

---

[2] https://forge.3gpp.org/rep/all/5G_APIs.

- The policy control function (PCF) is a unified entity providing policy rules (QoS, filtering, charging, etc.) to other control plane functions, such as SMF.
- The network slice selection function (NSSF) selects the set of network slice instances serving the UE, along with the best AMF for that purpose. Related access control to a slice is managed by the Network Slice-Specific Authentication and Authorization Function (NSSAAF)
- The network exposure function (NEF) exposes the capabilities of network and network/UE events for third-party, application functions.
- The network repository function (NRF) discovers network function instances. When it receives an NF discovery request from a NF instance, it provides the discovered NF instances.
- The application function (AF) resembles an application server that can interact with the other control-plane NFs. AFs can exist for different application services and can be owned by the network operator or by trusted third parties.

1.3 5G Path Towards 6G

The development of 5G technology has been and is quite dynamic, adding new features to the network, release after release (Fig. 2). In 2018, Release 15 specified the new radio (NR) 5G interface, 5G core, stand-alone (SA) and non-stand-alone (NSA) deployment options and provided support for Enhanced Mobile Broadband (eMBB) and basic Ultra Reliable Low Latency Communications (URLLC) services. Thanks to NSA that allows the combination of an NG-RAN with 4G cores already available, operators have been able to quickly deliver 5G access to customers, before deploying a complete standalone 5G (Fig. 1).

Release 16 (5G-Evolution) arrived in 2020 and provided complete support for all envisaged service scenarios, namely eMBB, URLLC, and massive IoT (mIoT), also known as Massive Machine-Type Communications (mMTC), to make the 3GPP 5G System (5GS) a communication-enabling platform suitable for a wide range of industries (“verticals”). Moreover, Release 16 also improved network slicing and features designed for private 5G networks, thus more focusing on enterprise

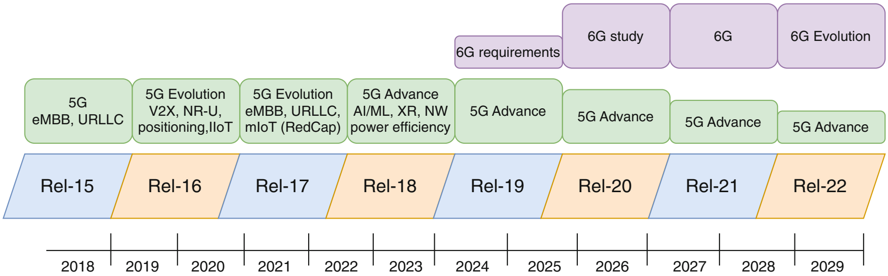


Fig. 2 Timeline of 3GPP releases

and business applications such as V2X, Industrial IoT (IIoT). Release 16 also made significant improvements to positioning services, such as integrating the Metropolitan Beacon System (MBS) [3].

Release 17 (5G-Evolution) is the most recent standardization effort on 5G at the time of preparing this contribution (early 2023). The main focus is on the performance improvement of service scenarios (eMBB,URLLC,mIoT), especially with respect to Industrial IoT, with the support of Reduced Capability devices (RedCap), Railways, Mission Critical (MC) and priority service, Drone/UAS/UAV/EAV, etc. The Radio interface and the Access Network have been significantly improved too, e.g., NR now supports 1024 QAM for downlink. Release 17 has also introduced system optimization for edge computing, cutting, self-organizing/ autonomous network (SON), and energy efficiency [5].

Release 18 has just started the evolution journey of 5G-Advanced, focusing on performance and more on the sustainability of 5G networks and data-driven and Artificial Intelligence (AI) powered control [31]. The work on Release 18 will surely lay a strong foundation for 6G design, whose standardization activities are expected to begin in 2024/2025.

1.4 6G Architectural Proposals and Driving Technologies

6G use cases require technological improvements that have highlighted the shortcomings of the current 5G architecture that 6G is expected to address. 6G should provide support for *massive digital twins*, with objects highly synchronized with their digital representation; *immersive communications*, through which people extend their sense into the digital domain; *connected intelligence*, whereby the network becomes a trusted domain in which to execute (third-party) AI capabilities; *cognition*, integrating sensor data to discover what humans want, their intentions and their desires [10].

To meet these challenges, 6G should offer reliable AI services and use AI for intelligent connected management and control functions. 6G should also introduce a certain level of programmability of the user equipment to allow customization of the air interface without the need to standardize too many details for each use case and/or to accelerate the time to market for new features. Integrated sensing and communication will allow massive antennas and higher 6G frequencies to be exploited for communications and sensing the environment, e.g., THz imaging and localization with centimeter accuracy. To ensure the sustainability of ICT, 6G should reduce its energy footprint and suppress the increase in energy consumption with the increase in traffic. Finally, 6G should be inclusive and affordable for people and operators. Coverage should be global, with greater integration of satellites to include remote areas, and a significant increase in the number of picocells is expected, which is only possible if the cost of these base stations is more affordable for operators. Low-cost devices should be available when only reduced capabilities are needed (RedCap).

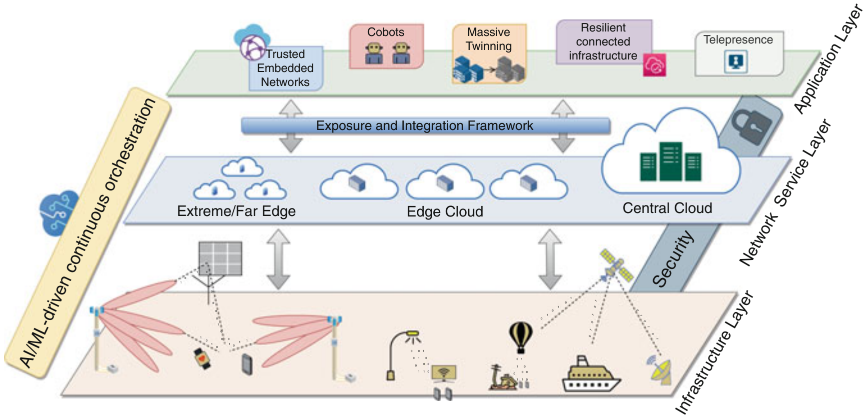


Fig. 3 High-level view of 6G architecture (adapted from [24])

1.4.1 High-Level View of 6G Architecture

Although we are in a very preliminary stage of envisioning what 6G will be, recently (December 2022), the 5G-PPP initiative has published a white paper on the landscape of 6G architecture [21]. This paper is mainly derived from the works of the Hexa-x EU project.[3]

Figure 3 shows a high-level view of the 6G architecture envisaged by 5G-PPP/Hexa-x. There are three layers: Infrastructure, Network Service, and Application.

The infrastructure layer consists of RANs, Core Networks, transport networks, edge/cloud data centers, sensors, and so on, and provides the physical resources for network services and application layer elements. 6G RAN will dramatically improve performance in terms of latency, reliability, availability, data rate, capacity, coverage, energy efficiency, and location accuracy. In addition, the Infrastructure layer should incorporate several (sub)networks into a 6G network of networks.

The Network Services layer is cloud and microservices-based, with 6G network functions and (micro)services spread from the central cloud to the cloud of the extreme edges, i.e., to devices beyond the RAN, such as personal devices (smartphones, laptops, etc.) and heterogeneous IoT devices (wearables, sensor networks, connected cars, connected appliances, etc.). Following the 5G Network Exposure Functions (NEF) idea, the 6G network services layer offers a powerful exposure framework and integration fabric that enables interoperability and seamless networking and computing across different domains (e.g., different operators) and with applications.

---

[3] https://hexa-x.eu/.

Network management and orchestration should implement an AI/Machine-Learning (ML) driven continuous orchestration framework that spans devices, edge, and cloud. This requires a 6G architecture with also a service-based RAN (in 5G only Core is service-based), enabling intelligent control algorithms to leverage a uniform Core/RAN web API. Security and privacy mechanisms will permeate the overall architecture, making 6G a trusted environment for communications and applications. Distributed Ledger platforms could be used to establish “distributed trus”, such as among domains and applications.

1.4.2 Driving Technologies

6G architecture will be driven by a set of emerging technologies and architectural paradigms that will address the challenges of performance and complexity. We present hereafter in a concise way a brief summary about 6G technologies/architectural paradigms, while we refer the reader to chapter 3, 4 and 5 of the book for an in-depth presentation.

1.4.2.1 Artificial Intelligence and Machine Learning

6G will be a AI-native network that will use AI/ML to automatically adapt network functions to new environments for which they were not originally planned. 6G is expected to have the capability of a self-organizing network (SON), since the number of antennas, devices, cells, slices, etc. will be large and dynamic enough to make human management difficult, requiring model-free automatic approaches such as Deep Learning and Deep Reinforcement Learning [18]. In addition to using AI for its control, 6G will offer AI-as-a-Service (AIaaS) to support cross-domains distributed AI services, such as federated learning and explainable AI, making AI engines as close to applications as possible.

1.4.2.2 THz Bands

6G air interface will be extremely flexible to cope with heterogeneous and dense environments, while providing Gbit/s performance per user, Tbit/s per cell, with submillisecond air latency [41]. Several radio approaches are being considered to achieve this capability, including the use of THz bands, which allow extremely wideband channels with a bandwidth of tens of GHz, and can potentially provide a high-precision positioning capability, too. However, achieving stable THz communications in practice is by no means an easy task because of severe path loss and atmospheric absorption, the complexity of the RF front-end, and the possible need to design a new suitable waveform in addition to OFDMA.

1.4.2.3 Novel Radio Technologies

Cell-free massive MIMO is being considered to cope with the difficult propagation characteristics of the THz band and to improve the capacity of the 6G system in dense and heterogeneous deployments. It provides that UEs can be served by several base stations controlled by one or several central units (CUs), thus offering macro diversity in MIMO systems, the possibility of better controlling inter-cell interference, and the absence of cell and cell boundaries [23]. Moreover, Reconfigurable Intelligent Surfaces (RISs) is considered for the control of radio signals. A reconfigurable intelligent surface (RIS) is a two-dimensional surface of engineered material whose properties are reconfigurable rather than static [11]. The scattering, absorption, reflection, and diffraction properties of a RIS can be thereby changed with time and controlled by software through a central unit to optimize signal propagation. Finally, Free-Space Optical (FSO) communications could provide an excellent opportunity to develop ultrafast data links that can be applied in a variety of 6G applications, including wireless backhauls for cellular systems [27].

1.4.2.4 Satellite and Aerial Coverage

6G could address the global coverage challenge by a combination of Terrestrial Networks (TN) and Non Terrestrial Networks (NTN). NTNs can likely provide a lower capacity than terrestrial networks, thus an NTN coverage is suitable for rural areas, including oceans, with low or very low population density. NTNs can use Low Earth Orbit (LEO) and Medium Earth Orbit (MEO) satellites, but it is necessary to find solutions that minimize the number of handovers and the signalling needs for mobility robustness as well as to cope with Doppler shift due to the speed of satellites.

Starting from the ground and before reaching the altitude of the satellite, lower-altitude UAV and higher-altitude platform systems (HAPS) can be deployed to augment ground-level communications with flying networks. These networks can provide coverage extension, emergency response, capacity enhancement, massive IoT data collection, backhauling solutions, etc. The HAPS are various types of platforms that float in the stratosphere, such as balloons or aircrafts powered by renewable energy sources, and are considered a trade-off between the high coverage of satellites and the relatively low latency of UAVs.

1.4.2.5 Localization and Sensing

As was the case in 5G, the technologies explored in 6G networks open up the possibility of using the mobile communication system itself for 3D location and sensing. However, unlike 5G, 6G joint communications and sensing also means the possibility of integrating additional sensing devices (e.g., radar-based systems) into the network and sharing available resources in terms of time, frequency, available antennas, transmission power, processing capacity, and so on. This sharing requires the development of new trade-offs between sensing and communication.

By integrating communications and sensing, sensing/location services can be offered to applications, on the one hand, and these sensing capabilities can be used to improve the performance of the 6G network itself, on the other hand, providing optimization inputs for various network control activities, such as beam steering.

The accuracy of 6G sensing requires careful integration of data from multiple access technologies (Received Signal Strength, time-difference-of-arrival, angle-of-arrival, etc.) and their processing into neighborhood clouds using AI/ML algorithms.

1.5 Conclusion

Shaping 6G will take many years, as we have seen with previous generations. We have presented the direction of the evolution of mobile architectures from the 5G state-of-the-art, through its improved and advanced versions, to 6G, which represents a vision of mobile service in 2030. Our vision will, of course, be updated as we proceed with research on 6G in the future.

2 First Glance on Use-Cases and Requirements of 6G Systems

Andrea Detti

Abstract 6G is the next horizon for cellular networks, bringing a hyper-connected global experience for applications that will provide a cyber-physical continuum experience anywhere, anytime. We are on the eve of 6G, so there is no global consensus on the use cases and performance requirements that will drive network development. However, something can be glimpsed, some consensus is already there, and this chapter lays out some visions, currently presented by various works of literature, forums, and companies.

2.1 Introduction

6G will provide an extreme connectivity experience for humans and everything by seamlessly integrating the physical, digital and human worlds. Compared to 5G, the bar of requirements and performance will be raised again to support cutting-edge network applications such as extended reality (XR), mobile holograms and massive digital twins, etc. Today in 2023, we are at the beginning of the 6G research era, and different groups of companies and researchers are providing different visions, use cases, and performance requirements. The objective of this short chapter is to provide an initial analysis of some use cases and key performance indicators (KPIs) of 6G, based on current literature works and perspectives of active projects.

2.2 *Proposed 6G Use-Cases*

In line with previous generations of mobile systems, the International Telecommunication Union (ITU) is taking the lead in preparing for the development of International Mobile Telecommunications (IMT) services for 2030 and beyond, commonly referred to as 6G. The ITU is currently outlining its vision for 6G and defining the initial specifications, which will then be followed by standardization efforts by the 3GPP and early experimental developments [45, 47].

As shown in Fig. 4, the concept behind 6G will be flexibility at all levels: from frequency management (including bands in the THz and visible light ranges) to the use of heterogeneous radio transmission technologies, from the seamless coexistence of heterogeneous networks to the full integration of sensors into communication infrastructures. Artificial intelligence will have to be pervasive in all network systems, sensors, and applications so that the network is fully cognitive and able to interpret the operational context and observed reality. All infrastructure components will have to minimize energy consumption. Security will have to be ensured already at the physical layer with a natively multi-domain security-oriented architecture.

In addition to ITU, another prominent initiative is the European Hexa-X project,[4] which aims to develop the foundation and contribute to the industry consensus that will lead to 6G. As shown in Fig. 5, their vision consists of a 6G that creates

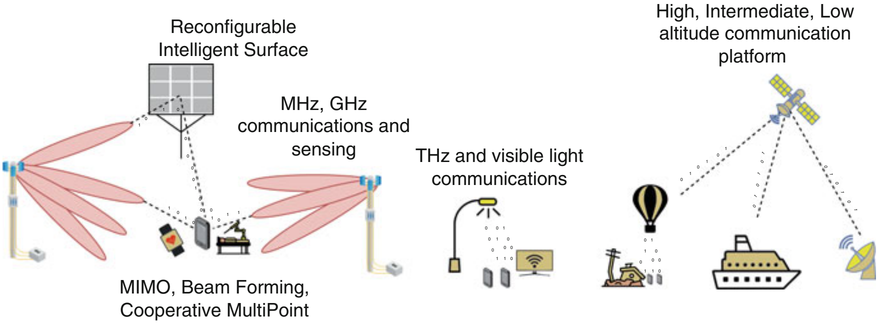


Fig. 4 The integration of many different sensor hardware with the heterogeneous communications networks under 6G systems

---

[4] https://hexa-x.eu/.

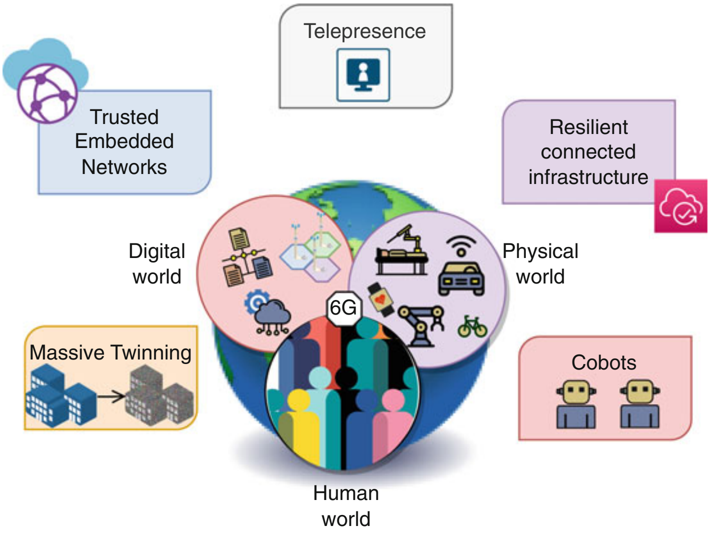


Fig. 5 Hexa-X 6G vision (adapted from [24])

a cyber-physical continuum by coupling and enhancing the interactions between three worlds: a human world of senses, bodies, intelligence, and human values; a digital world of information, communication, and computation; and a physical world of objects and organisms [24, 25, 43]. The key values that should influence the capabilities and requirements of 6G are: Sustainability of the 6G with regard to environmental, social and economic aspects; Inclusion, for a global 6G available to everyone and everywhere; and Trustworthiness, for a 6G that is a backbone of society.

Building on this vision, the Hexa-X project proposed 28 use cases requiring extreme network performance (latency, bandwidth, traffic volumes, etc.) beyond the current capabilities of 5G. These use cases were grouped into 6 families of applications:

- Enabling sustainability—applications that contribute to the sustainability of the environment and society.
- Massive twinning—applications that support secure, real-time digital representation of environments composed of a huge number of entities.
- Telepresence—applications that consist of being present and interacting at any time and in any place (real and virtual worlds), using multiple senses and encompassing human-to-human interactions, physical and digital things (e.g., metaverse).
- Robots to cobots—applications for robots that collaborate (cobots) with each other and humans to perform complex tasks.

- Trusted embedded networks—communication services for providing private subnets (body area network, ad-hoc network, etc.), possibly integrated with a cellular network, that require a high level of reliability.
- Hyperconnected resilient network infrastructures—use cases involving subnets or networks of networks (i.e., tens of thousands of networks in the same geographic area), requiring a high level of resilience.

From the vendors' perspective, Huawei envisions a 6G that will continue the transformation from connected people and things to connected intelligence [26]. Accordingly, their use cases extend the 5G scenarios toward the *extreme* cases of enhanced Mobile Broadband (eMBB+), Ultra Reliable Low Latency Communications (URLLC+), and massive Machine Type Communications (mMTC+) and add two vertices to the resulting "pyramid" of scenarios, namely: sensing and artificial intelligence (AI) that will emerge in 6G. The first involves a mutual integration of communications and sensing technologies whereby networks are used for sensing (e.g., indoor localization), and sensing technologies (e.g., Lindar, imaging) are used to help control the network. The last AI scenario aims to use AI to control the network and use the network to support AI-as-a-Service (AIaaS) functionality, i.e., the network for AI.

2.3 Preliminary 6G Performance Requirements

The use cases presented call for a significant improvement in the performance that the network is expected to deliver, which is mentioned in many papers/forums related to 6G [26, 41, 47]. Currently, there still needs to be a complete consensus on these requirements. Accordingly, we have selected some papers/presentations and reported their expected KPIs in Table 1. Note many KPIs of 6G will be continued from those of 5G but opportunely increased in their target value. In addition, it is expected that KPIs not yet standardized in 5G will emerge in the course of 6G developments that can be called "true" 6G KPIs [37].

To realize advanced multimedia services such as immersive XR, such as mobile holographic communications, and digital twin, a common estimate for 6G peak data rate is 1 Tbps, about 20 times that of 5G (20 Gbps) [29]. In addition to the peak rate, which is the maximum throughput achievable in a perfect environment, the user-experienced data rate that is achieved even in a harsh environment is introduced (e.g., cell edge). The target speeds of 6G are set to 1 Gbps, which is about 10 times faster than in 5G (100 Mbps).

Area traffic capacity in terms of bitrate and device density also needs to increase on the order of 10-1000 times over 5G.

To provide the ultimate experience of delay-sensitive real-time applications, latency-related performance needs to improve significantly, and there is consensus on an over-the-air latency target of less than ms and jitter in the order of mus, i.e., 10 times less than 5G. There is also a consensus view on reliability increasing to seven

Table 1 6G key performance indicators (KPI)

	Target range		
KPI	IMT-2030(6G) Promotion group [47]	Samsung [41]	Huawei [26]
Peak data rate	0.1–1 Tbps	1 Tbps	1 Tbps
User experienced data rate	1–10 Gbps	1 Gbps	10–100 Gbps
Area traffic capacity	0.1–10 Gbps/m^2	N/A	1000x w.r.t. 5G
Connection density x km^2	$10^7 - 10^8$	10^7	10^7
Air latency	0.1–1 ms	0.1 ms	0.1 ms
Jitter	μs-level	N/A	0.1 μs
Reliability	99.99999%	99.99999%[a]	99.99999%
Mobility	1000 km/h	> 500 km/h	N/A
Positioning accuracy	cm-level	N/A	50/1 cm outdoor/indoor
Spectral efficiency	1.5–3.5x w.r.t 5G	2x w.r.t 5G	N/A
Energy efficiency	20x w.r.t. 5G	2x w.r.t. 5G	100x w.r.t. 5G
Frequency bands	5G and THz	5G and THz	5G and THz

[a] Computed as 1—packet error rate.

nines, i.e., 100 times that of 5G. As for mobility, it is expected to support higher speeds than 5G, that is, more than 500 km/h.

6G raises the position accuracy requirement to 1 cm compared to the best 10 cm achieved with 5G (3GPP Rel. 16, frequency range 2, In Door Open Office, Time Difference of Arrival). This comes from the many 6G applications and services such as extended reality (XR), holographic communications, industrial IoT applications, ultra-precision manufacturing, and so on, which need high position accuracy, even indoors, where GNSS and A-GNSS-based solutions do not work well. In addition, highly precise localization can be used to optimize communications (modulation and coding scheme, beamforming, etc.) for different mobile terminals to achieve maximum cell throughput, maximum capacity, and lowest latency [9].

In addition, to support this high speed, the 6G frequency ranges considered include 7–20 GHz as the essential range for most wide-area 6G use cases, the upper mmWave 100–300 GHz range with a complementary role serving small-area niche scenarios with a very high bitrate, and also the THz bands (0.3–1 THz) for very small service areas at extreme bitrates. However, there is no broad consensus considering that THz communications suffer from propagation problems that are even more acute than those faced by the mmWave spectrum bands [43].

2.4 Conclusions

6G is a 2030 technology aimed at rather futuristic applications that create a continuum of real and digital worlds in which the inhabitants are humans, robots, and autonomous objects. The level of performance that 6G networks and network systems are expected to deliver at a worldwide level is extremely challenging, taking in mind that even 5G has not (yet) realized its fully potential. Meeting the ambitious goals of 6G will involve addressing numerous challenges, not only in communications, but also in areas such as software development, computing, and human-machine interfaces. Furthermore, gaining social acceptance for the use of virtual services in critical applications, such as healthcare and autonomous driving, will require significant effort.

3 Internet of Things Technology

Michele Nitti

3.1 The Internet of Things as Pervasive and Ubiquitous Technology

The Internet of Things (IoT) has become a reality with billions of devices able to send key information about the physical world and implementing simple actions, which leads to the paradigm of the anytime and anyplace connectivity for anything [32]. The massive amount of data flowing through the IoT has pushed forward the development of new applications in several domains, such as the management of industrial production plants, the logistics and transport supply chain, the e-health, the smart building, just to cite a few. Moreover, even inside these domains, there are different levels of complexity from simple monitoring from sensors, which enable alerts and notifications of changes to control of product functions and personalization of the user experience, to optimization that enables to enhance product performance and allows predictive diagnostics, service and repair to autonomous product operation with self-coordination of operation with other products and systems.

Given the diversity of applications and uses for the IoT, capturing the essence of this paradigm in one sentence is nearly impossible. Several definitions can be found in both scientific and non-scientific papers, which try to describe what the Internet of Things actually is. Different definitions for IoT have been presented, roughly represented by different keywords, as depicted in Fig. 6.

Aside from the virtual and digital words, that would require more attention, **Internet** and **objects** are among the keywords and could not be otherwise. Despite all its shortcomings as a best-effort network, the Internet represents the link between humans and the collection of available multimedia services and thus serves as the basis for the IoT [7]. In this regard, the term Internet in IoT indicates that the

Fig. 6 Word cloud for the Internet of Things

objects (or at least their services or data) can be consulted and processed by other applications through the existing internet infrastructure.

Another tricky definition lies in what the term object encompasses and its relations with the term things. From [39], an object for the IoT is “any object not precisely identified, which can be solid or intangible, static or mobile, living or inanimate” or in other words: anything. But the *Things* of the IoT are not simple objects, but smart objects which are digitally enhanced with a combination of one or more sensors/actuators, computation and communication interfaces. In this sense, the idea of the IoT as an enabler of connectivity for anything and not only for anyone comes exactly from this conception. A Thing for the IoT is then the union of an entity of the real world enhanced with an IoT device, i.e. a digital device connected to the Internet.

The same IoT concept evolved through the years. The phrase Internet of Things started as the title of a presentation made by Kevin Ashton, executive director of the Auto-ID Labs, at Procter&Gamble in 1999 [33]. Ahston was looking for a term to refer to a global chain of goods identified through Radio Frequency IDentification (RFID) to optimize logistics and supply chains. Since then, the vision of IoT has been further developed and extended and RFID are now not the only Things in the IoT. Moreover, there is a number of similar technologies that are very closely related to IoT, which include Machine-to-Machine (M2M), Cyber-Physical-Systems (CPS) and the Web of Things (WoT). These technologies have no clear differences between them and the term IoT is widely used to cover all of them.

M2M, as defined by ETSI [42], refers to the communication between two or more entities that do not necessarily need any direct human intervention. Hence, M2M is part of the IoT, since IoT brings people, processes, data and things together and comprises a broader range of interactions, not limited to M2M but including also Person-to-Machine (P2M) and Person-to-Person (P2P) interactions.

CPS has the highest degree of overlap with IoT as it represents a system of collaborating computational elements controlling physical entities. The main difference lies from a networking or communication point of view as the IoT targets a broader view of connecting objects in a global aspect whereas CPS targets the coordination of networked objects to achieve a specific goal [35].

Finally, the WoT is concerned with only the highest OSI layer (7), which handles applications, services, and data to enable access and control over IoT resources mainly using web technologies. However, IoT doesn't advocate a single Application-level protocol and usually focuses on the lower layers of the OSI stack [22].

Common features of all these technologies are the pervasiveness and the ubiquity of IoT technology in our everyday life. Meanwhile, different IoT platforms provide the programmatic tools necessary to integrate a rich set of functionalities via numerous Application Programming Interfaces (APIs). Due to the high diversity of devices, applications, and interests in IoT, dozens of these IoT platforms are active today [8], so there are many vertical solutions that implement their own solution for the IoT. This means that the problem of heterogeneity is shifted from objects to architecture. Currently, there is not only one Internet of Things but many Intranet of Things so a shift from a vertical to a horizontal platform is required as depicted in Fig. 7.

Vertical platforms build a separate platform for each application domain so that services are not shared from one platform to another (vertical and fragmented ecosystem). This leads to building standalone platforms, which require high maintenance costs, time and resources; moreover, for converged services, i.e. for applications requiring data from different services, platform integration issues can arise: whenever there is the need to create new services that make use of data from

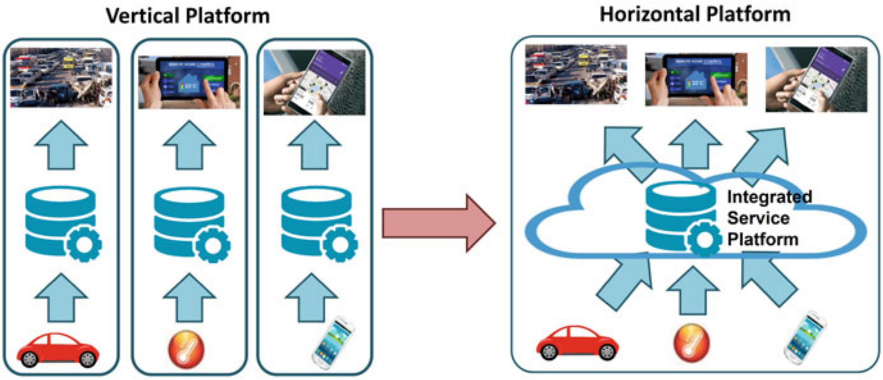


Fig. 7 Development of the IoT platforms, from vertical to horizontal ones

other domains, it is necessary to create a new vertical platform or to find a way to enable the communication between different platforms.

A horizontal platform is designed to operate independently of objects and services, hosting a variety of devices. It provides all the functionality commonly required by different services, such as service search and composition. The immediate advantage is related to the low maintenance cost since it does not depend on any specific service or device and the ease of convergence and connection between services through the use of common interfaces.

The absence of horizontal platforms is leading to a highly fragmented market, where current IoT applications make use of gateways to bridge the world of smart objects with the user through the use of custom applications. However, these apps have prior knowledge of the devices they will work with, so application-specific interaction patterns, communication protocols and data formats are already established. Eventually, this approach can not be sustained as end-users has to install and learn to use a variety of different applications and smart objects which only work seamlessly in their own environment and are not able to dynamically adapt to context conditions or to situations behind their original purpose.

To effectively support the strict requirements of future applications, the research community is promoting the Digital Twin (DT) technology [30].

3.1.1 Digital Twins

Digital Twins, also called Virtual Objects, are the digital representation, semantically enriched, of the services offered through devices attached to physical objects, which are able to acquire, analyse and interpret information about its context, to augment the potentialities of the associated services for the benefits of the quality of life of humans as a final consumer of the real world data [36].

The concept of DT is now widely spread, however, there is still no consensus regarding the direct association between a DT and its physical counterpart. It is still unclear if the DT should match all the services of a single real object or only a subset of them. Let us consider an apartment, where we have home automation in every room: we can create a DT for every room in the house, so to have the finest granularity and control over the house, e.g. related to the temperature, but then we would need to replicate or aggregate some data in order to consider the information at house level, e.g. if we want to optimize the electricity consumption. This needs to further scale down if we want to consider the same information regarding electricity at the building level.

It is then important to find a tradeoff between the number of replicas of the same information in different DTs and their reusability. On the one hand, the creation of a DT for each service can help the service discovery when a particular service is needed by an application. On the other hand, creating a single DT for each real-world object allows the same information to be available for different services and then lessens memory consumption. Indeed, a many-to-many association between physical objects and DTs can provide a higher degree of freedom in the design of

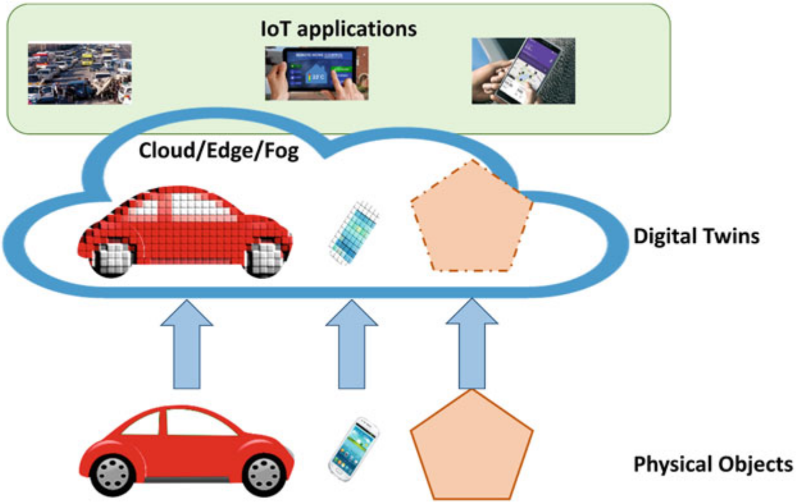


Fig. 8 Concept of digital twin in the IoT

an abstraction layer; the specific choice is related to many factors and will depend on the role that the DT will play in the coming years: as an endpoint to find the information useful for applications, creating a DT for every service could be the best solution; with the advent of the Information-Centric Networking (ICN) approach, the service discovery operations will be content-centric instead of location-centric, so the DTs will act only as gateways to access the virtual world and then a single DT for each real entity could be enough.

In IoT platforms, Digital Twins are usually implemented as web services [28] in the so-called abstraction layer, so that every deployed DT contains all the information needed to fully characterize a physical object and its intended or predicted behaviour. as shown in Fig. 8. The role of this layer differs from platform to platform, but some functionalities can be considered basic and are present in most of them. This applies to several functionalities: the semantic descriptions represents an object in the digital world; context awareness absorbs all incoming data from external objects and performs calculations and transformations on them. In the following years, DTs will be able to manage in a cognitive way the accumulated data in order to make decisions and act upon the physical devices.

DTs are created based on the catalogue of available templates stored in a repository, which describe their offered capabilities (e.g. offered functions). DT templates are created by the system administrator or by the device manufacturer, and can be specific for every device, as is the case for Lysis [20], or generic, as it happens for the channels in ThingSpeak [34]. A DT Registry contains metadata for each installed DT, which is preserved for a specific time period. DT registries store the semantically enriched data that are used for the description of the DTs, in order to be available anytime from anywhere. The stored information may include the instance name and the installation context, such as the DT identifier, associations with ICT and non-ICT objects, location and offered functions. DT, and their composition, registries may be distributed across several domains.

Finally, the expected upsurge in the number of objects involved in the IoT [12], is going to exacerbate the scalability issue in the next few years. DT life cycle will need to be thoroughly managed so that virtual objects are deleted as soon as they are not needed anymore. One of the future challenges related to digitalization will certainly be DT garbage collection.

3.2 The Evolution of IoT Towards the Internet of Everything Paradigm

The Internet of Everything (IoE) paradigm is advancing toward enriching people’s lives by adding value to the Internet of things (IoT), with connections among people, processes, data, and things. The IoE is a term that was first defined by CISCO in 2012 [19]. IoE has a bigger scope than IoT and it aims to look at IoT in a bigger manner, but it all comes down to the definition of *Things*: based on the definition we provide in the previous Section, we can use the terms IoT and IoE interchangeable.

The main goal of IoE systems is to make *everthing* cooperate in order to provide complex services to the final users. Today, the interaction model is based on humans looking for information provided by objects (human-object interaction), but in the IoT near future, this model will quickly shift to object-object interaction, where objects will autonomously look for others.

According to Cisco [13], “the number of devices connected to IP networks will be more than three times the global population by 2023”, thus leading to tens of billions of devices looking for services in a sea of data, without any clue to understanding how the information provided by other objects can be trusted and processed automatically.

The creation of an effective IoE will pass through the definition of an augmented DT; it will encompass the capability to autonomously and adaptively interact with the surrounding environment, in order to dynamically deploy applications for the benefit of humans, so as to improve their quality of life. Which principles and rules should govern the DTs’ behaviour is still to be understood. One of the proposed approaches is that of giving them a behaviour similar to that of humans in the real world, which, with more or less complex social rules, undertake effective interactions in a scalable way. This so-called Social IoT is expected to bring a social network of DTs that exchange data and services following the friendships among them in a scalable and trustworthy way [6]. Other solutions propose following the principles of mankind’s neural system to drive the interactions, as it is the most effective solution proposed by the nature to support intensive communications in complex networks [38]. The evolution and possible use cases that will be engendered by applications of these solutions are still unknown, but it is straightforward that they can be easily enabled by DTs, thanks to the capability of coping with complex networks.

3.2.1 Social Internet of Anything

An approach with the potential to efficiently manage the DTs is based on the exploitation of social networking notions into the IoE, as formalized by the Social Internet of Anything (SIoA) concept [40]. According to this vision, DTs create social relationships among themselves as humans do: this approach introduces the vision of social relationships among different devices so that they are more willing to collaborate with friends with respect to strangers. The idea is to exploit social awareness as a means to turn communicating objects into autonomous decision-making entities. The new social dimension is able to mimic interactions among users and motivate a drift from egoistic behaviour to altruism or reciprocity. Therefore, DTs can autonomously establish social links with each other (by adhering to rules set by their owners) so that "friend" DTs exchange data in a trustworthy manner. A set of forms of socialization among objects is foreseen. The parental object relationship (POR) is defined among similar objects, built in the same period by the same manufacturer (the role of the family is played by the production batch). Moreover, objects can establish a co-location object relationship (CLOR) and co-work object relationship (CWOR), like humans do when they share personal (e.g., cohabitation) or public (e.g., work) experiences. A further type of relationship is defined for objects owned by the same user (mobile phones, game consoles, etc.) that is named ownership object relationship (OOR). The last relationship is established when objects come into contact, sporadically or continuously, for reasons purely related to relations among their owners (e.g., devices/sensors belonging to friends); it is named social object relationship (SOR). These relationships are created and updated based on the objects' features (such as type, computational power, mobility capabilities, brand, etc.) and activities (frequency in meeting the other objects, mainly).

The benefits of this approach have been studied to show how Cloud and Edge nodes play a complementary role in information diffusion [46] and how SIoA can provide a possible solution for collaborative Edge Computing [17] or to exploit the availability of computing resources at the Edge. Furthermore, the exploitation of social networking notions into the IoV has brought to the definition of the Social IoV (SIoV) paradigm, i.e., a social network where every vehicle is capable of establishing social relationships in an autonomous way with other vehicles or road infrastructure equipment [44].

A reference SIoA architecture is shown in Fig. 9, which is based on four levels. The lower layer is made up of the "Things" of the real world, which have the role to sense the physical environment and provide data to the higher layers. The Virtualization layer is made up of Social Digital Twins (SDTs), representing the atomic services of the physical objects in the virtual world. The Micro Engine (ME) is the main entity of the Aggregation layer and represents a mash-up of one or more SDTs and other MEs; on this level, there is no notion of a device, and the

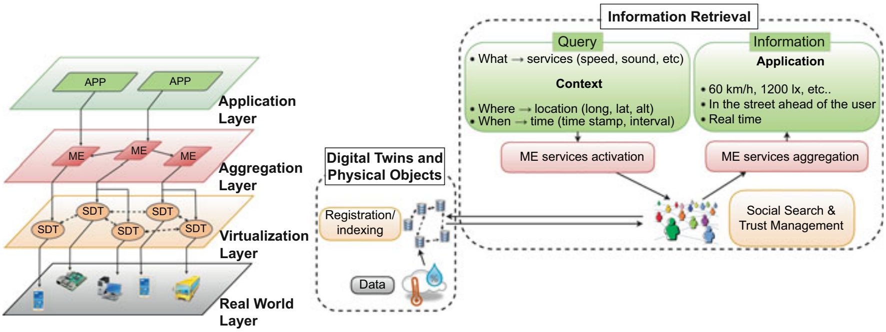


Fig. 9 Reference SIoA architecture

only visible entities are services. Unlike the underlying layers, which implement specific functionality of a single object (physical or digital), this layer implements functionality across all objects. Finally, the Application layer is installed in the Cloud/Edge and partially in the devices, so that applications can be deployed and executed by exploiting one or more MEs.

Whenever the application layer triggers some processing that requires looking for other services, it generates a relevant query. The query specifies what services are required and it is enhanced with context parameters, which represent the application requirements, such as a specific time (when) or a specific place (where). The generated query is then handled at the Aggregation Layer, where the needed MEs for data elaboration are activated. The creation of complex processes can be represented as a sequence of coordinated actions (workflows) performed by individual components. Workflows can be nested, so it is possible to call to workflow from inside another one.

After this, the query is taken over by the SDT of the device that triggered the process, which navigates its social network in order to search for other SDTs that can offer the data related to the desired services by the application. Indeed, in the SIoA, each SDT maintains information related to its friends and to the services that the corresponding physical object can provide. In this sense, SDTs can be seen as atomic registration/indexing servers.

In the following, we describe these perks, which are all desirable for DT-based applications.

Shaping the Network The SIoA structure can be shaped as required to guarantee the network navigability so that the DTs discovery is performed effectively and the scalability is guaranteed like in human social networks. Based on real objects’ movements and profiles, each DT can create its own set of relations with other DTs. All the relations depend on the rules set in the system: these rules have a direct impact on the overall navigability of the network; for the overall network to be navigable, i.e., to enable a DT to easily reach any other DT in the network, all, or the most of, the DTs must be connected, i.e., a giant component must exist in the network, and the effective diameter of the social network must be low. Moreover, the distribution of the number of connections each DT has with its peers, namely the degree distribution, should be close to a power-law distribution. This results in a scale-free network and indicates the presence of hubs, i.e., DTs with a large number of connections w.r.t. the average, in the network.

Trustworthiness Management Trust is typically regarded as essential to cooperation and it has thus been recognized as a critical factor for the IoT [16]. To this, it is essential to understand how the information provided by each DT can be processed automatically by any other peer in the system. For this interaction model, the concept of trust is of utmost importance, since it represents the degree of risk perceived by a DT during a collaboration: if the risk is too high, it means there is no trust between the parties, there is no incentive to enter a relationship and thus any form of cooperation is not feasible. This means that cooperation between DTs can only be effective if there is trust between them so information can efficiently circulate. All the trust models need some inputs and attributes that are associated with the main characteristics of the transaction (requester and provider) and the network. The trust attributes concern the Quality of Service (QoS) or the social area, which are composed in different ways to obtain a trust value. Trust is tied to the concept of reputation: for a DT to quickly collect trust information, it has to rely on the perception of other DTs, that is reputation. Indeed, trust can be gained on both direct and indirect bases, but in large networks, it takes time (sometimes, too much time) for a DT to collect enough direct experience, so a DT has to rely on the perception of other DTs, that is the reputation. Through reputation, it is possible to collect, distribute and aggregate feedback about participants' past behaviour and then provide a global perception of a DT. In particular, social awareness is expected to make the exchange of information and services among different DTs easier and to perform the identification of malicious nodes by creating a society-based view about the trust level of each member of the community. This concept is important for DTs interacting with one another so that they do not need to apply a try and error approach to discover reliable sources, but they can rely on the reputation collected by the community.

Dynamic Discovery and Composition The virtualization of every entity of the real world will lead to a dynamic network of billions of DTs so that the creation of smart services enabling 6G applications will foresee the composition of several dozen or even hundreds of DTs. The dynamic discovery of DTs is a fundamental component of the envisioned framework which is aimed at finding which DTs can provide the required service. In the SIoA, discovery happens in the same way humans seek for friendships and for any information in the social networking services. When the required DTs has been found, the service composition component enables the interaction between them. Most of the time, the interaction is related to a DT that wishes either to retrieve an information about the real world or to find a specific service provided by another DT. The main potential in deploying SIoA is its capability to foster such information retrieval. Leveraging the DT relationships, the discovery procedure finds the desired service. Moreover, every DT can autonomously navigate the network of social DTs so that the service composition is highly personalized based on the DT that started the application, offering to the users personalized experiences when exploiting their applications.

The creation of a social network among all the DTs supporting future 6G applications, as an overlay infrastructure that is exploited to foster trustworthy and effective interactions, following the SIoA paradigm. Groups of DTs with common interests are created to exchange information about events of common interest for the implementation of the services required by the applications. By exploiting the functionalities of the SIoA paradigm, a trustworthy view of these groups is created, which is then exploited to detect events and trigger actions needed to improve the user experience

References

1. 3GPP (2020) NG-RAN; architecture description. Technical Specification (TS) 38.401, 3rd Generation Partnership Project (3GPP), version 16.3.0
2. 3GPP (2020) NR; NR and NG-RAN overall description; stage-2. Technical Specification (TS) 38.300, 3rd Generation Partnership Project (3GPP), version 16.3.0
3. 3GPP (2020) Summary of Rel-16 work items. Technical Report (TR) 21.916, 3rd Generation Partnership Project (3GPP), version 16.2.0
4. 3GPP (2020) System architecture for the 5G system (5GS). Technical Specification (TS) 23.501, 3rd Generation Partnership Project (3GPP), version 16.6.0
5. 3GPP (2023) Summary of Rel-17 work items. Technical Report (TR) 21.917, 3rd Generation Partnership Project (3GPP), version 17.0.1
6. Atzori L, Iera A, Morabito G, Nitti M (2012) Comput Netw 56(16):3594
7. Atzori L, Iera A, Morabito G (2017) Ad Hoc Netw 56:122
8. Babun L, Denney K, Celik ZB, McDaniel P, Uluagac AS (2021) Comput Netw 192:108040
9. Behravan A, Yajnanarayana V, Keskin MF, Chen H, Shrestha D, Abrudan TE, Svensson T, Schindhelm K, Wolfgang A, Lindberg S, et al. (2022) Positioning and sensing in 6g: gaps, challenges, and opportunities. IEEE Veh Technol Mag 18:40–48
10. Bernardos CJ, Uusitalo MA (2021) European vision for the 6g network ecosystem. https://doi.org/10.5281/zenodo.5007671.
11. Björnson E, Wymeersch H, Matthiesen B, Popovski P, Sanguinetti L, de Carvalho E (2022) Reconfigurable intelligent surfaces: a signal processing perspective with wireless applications. IEEE Signal Process Mag 39(2):135–158. https://doi.org/10.1109/MSP.2021.3130549
12. Chettri L, Bera R (2019) IEEE Internet Things J 7(1):16
13. Cisco Annual Internet Report (2018-2023) White Paper. https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html
14. Cox C (2020) An introduction to 5G: the new radio, 5G network and beyond. Wiley, Hoboken
15. Detti A (2018) Functional architecture. In: Marsan MA, Melazzi NB, Buzzi S (eds) 5G Italy white book: from research to market, pp 59–68
16. Din IU, et al (2018) IEEE Access 7:29763
17. Dong P, Ge J, Wang X, Guo S (2021) IEEE Trans Comput Soc Syst 9(1):291
18. Erol-Kantarci M (2022) Deep and reinforcement learning in 5G and 6G networks. IEEE Future Networks World Forum 2022. https://oregon.chapters.comsoc.org/wp-content/uploads/sites/71/2022/05/ErolKantarci_IEEE_Oregon_2022_vd.pdf
19. Bradley J, Barbier J, Handler D (2013) Embracing the Internet of everything, White Paper. https://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoE_Economy.pdf

20. Girau R, Martis S, Atzori L (2016) IEEE Internet Things J 4(1):40
21. Group GPAW (2022) The 6G architecture landscape. Technical specification (TS), 5G PPP Architecture Working Group, version 1.0
22. Guinard D, Trifa V, Mattern F, Wilde E (2011) Architecting the Internet of Things. Springer, Berlin, pp 97–129
23. He H, Yu X, Zhang J, Song S, Letaief KB (2021) Cell-free massive mimo for 6g wireless communication networks. J Commun Inf Netw 6(4):321–335
24. Hexa Project (2022) D1.3: targets and requirements for 6G - initial E2E architecture. Hexa-X Deliverable. https://hexa-x.eu/wp-content/uploads/2022/03/Hexa-X_D1.3.pdf
25. Hexa Project (2022) D1.2: expanded 6G vision, use cases and societal values. Hexa-X Deliverable. https://hexa-x.eu/wp-content/uploads/2022/04/Hexa-X_D1.2_Edited.pdf
26. Huawei (2022) 6G: the next horizon. Huawei White Paper. https://www-file.huawei.com/-/media/corp2020/pdf/tech-insights/1/6g-white-paper-en.pdf
27. Jahid A, Alsharif MH, Hall TJ (2022) A contemporary survey on free space optical communication: potentials, technical challenges, recent advances and research direction. J Netw Comput Appl 200:103311
28. Kaur MJ, Mishra VP, Maheshwari P (2020) Digital twin technologies and smart cities. Springer, Berlin, pp 3–17
29. Kennedy D (2017) D2.6 final report on programme progress and KPIs. https://5g-ppp.eu/wp-content/uploads/2017/10/Euro-5G-D2.6_Final-report-on-programme-progress-and-KPIs.pdf
30. Khan LU, et al (2022) IEEE Commun Mag 60(1):74
31. Lin X (2022) An overview of 5g advanced evolution in 3gpp release 18. IEEE Commun Standards Mag 6(3):77–83. https://doi.org/10.1109/MCOMSTD.0001.2200001
32. Lin J, Yu W, Zhang N, Yang X, Zhang H, Zhao W (2017) IEEE Internet Things J 4(5):1125
33. Mattern F, Floerkemeier C (2010) From active data management to event-based systems and more. Springer, Berlin, pp 242–259
34. Maureira MAG (2014) ThingSpeak-an API and Web Service for the Internet of Things. https://api.semanticscholar.org/CorpusID:28029746
35. Minerva R, Biru A, Rotondi D (2015) IEEE Internet Init 1(1):1
36. Minerva R, Lee GM, Crespi N (2020) Proc IEEE 108(10):1785
37. Nielsen L, Gavras A, Dieudonne M, Mesogiti I, Roosipuu P, Houatra D, Kosmatos E (2022) Beyond 5G/6G KPIs and target values. https://doi.org/10.5281/zenodo.6577506
38. Ning H, Wang Z (2011) IEEE Commun Lett 15(4):461
39. Nitti M, Pilloni V, Colistra G, Atzori L (2015) IEEE Commun Surv Tuts 18(2):1228
40. Roopa M, et al (2019) Comput Commun 139:32
41. Samsung Research (2020) 6G: the next hyper connected experience for all. Samsung White Paper. https://cdn.codeground.org/nsr/downloads/researchareas/20201201_6G_Vision_web.pdf
42. ETSI TS 102 689 V2.1.1 (2013) RTS/M2M-00001ed211, Machine-to-machine communications (M2M) M2M service requirements. https://www.etsi.org/deliver/etsi_ts/102600_102699/102689/02.01.01_60/ts_102689v020101p.pdf
43. Uusitalo MA, et al (2021) 6G vision, value, use cases and technologies from European 6G flagship project hexa-X. IEEE Access 9:160004–160020. https://doi.org/10.1109/ACCESS.2021.3130030
44. Vegni AM, Loscri V (2015) IEEE Commun Surv Tuts 17(4):2397
45. Yazar A, Dogan-Tusha S, Arslan H (2020) 6G vision: an ultra-flexible perspective. ITU J Future Evol Technol 1(1):121–140
46. Yi Y, et al (2020) IEEE Internet Things J 8(4):2177
47. Ying DU IGPG (2022) Views towards IMT for 2030 and beyond. https://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5d/Pages/wsp-imt-vision-2030-and-beyond.aspx

Principles of 6G Wireless Networks

**Qinghe Du, Houbing Song, Xiao Tang, Zixiao Zhao, Chunge Xu,
Yuquan Xiao, and Zhongmin Ma**

1 Introduction

Early in 2018, the 6G technology was first discussed by US FCC in a public forum. Subsequently, Finland started to initiate the 6G research and held the world's first 6G summit from March 24 to 26, 2019, which released the "6G Ubiquitous Wireless Intelligence—Key Drivers and Research Challenges" white paper, claiming that key performance indicators of 6G will be 10 to 100 times higher than those of 5G. In the 6G era, it is possible to download 10 high-definition videos of the same type in one second. In the same year, countries such as the United Kingdom, Germany, Japan, China, and South Korea officially launched their 6G researches. On April 8, 2020, the Japanese Ministry of Internal Affairs and Communications released the strategic goal of establishing 6G major technologies by 2025, hoping to realize 6G practicality by 2030. On July 14, 2020, Samsung in South Korea released the "Next Generation Hyper-Connection Experience" white paper, which pointed out that the peak rate of 6G commercial use will reach 1000 Gbps and the latency will be as low as 100 µs in 2030. In August 2021, LG Electronics in South Korea successfully conducted a wireless signal transmission test in the 6G

Q. Du (✉) · Z. Zhao · C. Xu · Y. Xiao · Z. Ma
School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an, China
e-mail: duqinghe@mail.xjtu.edu.cn

H. Song
Department of Information Systems, University of Maryland Baltimore County, Catonsville, MD, USA

Electrical Engineering and Computer Science Department, Embry-Riddle Aeronautical University, Daytona Beach, FL, USA
e-mail: h.song@ieee.org

X. Tang
School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China

V. Loscri et al. (eds.), *The Road towards 6G: Opportunities, Challenges, and Applications*, https://doi.org/10.1007/978-3-031-42567-7_2

millimeter wave band, the test distance exceeded 100 m. At the Second Global 6G Technology Conference held from March 22 to 24, 2022, nearly 100 authoritative experts from universities and research institutions, telecom operators, equipment manufacturers, and other countries around the world, including the United States, the United Kingdom, Canada, Finland, Sweden, Japan, Singapore, Greece, and Saudi Arabia, conducted comprehensive exchanges and in-depth discussions on eleven topics related to 6G millimeter wave and Terahertz technology, 6G vision and technology requirements, 6G spectrum sharing coexistence technology, 6G network architecture and key technologies, 6G wireless coverage expansion technologies, 6G wireless air interface transmission technologies, 6G wireless network security architecture key, air-ground-integrated intelligent networking technology, and 6G all-scene on-demand service key technologies. In July 2022, the Skolkovo Institute of Science and Technology (SKOLTECH) and the Wireless Radio Research Institute (FSBI NIIR) in Russia prepared to jointly develop 6G network technologies, involving the development of prototypes for mass production as well as network communication security. On July 24, 2022, at the "5G Application and 6G Vision" sub-forum of the Digital China Summit, the IMT-2030 (6G) Promotion Group released the "6G Typical Scenarios and Key Capabilities". On October 12, 2022, a research team from Nagoya University in Japan conducted a research trial of 6G communication networks along the railways in the city. On November 22, 2022, Ericsson announced that it would invest tens of millions of pounds in 6G network research in the UK in the next 10 years, and the research fields included network elasticity and security, artificial intelligence, cognitive networks, and energy efficiency.

Along with the continuously proceeding and accelerating research progress of 6G technologies, it brings increasing anticipation of the 6G vision to further brighten the wireless future towards more diversified and immersive applications. Originating from the classical scenarios in the 5G era as eMMB, mMTC, and URLLC, the new enabling technologies of 6G are expected to fill the gaps between the conventional scenarios and spark new and more exciting applications. The newly emerging application scenarios are widely suggested by many members of IMT [14], which, although different in some ways, basically show commonly-recognized natures. In particular, new application scenarios are inspired by integrating functions of classic scenarios of 5G, targeting new and diversified conceptual applications such as digital twin, metaverse, extended reality, as well as many others. Some typical suggestions on application scenarios for the 6G era include ubiquitous mobile broadband (uMBB), ultra-reliable low-latency broadband communication (ULBC), and massive ultra-reliable low-latency communication (mULC) [16]. Particularly, uMBB intends for high-quality on-board communications and global ubiquitous connectivity, which is expected to cover the whole Earth's surface and remarkably boost the network capacity at hotspot areas. ULBC enables URLLC applications with high throughput, diversifying the critical applications with high-rate features, while mULC extends the URLLC devices towards massive deployment. Given the enhanced scenarios anticipated in 6G, more prosperous and fascinating applications can be enabled and well supported, such as 5D game, XR, digital twin under uMbb, tactile Internet, intelligent transport, smart logistics under mULC, and HTC, multi-sense experience, autonomous driving under ULBC.

Following the trend proceeding to the 6G era, in this chapter we briefly share discussions of principles of 6G wireless networks. While the concepts of 6G will definitely form a significantly-grand and new architecture, which is still under wide research and discussions, this chapter can only cover some typical or unique features or building blocks of 6G rather than give a complete and detailed illustration. The rest of this chapter is organized as follows. Section 2 introduces a brief comparison between 5G and 6G, elaborating the performance gap in terms of main performance indicators. Section 3 discusses the physical-layer technologies, introducing the new paradigm-shifting designs to enable enhanced transmissions. Section 4 addresses the resource allocation strategies, with new perspectives on efficient utilization and management of the multi-dimensional resources in 6G era. Section 5 investigates the networking related topics, emphasizing the new requirements for this classical role in protocol stacks to fit the 6G future.

2 The Gap from 5G to 6G

With the ever-increasing demands for faster data rate, ubiquitous access, and seamless services anywhere and anytime, the 6G network faces formidable challenges in accommodating various use cases and application scenarios that entail novel ideas and visionary goals. According to the latest updates through the ongoing discussions on IMT 2030 and beyond [14], novel and revolutionary use and application scenarios of 6G are desired globally, including ubiquitous intelligence, immersive multimedia, and multi-sensory communications, digital twin and extended world, smart industries and transportation, E-health and well-being, contiguous and ubiquitous connective, integrated sensing and communications, ubiquitous computing, and sustainability.

Nowadays, although commercialization and coverage of 5G is still under fast expansion all over the world, it is readily expected from a technological perspective that the capability of 5G may not be able to support the aforementioned 6G use and application scenarios. The gap from the 5G network to the future vision not only is huge for the crucial yet traditional key performance indicators (KPI), but also lies in new multi-dimensional functions. Many research organizations and industrial companies have released their proposals on KPIs required for 6G, which are also systematically integrated while refined at IMT. Specifically, the following KPIs are the particularly extracted from the vision of IMT-2030 [14], while the specific targets are still under wide negotiation.

2.1 Peak Data Rate

The peak data rate is no doubt the most important indicator for communications networks. The peak data rate requirement for 5G network needs to reach at least 20 Gbps for downlink transmission and 10 Gbps for uplink transmissions under ideal conditions per user or device [33]. In contrast, the peak data rate expected for 6G is roughly 10 times over that of 5G at least. Although the target has not been finalized at IMT-2030, the ambitious peak data rate requirement even reaches the Tbps level, to support information services in applications such as extended reality (XR), digital twin, Metaverse, etc. Although extremely challenging, this goal has been recognized widely across the research communities. The confidence in achieving the goal essentially comes from potential new transmission technologies such as super-large scale antennas, Tera-Hertz (THz) communications, orbital angular momentum (OAM), visible light communications, etc., which can significantly expand the dimension and range of signaling space over frequency, space, and electromagnetism domain, and thus gain an ultra-high data rate.

2.2 User-Experienced Data Rate

User-experienced data rate is required to reach 100 Mbit/s in urban or suburban areas, and is expected to reach 1 Gbit/s for hotspot areas, while these targets might have not been well realized averagely for practical 5G networks. As the potential limit of user-experienced data rate is largely confined by the peak data rate and the number of users served within the same coverage, it is not surprising that the achievable user-experienced data rate shall be improved 10 times in 6G compared to that in 5G. Therefore, 1 Gbit/s is likely to be the minimum requirement for user-experienced data for 6G, while the key to this goal shall highly rely on efficient interference management and continuously decreased cell size.

2.3 Spectrum Efficiency

The spectrum efficiency of 5G has been improved to a considerably high level, which attains 30 bps/Hz and 15 bps/Hz for uplink and downlink, respectively. Based on current discussions in the research community, 6G is expected to reach about 1.5 3 times higher spectrum efficiency than that of 5G. It is worth noting that the improvement of spectrum efficiency is not that large as compared to that of the data peak rate since the link-level potential has been extensively exploited, where the current link-level transmission has been approaching the Shannon limit. The main path to improving spectral efficiency would possibly be concentrating on digging the space-domain resources by introducing a super-large antenna array. In the meantime, thorough and efficient interference suppression across overlapped coverage is also crucial.

2.4 Area Traffic Capacity and Connection Density

Area traffic capacity refers to the total throughput offered in unit geographic area. Based on the 5G vision specified by IMT-2020, the area traffic capacity requirement is 10 Mbit/s/m^2, whereas the 6G network is expected to reach a 100-times bump. For different application scenarios, the area traffic capacity could vary from 100 Mbit/s/m^2 to 1000 Mbits/m^2. This KPI marks the revolutionary increment of capability in continuous coverage by mobile networks.

The connect density describes the number of accessible devices per unit area. Based on the 6G vision, the averagely estimated density is 10 times higher than that of 5G, and some more aggressive desire is 100 times larger, implying that the connection density can increase to 10^7~10^8 devices/km^2, i.e., there will have 10~100 devices per m^2. The explosive increment of connection density aims at connecting devices held by human as well as machine-type nodes, building a fully-connected smart cyber-physical system/network.

2.5 Mobility

Extending the coverage is one of the core missions for The 6G network. Aside from satellite coverage, service provided by terrestrial access points/stations need to be extended to the commercial airplane and high-speed trains. In these scenarios, the moving speed may reach 1000 km/hour with particular QoS requirements, which is twice as much as the highest mobility considered in 5G. The maximum Doppler frequency spread could go beyond 15 kHz at a carrier frequency of 2 GHz, which would be worse when the carrier frequency goes higher. OFDM-like modulation approaches encounter a rather stringent challenge than even in these scenarios, demanding the new ceiling-breaking techniques for traditional physical-layer in extremely-high mobility regimes.

2.6 Latency and Reliability

IMT-2020 visions to confine air-interface latency within 1 ms and 99.999% reliability in 5G systems, which is then able to facilitate ultra-reliable and low-latency communications (uRLLC), serving application scenarios in industrial automation as well as self-driving cars. However, due to the random access procedures for bursty traffic, collisions among different users' requests are unavoidable, and 1-ms latency remains extremely challenging for The 6G network. Although delay caused by collisions is often categorized as access delay rather than air-interface latency, the accumulated latency as a whole, which is typically attributed to queuing delay, access delay, and air-interface latency (transmission delay as well as processing delay), will severely affect users' quality-of-experience (QoE). In such a case, shortening air-interface latency is still considered an important contribution to delay QoE enhancement, as it determines how quickly each packet can get through the link bottleneck over the air. Correspondingly, wide discussions are suggesting lowering the air-interface latency as much as possible towards the challenging level of 0.1 ms.

2.7 Positioning and Sensing Capabilities

Positioning is a new KPI formally introduced for 6G in contrast to 5G visions. The research communities as well as the industries hope to achieve accurate positioning function via capturing and extracting propagation features, which would no doubt be beneficial for resource allocation, security monitoring, etc. Positioning functions can be further divided into horizontal positioning and vertical positioning which are challenging in different aspects. The well-recognized accuracy for positioning needs to be around cm-level to enable practical use cases.

The sensing function marks a more generic yet powerful capability for The 6G network. Sensing functions include ranging, velocity and angle estimation, localization, imaging, object detection, and mapping. All the information can be tailored, fused, and distributed, to facilitate diversified applications, and eventually be crucial to building the digital-twin system in a wide sense.

2.8 AI Capability

AI capability has been envisioned to be incorporated into IMT-2030 vision, which is expected to offer smart functionalities over IMT systems. These functionalities are targeted to serve AI-enabled user applications. The applications include data collection, AI training, inference, computing, etc., while specific applications or modes have neither been well formulated nor demonstrated. It is not a surprise that AI capability has introduced new tasks for mobile communications networks, and therefore studies on how to integrate essential communications functions with AI-supported abilities will persistently attract extensive research attention.

2.9 Security, Privacy, and Trustworthiness

From a commercial perspective, security protection for mobile communications networks has functioned well along with the evolution to 5G. However, 6G has raised much more aggressive and ambitious requirements on security. First, security needs to be assured over the entire lifecycle of IMT-2030. Second, the protection of user privacy becomes more and more important, not only for the user's data or information, but also extended to the user's actions and intentions. Third, two extremely challenging targets have been widely agreed upon as features for the security of 6G network, namely the zero-trust principle and one-time pad secure transmissions. Last but not least, security towards continuous running of the network has now been put into an ultra critical role. It is crucial to keep a considerable percentage of the network operating and functioning normally even in extreme weather events, faults, and/or natural disaster situations.

2.10 Sustainability

One controversial issue for 5G networks is the high energy consumption in maintaining the large scale of infrastructure. Persistently improving energy efficiency and lowering energy consumption is no doubt the mission of 6G as well as the active responses to global demands on carbon neutrality. It is highly desired for 6G to continuously reduce the negative impact on the environment, contributing to sustainable and harmonized living conditions. In the meantime, sustainability requirements are also imposed on some network services itself. Specifically, extremely massive machine-type devices are the essential to implement a fully-connected cyber-physical network with sensing capability. As a consequence, extending the lifetime of not only the battery, but also the devices' hardware becomes a more serious problem than ever. In addition, software extensibility will also be critical to prolong the devices' lifetime. In summary, the vision for 6G is getting clear, which is rather exciting yet also faced with a huge gap with current 5G systems, motivating ten years of continuing efforts from standardization organizations, academic researchers, as well as industrial bodies in related areas.

3 PHY Layer Design

While the evolution of mobile communications networks towards 6G is driven and inspired by diverse emerging and visioned highly-resource-consuming applications, persistently-surprising innovation of physical-layer technologies is the fundamental power initiating as well as supporting the revolution. There have been so many proposals and research results towards 6G, which can hardly be comprehensively covered in one chapter, and thus next we mainly only share discussions on several crucial paradigm-changing PHY designs, namely digging new signal spaces, reshaping transmission over the air, building new coverage paradigm, integrating communications with sensing and computing, etc.

3.1 Dig New Signal Spaces

The 6G infrastructure expects to support higher data rate, offer better experience, accommodate more users, and serve more diversified applications. To achieve these ambitious goals, the first priority is to be able to dig new signal spaces via expansions of new wireless resources. We in this section will discuss enabling technologies of resource expansions in frequency domain, space domain, magnetoelectric structure domain, where the representative technologies include mmWave and Terahertz communications, extremely large-scale MIMO systems, orbital angular momentum (OAM) based communications, etc.

3.1.1 Communications Over Higher Frequency Band

The accommodation of communications services for more devices and people reply on the availability of sufficient amount of access and transmission resources. Sub-6 GHz frequency band has been fully occupied and extremely crowded. Before cognitive radio technologies become mature and applicable for practical systems, very little room exists for sub-6 GHz band. As a consequence, moving to higher frequency band is a direct and effective way [6, 29–31, 35, 48]. The MmWave communications technology covering around 30 to 300 GHz frequency band have been extensively studied for the 5G system, but have not been applied and embedded for practical product, leaving mmWave communication still a key enabler for the 6G system serving indoor coverage, backhaul transmissions, etc. Terahertz ranges from 300 GHz to 3 THz, which offers extremely rich band having not been explored. For these high-frequency bands, either propagation model, signal processing, or system implementation techniques have not been fully conquered, which will be concentrated topics attracting many research efforts.

3.1.2 Extremely Large-Scale MIMO Systems

Extremely large-scale MIMO systems expect to deploy a very large number of antennas in a compact space [47], creating new wireless resources in space domain. For 6G systems, the number of antenna equipped at a practical access point can go up to one thousand. Even with cheap radio module, the number of simultaneously-serving users is at least beyond. The drastic increment in system capacity comes from the ultra-narrow spatial beam via beamforming over massive antenna array. In the meantime, the extremely-large scale antenna can maximally reduce the cross interference across coexisting signal streaming. In contrast to the very attractive user-accommodation capability, there exist several crucial yet fatal challenging for implementation. First, signal space with such a huge dimension introduces too much computing burden for signal processing. Even the linear receiver's complexity might go beyond the chip's computing resources. Second, channel estimation and feedback need too much overhead, significantly lowering the spectrum efficiency. Third, interference residue will severely degrade the detection performance of target signals. Yet there is still a long way towards extremely large-scale MIMO systems, the promising feature is worth persistent research efforts. Other technologies related to extremely large-scale antennas include holographic MIMO, reconfigurable intelligent surface, etc, which brings new functions for 6G and will be discussed in Sect. 3.2.

3.1.3 Orbital Angular Momentum Based Communications

Orbital angular momentum is a type of electromagnetic wave's wavefront structure with helical phase [3, 44]. Following the study on this topic, certain modes of OAM-based vortex waves were found to have different yet orthogonal structures, and thus inspiring new chance for very high-rate multiplexing in digital communications links [3, 36, 44]. OAM-based communications technologies were initially proposed in optic communications, which have been demonstrated also effective for wireless transmissions [3, 28]. Key challenges of OAM-based wireless communications applied to 6G are summarized as follows. On one hand, QAM-based vortex wave is subject to losing orthogonal properties across different modes after long-distance transmissions. On the other hand, the numbers of modes are limited and thus somehow confine the data rate from further improving. Aside from challenges for QAM, QAM does offer a new angle to obtain extra signal space dimensions to achieve high data rate, which might inspire new research efforts on features of electromagnetic wave.

3.2 Reshape the Propagation Environments

In recent years, intelligent reflective surface (IRS) is considered as a technology to break the traditional communication paradigm by integrating a large number of low-cost passive reflective elements. Each element can adjust the reflection coefficient separately such that the wireless transmission channels can be shaped with beam directly better aiming at the receiver. With fast development of metamaterials and micro electro mechanical systems (MEMS) [22], the reflecting phase/coefficients can be adjusted in a realtime manner, enabling fast reconfiguration of IRS. As a consequence, IRS has been widely accepted as a key enabling technology for 6G. Specifically, the low cost features of IRS make it readily deployed in many scenarios. More importantly, devices even with omnidirectional antenna still can gain more concentrated beam with the aide of IRS. In addition, under the IRS with relatively large scale of reflective units, multiple coexisting links can all benefit simultaneously with optimization over just one IRS.

However, it is still worth noting that due to the passive nature of reflective surface, some harsh challenges have not been overcome. In particular, IRS actually introduces two concatenated reflecting links to each transmitter-receiver connection. On one hand, channel estimation based on pilots for the two reflecting links separately exists unavoidable vagueness because the cascaded channel matrix results from multiplication over the two reflecting links' channel matrices. Although this vagueness is not crucial for a single transmitter-receiver pair, it brings hurdle for optimization in multi-user cases. On the other hand, channel estimation for IRS currently lacks an effective approach for large-scale reflecting unit in that confined by the number of transmitting antennas, IRS needs to turn the reflecting unit on and off to finish probing all elements of the channel matrices. Aside from the issues in channel estimation, the passive nature of IRS leaves many inconveniences towards optimization in a more complicated connecting environment.

To enhance the functions of reflective surfaces, holographic MIMO technologies [13] have been proposed. In contrast to the conventional passive reflectors based IRS systems, holographic MIMO technologies can integrate proactive reflecting units, which not only improve the controllability for beamforming, but also create conditions for easier channel estimation of two separate reflecting links. Moreover, continuous surface materials are also developed and can be applied to further enhance the capability of reflecting control. Many applications like connections across buildings, in-building coverage and positioning, energy-efficient beamforming, physical-layer security, etc., can then benefit significantly from the holographic MIMO technologies.

3.3 Build New Coverage Paradigm

Cellular coverage pattern has been faithfully followed from 2G to the current 5G networks design. However, the limited range, inter-cell interference and handoff, as well as indoor coverage are often the main issues to result in the bad experience for users. To tackle these issues, the new coverage paradigms are being studied, such as massive-MIMO-enhanced coverage, IRS-enhanced coverage, and cell-free coverage pattern [1, 32, 38]. In light of the increased antenna array, massive MIMO techniques are expected to support the massive access of ultra-dense users. Also, empowered by advanced signal processing techniques, beamforming techniques can be employed to direct the signal to specific users, thus overcoming obstacles such as buildings or other obstructions. IRS techniques can achieve single amplification and interference reduction. Furthermore, with the low-cost and flexible characteristics, it can be broad-range deployed to significantly enhance the coverage of traditional cellular networks. Except for the enhancement of traditional cellular coverage, a more revolutionary cell-free coverage paradigm has been developed, where a large number of access points are spread out over a wide area to support more users' and devices' connectivity. Compared with the traditional cellular networks, the issues of dead zones can be eliminated. Moreover, it has been shown that cell-free paradigm can guarantee the lower latency and higher energy efficiency than cellular networks. Also, the cell-free system based on a large number of access points can be regarded as a distributed massive MIMO system, called cell-free massive MIMO system, which has superior performance to the co-located massive MIMO systems in terms of coverage.

3.4 Integration of Communications with Sensing

Integrated sensing and communication(ISAC) is regarded as one of the signature features of 6G [5, 25, 26, 52]. As radio sensing and communication (S&C) systems are all evolving to higher frequency bands and larger antenna arrays, they are becoming more similar in terms of hardware architecture, channel characteristics, and signal processing. Therefore, radio sensing functions can be effectively integrated into 6G Radio Access Networks (RANs) and future cellular networks will enable imaging, monitoring, and location capabilities of the surrounding environment.

As a new topic, ISAC has many technical issues to be investigated in physical layer, and the key technologies include the integrated waveform design [23], fundamental information theoretical limits for ISAC [27], joint signal detection, and interference cancellation algorithms [45], and basic performance trade-off analysis, etc [5]. The integrated waveform design can be divided into Non-Overlapped design and fully unified waveform design.The Non-Overlapped design allocates orthogonal wireless resources to the communication and sensing functions which can be realized over temporal, spectral, spatial, or even code domains. In contrast, fully unified ISAC waveform enables communication and radio sensing functions on the same transmitted waveform which can be divided into three categories namely, sensing-centric design (SCD), communication-centric design (CCD), and joint design (JD) approaches with different functional emphasis. SCD and CCD aim to embed communication functions into existing sensing waveforms/structures. JD does not rely on existing sensing or communication waveforms and it conceives entirely new ISAC waveforms which provides additional freedom and flexibility. Information theory is essential for performance limits of wireless communication systems but is not exactly applicable to ISAC systems [27]. The most famous information-theoretic result related to ISAC is from the seminal paper by Guo et al [10]. That paper connected the input-output mutual information as a communication metric and the minimum mean square error (MMSE) as a sensing metric. Finally for the analysis and measurement of performance gain, it is clear that ISAC systems can gain overall performance by sharing wireless resources and hardware devices, but there is not a meaningful metric to measure the performance gain yet.

4 Resources Allocation

Resources allocations in 6G are extremely challenging due to the significantly improved performance metrics, such as very low end-to-end latency, much higher reliability, ever-increasing spectral efficiency, etc. In the meantime, the 6G vision inclines to integrate diversified heterogeneous resources together to serve the large population of mobile users and terminals, making resource allocations a considerably complicated task. While many resource allocation theories, strategies, and schemes have been proposed for The 6G network [12, 15, 19, 24, 49, 53, 55, 56], we in this chapter concentrate on two key aspects in implementing the effective and efficient resource allocations for The 6G network, including distributed AI architecture and integration of heterogeneous resources. Specifically, the distributed AI architecture is responsible for connecting the cutting-edge learning technologies with the communication systems as well as the networking protocol stack; on the other hand, fulfillment of QoS/QoE for the persistently-expanding user scales essentially rely on exploiting diverse available and new resources, which are the main reasons we hope to address these two issues in this space-limited chapter.

4.1 Distributed AI for Resource Allocation

The distributed artificial intelligence (AI) is natively embedded in the 6G network and provides integrated coordination for both cross-layer and internal function units, which will be beneficial to deal with the challenges mentioned previously. The missions of distributed AI of the 6G network are not only to assist the prediction of networks' behaviors accurately and conduct efficient optimization for faster responses, but also to sense heterogeneous resources and coordinate various users' requirements. One possible logical structure of the distributed AI is illustrated in Fig. 1 across two dimensions. One is AI's vertical dimension in each function layer of a communication system, and the other is AI's horizontal dimension between multiple users and between users and the upper-layer.

As shown in Fig. 1a, we suggest that the adoption of the AI structure should be driven by data and model simultaneously. Each layer has a sensor node that is used to sense and collect available information from the space–air–ground–sea environment. The top application layer can analyze and extract features from the service's QoS and then deliver these to other layers based on different models' requirements. Data from both sensing and services are used to train and update the AI models adaptively, i.e., model self-learning. Typical schemes employed to fulfill the self-learning of model include GAN, deep learning (DL), and LSTM. GAN has the ability to dynamically monitor and simulate the internal features of observations and then derive an abstract model to depict statistical rules. DL can seize the deep connection of input data and iterate towards expected outputs. LSTM is suitable for models with correlations in the time domain; thus, it can predict the model's future

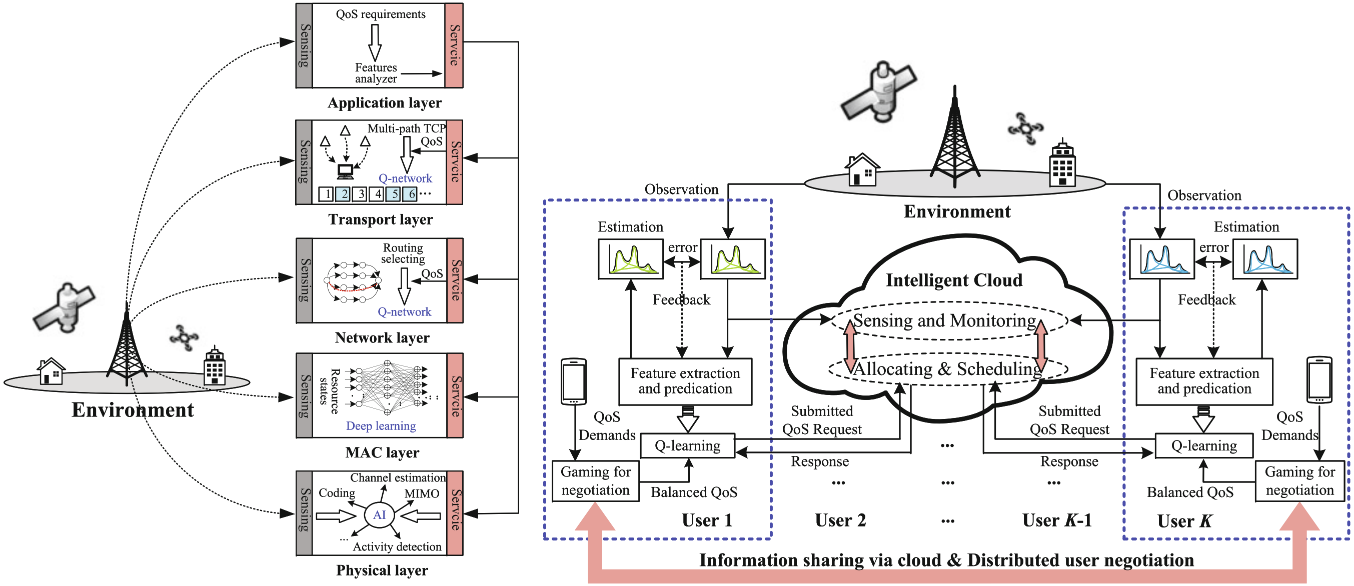


Fig. 1 The model of the distributed AI: the left graph denoted as (**a**) depicts the vertical dimension; the right graph denoted as (**b**) depicts the horizontal dimension

trends based on its historical performances. Specifically, the traditional models in the physical layer, which were mainly derived according to the characteristics of transmitters and common protocols, can be generated and updated adaptively with the assistance of AI techniques in The 6G network. In this regard, GAN has shown the potential to model the channels [40]. For the MAC layer, employing DL to extract implicit information from resource states can be considered as the solution. In the transport layer and network layer, on the one hand, it can use the Q-learning network to solve multi-scale computing problems (e.g., routing selections and multipath transmissions), and on the other hand, these models can learn from external data driven by GAN, DL, LSTM, etc.

In Fig. 1b, the process of information sharing and negotiation across users is depicted, which is driven by the distributed AI framework. First, each user obtains observations from the network environments, including information of diverse resources (e.g., spectrum, caching, and computing resources) and transmission conditions. These observations are injected into an extraction/prediction module to get key features of the transmission and network environments. Also, future environments can be estimated via the prediction function, which is used to guide user's transmission strategy. The estimation is further compared with the observations, and the obtained error is also fed back to the extraction/prediction module, which will be used to optimize the local extraction and prediction algorithms. In the meantime, the observations by users are also uploaded to the upper-layer, so that it can sense and monitor environment information over the entire network. Second, the upper-layer will help distribute user's QoS demands to each other, together with the resource and network environment information. With this information, each user will conduct negotiation in a distributive manner, through which each user will adjust its own QoS requirements to a more balanced level, i.e., acceptable for itself yet potentially feasible for the network. For QoS negotiation, users will adopt game-theory based approach, as the game theory has been proven to be capable of effectively solving complicated problems such as resource competition, coordination, allocation, etc. over wireless networks and the game theory itself is persistently evolving with more intelligence [11, 37, 41, 42, 51]. The negotiated QoS and extracted features from observations and estimation are used for the user to finalize its QoS requirements by reinforcement learning approach. Then, by analyzing users' QoS requests and the available resources over the network, the intelligent architecture can effectively optimize resource allocation and scheduling improving the network performances while meeting users' QoS requirements. Moreover, the optimized scheduling and resource allocation strategies will be sent back to users such that they can better train their learning module.

4.2 The Aggregation of Heterogeneous Resources

Delay-sensitive and real-time services will occupy a large proportion of applications served by the 6G network. To meet the QoS/QoE requirements for these services,

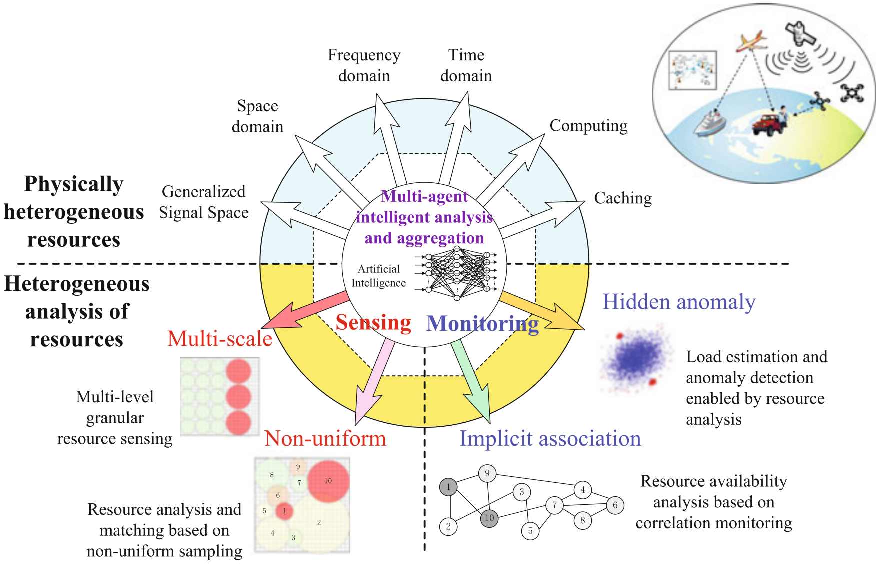


Fig. 2 Heterogeneous resource analysis and aggregation via intelligent sensing and monitoring

we expect that continuously distributed yet stable resource blocks be allocated to them. To satisfy this requirement, the network needs to dig out all potential resources embedded in diverse domains. As shown in Fig. 2, the precious air-interface resources exist not only in the traditional domains [54] such as time domain, frequency domain, spatial domain, but also hide over generalized signal space domain, waiting to be activated. Moreover, computing and caching resources [20] have been widely recognized as crucial supporting components for mobile communications, which further extend the concept of resources towards the success of 6G yet increase resources' heterogeneousness. As a consequence, how to aggregate these heterogeneous together to jointly serve various applications becomes a new issue in The 6G network.

On the other hand, aggregation of heterogeneous resources imposes another task for resource allocation. As aforementioned, pursuing allocation of resources in a continuous manner is desired for many services. But an evident negative consequence will appear, i.e., there will exist a great deal of resource fragments, showing higher heterogeneousness in terms of multiple scales and the non-uniform feature, causing low-efficiency usage of the precious wireless resources. Identification of these heterogeneous fragments, needless to say, is useful to avoid capacity loss but needs two-fold innovations in enabling technologies: multi-scale resource sensing and non-uniform resource sampling. The multi-scale characteristic emphasizes the diversities of methods (e.g., satellites, UAVs, robots, smartphones) and content (different resource domains) as well as accuracy (in meters, even decimeters) and timeliness (real-time or intermittent) in resource sensing. The non-uniform characteristic leads to higher requirements for heterogeneous resource analysis and matching. These fragments, if well organized, can effectively support many services that do not have stringent QoS/QoE requirements. Based on the above discussion, a potential strategy is to build a database for users' resource-usage patterns that reflects the users' access types and major service categories and maintains the resource-fragment map along with multiple resource domains. Then, a quick matching algorithm between the users and the resource-fragment map with relatively low complexity can be designed.

In addition, when aggregating the discovered available resources together, we shall note that the correlation as well as the anomaly hidden underneath the surface of observations. In particular, as shown in Fig. 2, there exist implicit associations across resources within or belonging to different domains. Also, anomaly of resource status sometimes might not be readily observed, which then will be very harmful if the corresponding resources are allocated. Therefore, advanced technologies integrating machine intelligence are urgently needed in The 6G network. In summary, we can foresee that the design and implementation of fast and efficient intelligent architecture as well as intelligent algorithms/schemes will become a high-priority task to deal with the explosively-increasing scale and loads in the densely deployed heterogeneous 6G network with all heterogeneous resources.

5 Networking-Related Issues

The above-mentioned issues mainly focus on the lower-layer design, such as PHY layer design, data-link layer design, and the joint resource optimization between them. Since the 6G network is expected to have strong capability of service awareness, not functioning simply as the tunnel for information, core networks and end-to-end upper-layer protocols (from the network layer to application layer), which have not sufficiently considered services' features, need to be adjusted and innovated. Towards this end, we will first investigate the vision on 6G core network. Then, we discuss the collaboration between lower and upper protocol layers and novel transmission control protocols for emerging services. Specifically, we will cover topics including tightly coupled cross-layer design, service-aware transmission over heterogeneous radio-access technologies (RAT), enabling user socialization for cloud extension, age control protocol for synchronization services, etc.

5.1 6G Core Network

Along with the various emerging technologies, a novel 6G core network architecture is necessary for the flexible and intelligent 6G communication networks [4, 21, 46]. The future 6G core network should have the following characteristics. Firstly, the complex and changing access scenarios require that the 6G core network can cope with the rapidly changing access to different types of networks. Also the recognition and prediction capability of access network environment changes is essential for the core network. Secondly numerous sensing scenarios demand huge amounts of computing resources and fast response, so the 6G core network should be enabled to redispatch communication and computing resources at fine granularity quickly after identifying changes in the network environment and different services. Thirdly when users move across two edge core networks, the 6G core network requires addressing cross-domain switching of the network and migration and prediction of services. Service migration technology can realize network reconstruction, service state preservation, data transmission, and state recovery during service migration by unified scheduling of communication and computing resources. By learning the trajectories of users, satellites, and drones, prediction technology can predict the occurrence time and destination nodes of service migration to reduce costs and improve the utilization of system resources. Finally, the 6G core network will be integrated with AI. Traditional cloud-driven AI places a heavy load on the backbone network with long latency, and the challenge of developing compute resources at the edge of the network with resource constraints. In 6G, AI will empower the network with computing power, and intelligent controllers will dispatch all devices within the network using the power of AI.

5.2 Tightly Coupled Cross-Layer Design with Service Awareness

The RAN architecture and diversified heterogeneous resources of the 6G network exhibit significant differences from conventional networks. However, the network protocol stack in each network node is almost unchanged. Thus, the question of “whether the current protocol stack fits well with the network environments nowadays?” arises. Two attractive new features for the 6G network are service-aware access with seamless coverage and the presence of various heterogeneous RATs. Correspondingly, the moving users and varying mobile channels, even those moving at a very slow speed, result in highly varying network statuses [34].

Figure 3 shows how the conventional protocol stack reacts to fast changes in the network. The higher the protocol layer is, the longer the response time is, and this may reach some tens of seconds. This significantly lowers the efficiency of mobile networks. The inconsistent problem across higher and lower protocol layers calls for the cross-layer design towards a tightly-coupled protocol stack. The tightly-coupled protocol stack shall have the following characteristics. First, the upper protocol layer can obtain and analyze the information from lower layers in a smaller time scale. Second, different protocol layers shall have the mechanism to check the consistence for users’ experiences. Third, it might be better to initialize a virtual protocol stack in addition to the working protocol stack entity, which integrates the concept of digital twin and enables fast predictive simulation, speeding the process for consistent

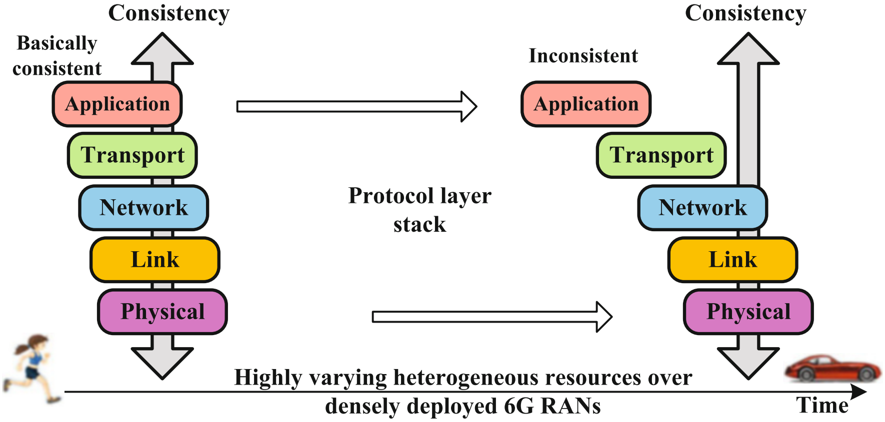


Fig. 3 Inconsistent responses to the highly varying heterogeneous network status caused by the loosely coupled protocol stack

decision for all layers. In summary, the protocol stack does need a consideration revolution in architecture level to adapt the persistently-fast-varying environment in 6G.

5.3 Service-Aware Transmission over Heterogeneous RATs

In the 5G network, users can switch connections between the NR and LTE base stations depending on which one can provide better QoS. In the heterogeneous 6G network, a mobile user might be able to connect via more air interfaces (different RATs) to receive single or multiple services from the cloud, and this will be more complicated than the switch in 5G. We have discussed the tightly coupled protocol design in the above, and it has the potential to support more frequent and diverse switching. If the end user is requesting multiple services, each service can be simply allocated to one RAT that is separated from the others. Clearly, the challenge lies in the attempt to obtain a single service from multiple RATs. Data packets belonging to a single service might come from different RRHs due to the diversified RATs present in the 6G network. However, the packet sequence is not consecutive in each path, and thus, packet reordering in the mobile user is needed for data recovery.

One of the most popular approaches to support multi-path transmission is the multi-path TCP [18]. The essential issue in multi-path TCP is the disharmony between the TCP's error control mechanism and the independent channel-quality/path-bandwidth of different paths. The TCP typically maintains a sliding window for the transmitter and receiver. At the receiver end, if a packet is lost or delayed, causing timeout, retransmission of all packets with proceeding sequence numbers, even some of which have correctly arrived, will be requested. This problem may occur often in the 6G network because the heterogeneous RATs have independent channel fading and congestion statuses. Moreover, the imbalanced bandwidths across different RATs might severely aggravate this problem. Aside from the multi-path TCP, the P2P approach via a consistent hash table and resource discovery and cooperation, might be an alternative approach. However, it is not suitable for the 6G RAN, as it requires cooperation across multiple paths, yielding a considerable overhead and thus not making itself a good candidate. The fountain code has the potential to increase the speed of delivery for single service over multiple air interfaces with different RATs. Multiple paths are analogous to the distributed servers mentioned above. The coded packets from different RATs are not restricted by their arrival order. Furthermore, unlike the multi-path TCP, the decoding of the fountain code concentrates on the accumulation of coded packets and is not affected by the bandwidth imbalance across multiple RATs.

5.4 Enabling User Socialization for Cloud Extension

Regardless of how robust and delicate the design architecture is, the provision of seamless high-quality services for users in various environments and locations is still extremely difficult. One promising solution is to enable user-centric networking in the 6G network, where the user plays the role of not only the consumer but also the provider, such that the cloud-based architecture of RAN can be extended to the user level [2, 7, 9, 17, 43, 57]. Device-to-device (D2D) communications provide strong support to achieve this goal, as they allow a device to relay the data to other devices in proximity, effectively benefiting the dead-corner coverage, data sharing, and even the handover in the RAN of The 6G network [58]. Unfortunately, the standardization of D2D has recently progressed slowly due to a lack of economic incentives for both service providers and mobile users. To solve this problem, data relay among users requires some incentive [8]. In contrast, the beneficiaries need to pay the helper in some way, enabling socialization across end users. This can be also easily integrated into the cloud architecture from the provider side. However, there are still many open problems to be researched and studied, including how to define and manage the payment, whether part of payment should be forwarded to the provider, how the payment is associated with service types, etc.

5.5 Age Control Protocol for Synchronization Services

The concept of age of information (AoI) has been prevailing in timeliness-sensitive system design, which is an end-to-end metric to evaluate the freshness of information [50]. AoI is defined as how long time has been elapsing since the information was generated at the source. How to achieve the wonderful AoI performance or guarantee the diverse AoI requirements for different applications is a crucial issue. Toward this end, in the transportation layer, a novel control protocol, called age control protocol (ACP), has been developed. ACP is used to regulate the packet updating rate from the source to the destination over the Internet aiming to achieve the minimum average AoI with the constrained network resource but no prior information about it. Different from the traditional TCP expected to guarantee the reliable delivery by using the retransmission mechanism, ACP does not consider any retransmission. Once the error transmission is aware at the transmitter, the newly freshest information is preferred to be generated and transmitted. The congestion control mechanism together with ACP should be also carefully considered, which is significantly distinctive from that of TCP aiming to guarantee the reliable and ordered packet transmission. Consider the case in which N-hop path between the source and destination. On the assumption that there does not exist any other traffic, the optimal congestion control mechanism is to assure in each server only one status packet is being severed meanwhile the waiting queue of each server is idle. Although the assumption is ideal, it has been shown via the simulation results that a good age control algorithm must strive to have as many update packets in transit as possible while simultaneously ensuring that these updates avoid waiting for other previously queued updates. The work [39] shows an alternative strategy that uses UDP as the basic transmission protocol and then considers adapting the updating rate. In order to obtain the rate which matches the current network state, the ACK containing the timestamp of the update being received at the destination should feedback to the source. Then, the source according to the recently received ACK adjusts the updating rate. Nevertheless, the ACP is still a relatively new protocol, and more research is needed to further optimize its performance and overcome its challenges.

We have briefly presented the principle of the 6G network in this whole chapter, with emphases on vision and discussions to give a large picture of 6G from the PHY, resource allocation, and networking perspectives. As the 6G network is a grand project with a huge amount of innovation covering diverse aspects of mobile communications networks, only some essentially-crucial enabling technologies and new concepts were shared in this chapter. Among many features expected and desired for the 6G network, exploiting new resources, equipping service awareness, developing AI-native architecture, and end-to-end protocol-stack innovation would be critically important. The standardization process of 6G still has a long way to go, and thus the ideas and designs on 6G will no doubt keep evolving. While the 5G network does not really achieve every goal it promises, research on the 6G network will still carry these tasks as well as fight for many newly-encounter challenges.

References

1. Ammar HA, Adve R, Shahbazpanahi S, Boudreau G, Srinivas KV (2022) User-centric cell-free massive MIMO networks: a survey of opportunities, challenges and solutions. IEEE Commun Surv Tuts 24(1):611–652

2. Cheng W, Zhang X, Zhang H (2016) Optimal power allocation with statistical QoS provisioning for D2D and cellular communications over underlaying wireless networks. IEEE J Sel Areas Commun 34(1):151–162
3. Cheng W, Zhang W, Jing H, Gao S, Zhang H (2019) Orbital angular momentum for wireless communications. IEEE Wireless Commun. 26(1):100–107. https://doi.org/10.1109/MWC.2017.1700370
4. Corici M, Troudt E, Magedanz T, Schotten H (2022) Organic 6G networks: decomplexification of software-based core networks. In: Proceedings 2022 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit), Grenoble
5. Cui Y, Liu F, Jing X, Mu J (2021) Integrating sensing and communications for ubiquitous IoT: applications, trends, and challenges. IEEE Netw 35(5):158–167. https://doi.org/10.1109/MNET.010.2100152
6. Cui M, Wu Z, Lu Y, Wei X, Dai L (2023) Near-field MIMO communications for 6G: fundamentals, challenges, potentials, and future directions. IEEE Commun Mag 61(1):40–46. https://doi.org/10.1109/MCOM.004.2200136
7. Du Q, Song H, Zhu X (2019) Social-feature enabled communications among devices toward the smart IoT community. IEEE Commun Mag 57(1):130–137
8. Gao Y, Xiao Y, Wu M, Xiao M, Shao J (2019) Dynamic social-aware peer selection for cooperative relay management with D2D communications. IEEE Trans Commun 67(5):3124–3139
9. Garcia V, Zhou Y, Shi JL (2014) Coordinated multipoint transmission in dense cellular networks with user-centric adaptive clustering. IEEE Trans Wireless Commun 13(8):4297–4308
10. Guo D, Shamai S, Verdu S (2005) Mutual information and minimum mean-square error in Gaussian channels. IEEE Trans Inf Theory 51(4):1261–1282. https://doi.org/10.1109/TIT.2005.844072
11. Han Z, Niyato D, Saad W, Başar T, Hjørungnes A (2012) Game theory in wireless and communication networks: theory, models, and applications. Cambridge University Press, Cambridge
12. He G, Cui S, Dai Y, Jiang T (2022) Learning task-oriented channel allocation for multi-agent communication. IEEE Trans Veh Technol 71(11):12016–12029
13. Huang C, et al (2020) Holographic MIMO surfaces for 6G wireless networks: opportunities, challenges, and trends. IEEE Wireless Commun 27(5):118–125
14. IMT Vision (2022) Framework and overall objectives of the future develoment of IMT for 2030 and beyond. Annex 3.5 to Docuement 5D/1555-E. Radiocommunication Study Groupus, International Mobile Telecommunications
15. Jamil SU, Khan MA, Rehman SU (2022) Resource allocation and task off-loading for 6G enabled smart edge environments. IEEE Access 10:93542–93563
16. Jiang W, Han B, Habibi MA, Schotten HD (2021) The road towards 6G: a comprehensive survey. IEEE Open J Commun Soc 2:334–366
17. Jo M, Maksymyuk T, Strykhalyuk B, Cho CH (2015) Device-to-device-based heterogeneous radio access network architecture for mobile cloud computing. IEEE Wireliess Commun 22(3):50–58
18. Leung KC, Li VOK, Yang D (2007) An overview of packet reordering in transmission control protocol (TCP): problems, solutions, and challenges. IEEE Trans Parallel Distrib Syst 18(4):522–535
19. Li B, Fei Z, Xu X, Chu Z (2018) Resource allocations for secure cognitive satellite-terrestrial networks. IEEE Wireless Commun Lett 7(1):78–81
20. Li L, Wang M, Xue K, Cheng Q, Wang D, Chen W, Pan M, Han Z (2021) Delay optimization in multi-UAV edge caching networks: a robust mean field game. IEEE Trans Veh Technol 70(1):808–819
21. Li Y, Huang J, Sun Q, Sun T, Wang S (2021) Cognitive service architecture for 6G core network. IEEE Trans Ind Inf 17(10):7193–7203
22. Liu S, Cui T (2017) Concepts, working principles, and applications of coding and programmable metamaterials. Adv Opt Mater 5:1700624

23. Liu F, Masouros C, Li A, Sun H, Hanzo L (2018) MU-MIMO communications with MIMO radar: from co-existence to joint transmission. IEEE Trans Wireless Commun 17(4):2755–2770. https://doi.org/10.1109/TWC.2018.2803045
24. Liu L, Zhou Y, Garcia V, Tian L, Shi J (2018) Load aware joint CoMP clustering and inter-cell resource scheduling in heterogeneous ultra dense cellular networks. IEEE Trans Veh Technol 67(3):2741–2755
25. Liu F, Masouros C, Petropulu AP, Griffiths H, Hanzo L (2020) Joint radar and communication design: applications, state-of-the-art, and the road ahead. IEEE Trans Commun 68(6):3834–3862. https://doi.org/10.1109/TCOMM.2020.2973976
26. Liu F, et al (2022) Integrated sensing and communications: toward dual-functional wireless networks for 6G and beyond. IEEE J Sel Areas Commun 40(6):1728–1767. https://doi.org/10.1109/JSAC.2022.3156632
27. Liu A, et al (2022) A survey on fundamental limits of integrated sensing and communication. IEEE Commun Surv Tuts 24(2):994–1034. https://doi.org/10.1109/COMST.2022.3149272
28. Lyu R, Cheng W, Shen B, Ren Z, Zhang H (2022) AM-SWIPT for IoE-driven 6G. IEEE Commun Mag 60(3):19–25
29. Mao T, Zhang L, Xiao Z, Han Z, Xia X-G (2024, February) Terahertz-band near-space communications: from a physical-layer perspective. IEEE Commun Mag 62(2):110–116. https://doi.org/10.1109/MCOM.004.2200429
30. Matthaiou M, Yurduseven O, Ngo HQ, Morales-Jimenez D, Cotton SL, Fusco VF (2021) The road to 6G: ten physical layer challenges for communications engineers. IEEE Commun Mag 59(1):64–69
31. Moerman A, et al (2022) Beyond 5G without obstacles: mmWave-over-fiber distributed antenna systems. IEEE Commun Mag 60(1):27–33
32. Nguyen NT, Lee K (2018) Coverage and cell-edge sum-rate analysis of mmWave massive MIMO systems with ORP schemes and MMSE receivers. IEEE Trans Signal Process 66(20):5349–5363
33. Rec. ITU-R M.2083-0 (2015) IMT vision - Framework and overall objectives of the future development of IMT for 2020 and beyond. Radiocommunication Study Groupus. International Mobile Telecommunications
34. Saad, MM, Khan MTR, Shah SHA, Kim D (2021) Advancements in vehicular communication technologies: C-V2X and NR-V2X comparison. IEEE Commun Mag 59(8):107–113
35. Sadovaya Y, et al (2022) Integrated access and backhaul in millimeter-wave cellular: benefits and challenges. IEEE Commun Mag 60(9):81–86
36. Sawant A, Lee I, Jung BC, Choi E (2021) Ultimate capacity analysis of orbital angular momentum channels. IEEE Wireless Commun 28(1):90–96
37. Shen J, Yang C, Li T, Wang X, Song Y, Guizani M (2021) Interactive artificial intelligence meets game theory in next-generation communication networks. IEEE Wireless Commun 28(2):128–135
38. Shi X, Deng N, Zhao N, Niyato D (2023) Coverage enhancement in millimeter-wave cellular networks via distributed IRSs. IEEE Trans Commun 71(2):1153–1167
39. Shreedhar T, Kaul SK, Yates RD (2022) ACP+: an age control protocol for the internet. arXiv:2210.12539
40. Sun L, Wang Y, Swindlehurst AL, Tang X (2021) Generative-adversarial-network enabled signal detection for communication systems with unknown channel models. IEEE J Sel Areas Commun 39(1):47–60
41. Tang X, Ren P, Gao F, Du Q (2017) Interference-aware resource competition toward power-efficient ultra-dense networks. IEEE Trans Commun 65(12):5415–5428
42. Tang X, Wang D, Zhang R, Chu Z, Han Z (2021) Jamming mitigation via aerial reconfigurable intelligent surface: passive beamforming and deployment optimization. IEEE Trans Veh Technol 70(6):6232–6237
43. Wang Y, Lin X (2015) User-provided networking for QoE provisioning in mobile networks. IEEE Wireless Commun 22(4):26–33 (2015)

44. Wang J et al (2012) Terabit free-space data transmission employing orbital angular momentum multiplexing. Nat Photon 6(7):488–496
45. Wang Z, Liu Y, Mu X, Ding Z (2022) NOMA inspired interference cancellation for integrated sensing and communication. In: Proceedings of ICC 2022 - IEEE international conference on communications, Seoul, pp 3154–3159. https://doi.org/10.1109/ICC45855.2022.9839026
46. Wang H, Xia X, Qi W, Liu Y, Xing Y, Qi X (2022) Survey on future mobile core network using live service migration. In Proceedings of 2022 IEEE/CIC International Conference on Communications in China (ICCC Workshops), Sanshui, Foshan
47. Wang Z, et al (2022) Extremely large-scale MIMO: fundamentals, challenges, solutions, and future directions. Accessed Mar 06, 2023. http://arxiv.org/abs/2209.12131
48. Yang N, Shafie A (2024, February) Terahertz communications for massive connectivity and security in 6G and beyond Era. IEEE Commun Mag 62(2):72–78. https://doi.org/10.1109/MCOM.001.2200421
49. Yao Z, Cheng W, Zhang W, Zhang H (2021) Resource allocation for 5G-UAV-based emergency wireless communications. IEEE J Sel Areas Commun 39(11):3395–3410
50. Yates RD, Sun Y, Brown DR, Kaul SK, Modiano E, Ulukus S (2021) Age of information: an introduction and survey. IEEE J Sel Areas Commun 39(5):1183–1210
51. You T, Zhang H, Zhang Y, Li Q, Zhang P, Yang M (2022) The influence of experienced guider on cooperative behavior in the prisoner's dilemma game. Appl Math Comput 426:127093
52. Zhang JA, et al (2021) An overview of signal processing techniques for joint communication and radar sensing. IEEE J Sel Top Signal Process 15(6):1295–1315. https://doi.org/10.1109/JSTSP.2021.3113120
53. Zhang X, Wang J, Poor HV (2021) Optimal resource allocations for statistical QoS provisioning to support mURLLC over FBC-EH-based 6G THz wireless nano-networks. IEEE J Sel Areas Commun 39(6):1544–1560
54. Zhao H, Ding K, Sarkar NI, Wei J, Xiong J (2018) A simple distributed channel allocation algorithm for D2D communication pairs. IEEE Trans Veh Technol 67(11):10960–10969
55. Zhao Z, Du Q, Sun L (2021) Network-load estimation for K-repetition grant-free access enabling adaptive resource allocation towards QoS enhancement. In: Proceedings of 2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Helsinki
56. Zhao Z, Du Q, Karagiannidis GK (2023) Improved grant-free access for URLLC via multi-tier-driven computing: network-load learning, prediction, and resource allocation. IEEE J Sel Areas Commun 41(3):607–622
57. Zhou Y, Liu H, Pan Z, Tian L, Shi JL (2016) Energy efficient two-stage cooperative multicast based on device to device transmissions: effect of user density. IEEE Trans Veh Technol 65(9):7297–7307
58. Zhou L, Wu D, Chen J, Dong Z (2018) Greening the smart cities: energy-efficient massive content delivery via D2D communications. IEEE Trans Ind Inf 14(4):1626–1634

6G Wireless Technologies

Carmen D'Andrea, Josep Miquel Jornet, Arjun Singh, Priyangshu Sen, Zabih Ghassemlooy, Stanislav Zvanovec, Shivani Rajendra Teli, and Asghar Gholami

1 Cell-Free Massive MIMO

Carmen D'Andrea

The users of a traditional cellular network experience large performance disparity between cell-center and cell-edge positions and this is one of the main drawbacks of the approach conceived in 1G to 5G wireless networks. The concept of cell-free massive MIMO (CF-mMIMO) has emerged, firstly in academia and then in

---

C. D'Andrea
Università degli Studi di Cassino e del Lazio Meridionale, Cassino, FR, Italy
e-mail: carmen.dandrea@unicas.it

J. M. Jornet
Department of Electrical and Computer Engineering, Institute for the Wireless Internet of Things, Northeastern University, Boston, MA, USA
e-mail: j.jornet@northeastern.edu

A. Singh · P. Sen
Department of Engineering, SUNY Polytechnic Institute, Utica, NY, USA
e-mail: singha8@sunypoly.edu; senp@sunypoly.edu

Z. Ghassemlooy (✉)
Optical Communications Research Group, Faculty of Engineering and Environment, Northumbria University, Newcastle-upon-Tyne, UK
e-mail: z.ghassemlooy@northumbria.ac.uk

S. Zvanovec · S. R. Teli
Department of Electromagnetic Field, Faculty of Electrical Engineering, Czech Technical University in Prague, Prague, Czech Republic

A. Gholami
Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran

V. Loscri et al. (eds.), *The Road towards 6G: Opportunities, Challenges, and Applications*, https://doi.org/10.1007/978-3-031-42567-7_3

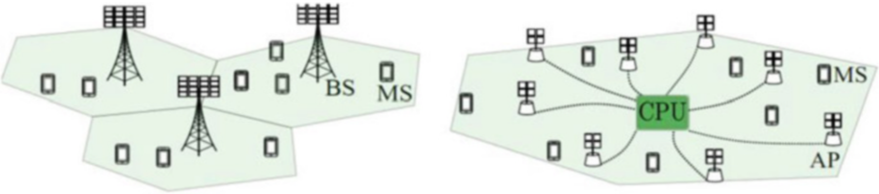


Fig. 1 Topologies: traditional cellular network on the left and CF-mMIMO network on the right

industry, to overcome this long-standing issue by providing consistent performance and seamless handover regardless of the users' positions [1–4]. By conveniently combining elements from massive MIMO, small cells, and coordinated multipoint (CoMP) with joint transmission/reception, CF-mMIMO gives rise to a cell-less architecture characterized by almost uniform achievable rates across the coverage area. For the ability of improving fairness in wireless communications, it is widely regarded as a potential physical-layer paradigm shift on the road to 6G wireless systems.

1.1 Network Topology

In CF-mMIMO, conventional base station (BSs) equipped with massive co-located antenna arrays are replaced by *many* low-cost access points (APs) equipped with few antennas, which cooperate to *jointly* serve the users. To enable such cooperation, the APs are connected to one or more central processing units (CPUs) via fronthaul links, as illustrated in Fig. 1. The size of a CF-mMIMO system can be comparable with a macro-cell and even bigger thanks to the use of multiple CPUs to connect the APs in the network. The fronthaul links between the CPUs and the APs can be realized in wired (fiber optics or radio stripes [4]) or eventually in wireless [5]. An important topic in the CF-mMIMO literature is the investigation on the case of limited capacity of the fronthaul links, some examples are references [6–8].

The distinguishing features of the originally formulated version of the CF-mMIMO are the following [1, 2]:

(a) The time division duplex (TDD) protocol is used to exploit channel reciprocity on the uplink and downlink.
(b) Uplink channel estimates are computed locally at each AP and exploited locally, which means that they are not sent on the fronthaul link and shared with the other APs.
(c) The beamforming vectors to be used at the APs for transmission/reception of the data are computed locally and not shared with the CPU.

(d) The fronthaul link is used to send data symbols on the downlink, and sufficient statistics on the uplink to perform centralized uplink data decisions; however, it is not used to share the channel estimates and/or the beamforming vectors.

Distributing the transmit/receive antennas over many cooperating APs has a twofold advantage over traditional cellular networks:

(i) It enables each UE to be located near one or a few APs with high probability and then to be ***jointly served*** by a reasonable number of favorable antennas with reduced path loss
(ii) As the serving antennas for each UE belong to spatially separated APs (which are usually seen with different angles), it brings an improved ***diversity against blockage*** and large-scale fading.

1.2 User-Centric Approach and Centralized Processing

CF-mMIMO is a recent research topic that has been gaining a lot of attention in the last few years and the original approach proposed in [1, 2], were each user is server by *all the APs* in the system, was significantly enforced with further research both in academia and industry. The two main modifications that contributed to enforce the original concept are the ***user-centric approach*** and the ***centralized processing***.

The assumption that all the APs serve all the MSs in the system may lead to some inefficiencies in the system as the size of the considered area grows. Indeed, it appears clearly pointless to waste power and computational resources at an AP to decode MSs that are very far and that are presumably received with a very low SINR. To overcome this limitation, a user-centric approach has been introduced in [3, 9]. In this approach, each user is served not by all the APs in the system, but *just by a sub-set of them* also named *AP cluster*, for example containing the ones that are in the neighborhood. Similar APs selection strategies can be used to choose the best AP cluster used to serve each user, i.e., received-power-based selection, largest large-scale-fading-based selection, policies based on energy-efficiency of the overall system. Common features between the APs' selection strategies are: (i) the clusters must be *dynamic*, i.e., they are adapted to the user's position and motion and (ii) the clusters relative to different users can be *overlapped.*

The authors of [1, 2] show that the original version of CF-mMIMO provides better performance than a small-cell system for the majority of the users in the system, thus confirming that the scheme is effective in alleviating the cell-edge user problem and in providing a more uniform service across users. These results are obtained with a simple maximum ratio transmission (MRT) on the downlink and maximum ratio combining (MRC) for the uplink and distributed processing, i.e., each AP uses only its local channel estimates to process data. Going on with research on that topic other transmission and reception schemes were introduced such as local partial zero-forcing (PZF), local minimum mean square errors (MMSE) assuming *distributed processing* [10, 11], and PZF and MMSE approaches assuming *centralized processing* [12]. The fully centralized approach, in which all the data collected by all the APs is processed in centralized manner at the CPU improves the system's performance, but it is not practical when the coverage area grows, i.e., it is not scalable [13]. A centralized approach among the APs must be jointly used with the AP-clustering to improve the performance by maintaining affordable the complexity of the approach.

Given these two add-ons to the original concept we can see some similarities and differences between CF-mMIMO and CoMP. Unlike CoMP with joint transmission/reception, which is traditionally implemented in a network-centric fashion with well-defined edges between clusters of cooperating APs [14], CF-mMIMO adopts a user-centric approach, where the above clusters are formed so that each user is served by its nearest APs [9, 13, 15]. Furthermore, as for CoMP systems, the CF-mMIMO approach greatly benefits from TDD operations, which allow to use uplink pilot signals for both uplink and downlink channel estimation. While CoMP was designed as an add-on to an existing cellular network, the deployment architecture and protocols are co-designed in a CF-mMIMO network to deliver uniform service quality.

1.3 The Communication Process

The communication process in CF-mMIMO consists in three phases:

1. ***Uplink training***: Assuming TDD the users send pilot signals to the APs and each of them, locally estimates the uplink channel. The common technique in literature assumes MMSE approach where the APs have knowledge about the large-scale fading coefficients to improve the performance of the channel estimation. Thanks to the TDD assumption, the downlink channel is the conjugate-transpose of the uplink one, thus the channel estimates obtained during the uplink training phase are then used both for downlink data transmission and for uplink data detection.
2. ***Uplink data transmission (UL)***: The original version of CF-mMIMO network can be implemented to carry out as much of the signal processing as possible in a *distributed operation*. A motivating factor is that each AP can easily be equipped with a baseband processor that is (at least) as powerful as those in the UEs, thus algorithms that can make efficient use of it should be developed. Since a key property of cell-free networks is that multiple APs are involved in the service of each user, the final data detection must be carried out at a point where the inputs from multiple APs are combined, i.e., at the CPU. An advanced uplink implementation of CF-mMIMO is a *centralized operation*, where the APs only act as relays that forward their received baseband signals to the CPU for processing. In this implementation, the CPU is a logical entity, and its tasks can be divided between many geographically distributed edge-cloud processors [16].
3. ***Downlink data transmission (DL)***: In the *distributed operation*, each AP is receiving the downlink data from the CPU and designs the locally transmitted

signal using local precoding based on the locally available channel estimates. The signals transmitted from the serving APs are coherently received at the desired users, but the interference suppression capability is limited since each AP can only suppress the interference that itself is generating. This is a limiting factor of the distributed operation since each AP is envisioned to have few antennas. As in the uplink, the most advanced downlink implementation of CF-mMIMO is a *centralized operation*, where the APs only act as relays that transmit signals that the CPU has generated and sent out over the fronthaul links. In this case, the signal is generated for each AP using these precoding vectors and the downlink data. Each AP must only be aware of its piece of the collective precoding vector, but the CPU can design the collective vector so that the different pieces fit well together. In particular, the CPU can select the precoding so that the APs can cancel out each other's interference in a way that each AP is unable to figure out individually. In the distributed operation, each AP is receiving the downlink data from the CPU and designs the locally transmitted signal using local precoding based on the locally available channel estimates.

1.4 Performance Assessment

To quantify the effective performance gain offered in CF-mMIMO systems, consider a deployment with single-antenna UEs, where MMSE channel estimation is performed, MRT and MRC are used at the APs for DL and UL, respectively, and the small-scale fading coefficients are i.i.d. complex Gaussian distributed (i.e., the usual Rayleigh fading assumption). Assume a communication bandwidth of 20 MHz; carrier frequency of 1.9 GHz; antenna heights at the APs and at the users of 10 m and 1.65 m, respectively; thermal noise with power spectral density of -174 dBm/Hz; front-end receiver at the APs and at the users with noise figure of 9 dB. The number of APs is 100, each equipped with 4 antennas, and the number of users simultaneously served in the system on the same time-frequency coherence block is 30. The APs and users are deployed at random positions on a square area of 1000×1000 m^2. The following deployments are considered for comparison:

- *User-centric cell-free massive MIMO (UC-CF)*: each user is associated to the $\mathrm{N_{UC}}$ APs that it receives with the strongest power.
- *Full cell-free massive MIMO (Full-CF)*: each user is associated to all the APs in the network.
- *Small-cells*: each user is associated to the AP that it receives with the strongest power.
- *Cellular massive MIMO (Cellular mMIMO)*: The same area is covered by 4 cells, each with one BSs equipped with 100 antennas, and each user is associated to the BS that it receives with the strongest power.

Figure 2 reports the Cumulative distribution function (CDF) of the rate per user for the considered deployments, where the UC-CF is reported with different values of

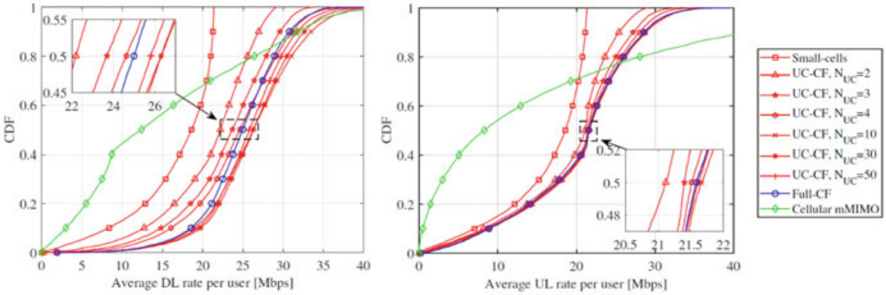


Fig. 2 Cumulative distribution functions (CDFs) of the DL and UL rate per user of the small-cells, full-CF, UC-CF and cellular mMIMO deployments

Table 1 x%-likely DL and UL rates in Mbps for the UC-CF, full-CF, small-cells, and cellular mMIMO deployments

	DL, x = 95	DL, x = 50	DL, x = 1	UL, x = 95	UL, x = 50	UL, x = 1
UC-CF, $N_{UC} = 10$	16	26.2	39.9	5.6	21.7	34
Full-CF	15.7	25	35.9	5.7	21.6	33.7
Small-cells	4.7	18.6	21.3	3.7	18.5	21.2
Cellular-mMIMO	1.4	12.4	40.2	0.13	8.3	59.3

N_{UC}. It is clearly seen that Cellular mMIMO, while providing very large rates to the lucky users, the ones that are very close to the BSs (this corresponds to the upper-right part of the figure), is largely outperformed by Full-CF and UC-CF deployments when considering the vast majority of the users (this corresponds to the lower-left part of the figure). Indeed, the figure reveals that about 97% of the users on the DL and 80% of the users on the UL enjoy much better rates with Full-CF and UC-CF deployments than with Cellular mMIMO. Additionally, the small-cells deployment outperforms Cellular mMIMO for about 70% of the users both on UL and DL, while its performance is always exceeded by increasing the number of serving APs.

This situation is again clearly represented by the numbers reported in Table 1, where the 1%, 50% and 95%-likely per-user rates in UL and DL are reported. The x%-likely rate is defined as the rate provided to the x% of the users in the system. It is indeed seen that Cellular mMIMO outperforms Full-CF and UC-CF deployments only when considering the 1%-likely per user rate, especially for the UL, while being practically not-able to serve a good share of the users. The fairness is thus preserved only when a distributed antenna system is considered. The reader should however not be led to draw the conclusion that “Cellular mMIMO does not work”. Indeed, the situation here considered is somehow extreme with simple MRT and MRC processing and with many users being served on the coherence block. In a less loaded scenario, with a smaller number of users and more advanced processing schemes such as regularized zero-forcing beamforming [17], the performance of Cellular mMIMO is restored and the gap with Full-CF and UC-CF deployments gets reduced.

1.5 Applications of Cell-Free Massive MIMO in 6G

As already said, CF-mMIMO is a promising technology on the road to 6G and it can give support in a variety of potential applications of interest in 6G scenarios thanks to its ubiquitous coverage and users' proximity. Some of these applications are briefly discussed in the following.

- ***Spatial Multiplexing of Heterogeneous Services***: The three cornerstones of 5G identified by the 3rd Generation Partnership Project (3GPP) consist of the *enhanced mobile broadband* (eMBB), *ultra-reliable low-latency communications* (URLLC), and *massive machine-type communications* (mMTC) [18]. CF-mMIMO, thanks to its ubiquitous nature, extraordinary macro-diversity gain and proximity to the users can of course increase the data rates, i.e., targeted towards the eMBB requirements. However, it is not only beneficial for the eMBB use case. By reducing the outage probability, and therefore increasing the reliability, it also provides significant benefits for URLLC compared to a single transmitter with single antenna. Due to the increased reliability, there is less need for retransmitting packets, which means lower latency. Finally, the mMTC use case also benefits from the CF-mMIMO, by capitalizing on the presence of multiple APs they are able to more easily separate the mMTC devices in the spatial domain, and thereby to detect them at the system's random-access stage.
- ***Multi-Access Edge Computing***: User's demands for additional computational resources and network's need of supporting URLLC applications converge to a paradigm called *multi-access/mobile edge computing (MEC)*. Its goal is to bring the application (and all its computing) closer to the user. CF-mMIMO, thanks to its dense distributed topology and users' proximity, can greatly facilitate the implementation of MEC providing the computation offloading, by enabling mobile devices to delegate either all or part of their computational tasks to multiple APs, each of which may be equipped with a MEC edge server. Moreover, the CPU of a CF-mMIMO, which is generally equipped with a more powerful server, may serve as a backup edge computing, to give, in turn, computation offloading support to the APs or directly to users with less strict latency requirements [19].
- ***Support to millimeter-wave communications and user mobility***: The use of millimeter-wave frequencies guarantees large bandwidths and thus large data rates, for this reason it should be mentioned on the road to 6G. These frequencies, above 30 GHz, suffer propagation impairments, such as high path loss, increased blockage, and severe atmospheric absorption. Practically, millimeter-wave frequencies can be used in urban and indoor scenarios with low-range communications and in conjunction with sub-6 GHz frequencies. In the millimeter-wave bands, where achieving high spectral efficiency is not a concern thanks to the large available bandwidth, the distributed cell-free operation can certainly better cope with the *hostile propagation environment* experienced at such high frequencies, as the presence of many serving APs in the users' proximity enhances coverage and link reliability. An aspect that deserves a closer look in

this context is the impact of the users’s mobility on the performance. As the user moves, the propagation channels towards the APs are subject to time variations and the user-AP associations need to change dynamically [20].

2 Terahertz Communications for 6G Networks: How Far Are We?

Josep Miquel Jornet • Arjun Singh • Priyangshu Sen

2.1 Introduction

The sixth generation (6G) of wireless cellular networks is going to be realized within the next 5 years, i.e., 10years after the 3rd Generation Partnership Project (3GPP) defined in Release 15 the first iteration of the fifth-generation (5G) networks. As of today, both academia and industry agree on one terabit-per-second (Tbps) as the target big data rate for 6G systems. The very first question is *how much bandwidth is needed to achieve this goal?*

In Fig. 3, we illustrate the bandwidth as a function of both the modulation order and the number of parallel channels. Clearly, unless we can support very high

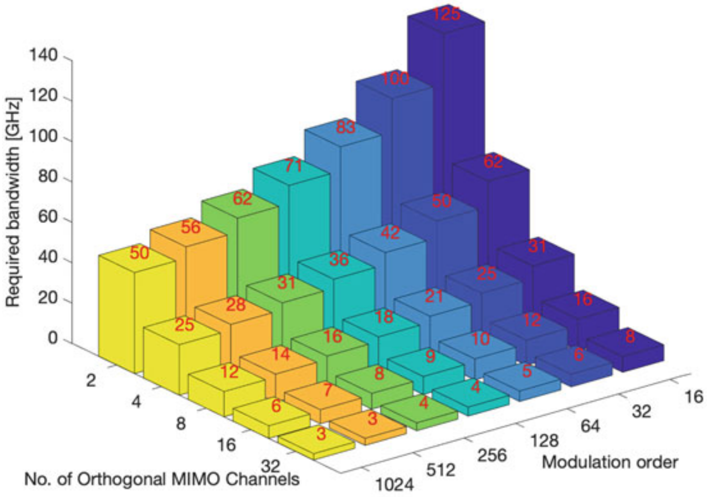


Fig. 3 Required bandwidth to achieve a data rate of 1 Tbps, as a function of the number of orthogonal channels and the modulation order per sub-carrier when a MIMO-OFDM system is considered

modulation orders over a very high number of parallel channels, the bandwidth requirements are going to be in the order of tens of gigahertz.

The second question is then *where are we going to find such bandwidth?* Currently, the majority of wireless communication and sensing systems are allocated at frequencies under 100 GHz, such as 5G NR Frequency Range 2 (up to 71 GHz), IEEE 802.11ay (between 57 and 71 GHz), and vehicular radars (71–86 GHz). As a result, there is no contiguous band that could accommodate the requirements for 6G networks. If bandwidth is the only driver, one could always adopt optical wireless systems (i.e., 200 THz for infrared systems or up to 600 THz for visible blue light systems). This is a path that should certainly be explored, but one last question immediately comes up: *what is in between the two ends of the spectrum?*

This is the terahertz band, the frequency range broadly defined between 100 GHz and 10 THz and, today, the core of many active research groups [21–25]. The applications of the terahertz band have drastically evolved in the last decade. Originally, the rather low power of terahertz transceivers suggested that terahertz communications were only applicable in the context of very short-range networks, such as multimedia kiosks utilized to quickly download large multimedia content on a mobile device [26]. With improvements in both technology as well as on the understanding of the propagation of terahertz waves, much more exciting applications are possible. For example, today, terahertz communications are considered for next-generation WiFi and dense cellular network applications that require very high data rates, such as wireless extended reality (XR) headsets [27]. Moreover, terahertz communications are also being considered for backhaul applications, as a potential technology to bring ultra-broadband connectivity to rural areas thus bridging the digital divide [28, 29]. At the forefront of the field, today, terahertz communications are also being explored as an enabler of much faster satellite communication networks [30, 31].

Beyond communications, or, in fact, before communications, the terahertz band has been widely considered for sensing applications. The very small wavelength of terahertz waves –from 3 mm down to 30 μm– combined with the very large available bandwidth available at these frequencies, can lead to radar and localization systems with much higher resolution and accuracy than wireless communication systems at microwave and even millimeter-wave frequencies, with deeper penetration than optical sensing systems. In addition, the low but non-negligible photon energy of terahertz waves –from 0.41 millielectronvolts (meV) up to 41 meV– leads to very unique electromagnetic signatures in different elements (from gases in the atmo sphere to nano and bio-materials) [32–34]. Such signatures can be leveraged, for example, in target classification. In a similar context, at this precise moment, there are multiple satellites orbiting the Earth that incorporate terahertz sensors. These are being utilized both for atmospheric studies and to receive signals from space that can help us study the origin of the universe and life [35].

Moreover, besides communications and sensing at the macroscale level, there is an entire set of applications of the terahertz band at the nanoscale level. Indeed, the very small wavelength of terahertz waves enables the development of very small antennas. Moreover, when leveraging new materials such as graphene, terahertz transceivers and antennas with micrometric footprints can be developed [36, 37]. The resulting nano-radios can be integrated in embedded nano-systems or nanomachines and are at the basis of transformative applications. For example, nanomachines with biological, chemical, and physical nanosensing capabilities can be interconnected in wireless nanosensor networks [38] and, ultimately, connected to macro-scale networks in the Internet of Nano-Things [39]. Alternatively, nano-radios can be found at the basis of wireless networks on chip (WNOC) in massive multi-core processing architectures [40]. Looking ahead, there is an opportunity for the terahertz photons to be used in quantum wireless communication networks, as a middle ground between microwave and optical qubits.

Despite these promising applications, the terahertz band had been largely disregarded for decades, mostly due to the lack of device technologies that could effectively exploit it, but also due to some pre-conceived misconceptions that discouraged its adoption. However, major progress in the large decade on all fronts has brought terahertz communication and sensing systems much closer to commercial applications. In this chapter, we follow a bottom-up approach, from materials and devices all the way to spectrum policy, to provide an updated view on terahertz communications and sensing systems, highlighting what has already been solved and what the immediate next challenges are. In Sect. 2.2, we present the state of the art in terahertz hardware technologies, highlighting the fight that the terahertz technology gap is almost closed. In Sect. 2.3, we review the basics of terahertz wave propagation and provide an updated view on the most recent channel modeling findings. In Sect. 2.4, we discuss critical aspects at the physical of terahertz communication networks, including ultra-massive MIMO and wavefront and waveform design. In Sect. 2.5, we describe the critical challenges at the higher layers of the protocol stack that need to be addressed to ensure that the user experience through- puts in line with what the physical layer supports. Finally, in Sect. 2.6, we highlight the critical role that spectrum policy and standardization will play towards ensuring that terahertz systems are ready for 6G, and we conclude the chapter in Sect. 2.7 (Fig. 4).

2.2 Terahertz Materials, Devices and Testbeds

For decades, the terahertz band was referred to as the terahertz gap because of the lack of device technologies able to support communication and sensing applications at these frequencies. However, the situation is much different today and, as we summarize in the next sections, there are multiple solutions to realize the critical hardware building blocks of a terahertz system, including analog front-ends, antenna systems and digital signal processing back-ends.

Federal Communications Commission FCC 19-19

Before the
Federal Communications Commission
Washington, D.C. 20554

In the Matter of

Spectrum Horizons ET Docket No. 18-21

James Edwin Whedbee Petition for Rulemaking to Allow Unlicensed Operation in the 95-1,000 GHz Band RM-11795 (Proceeding terminated)

FIRST REPORT AND ORDER

Adopted: March 15, 2019 Released: March 21, 2019

IEEE Standard for High Data Rate Wireless Multi-Media Networks

Amendment 2: 100 Gb/s Wireless Switched Point-to-Point Physical Layer

IEEE Computer Society

Fig. 4 THz through the years: (**a**) Expansion in THz use cases from the nano to the macro, fueled by developments in (**b**) device technology (Sect. 2.2) and channel studies (Sect. 2.3); (**c**) communication algorithms (Sect. 2.4); (**d**) networking protocols (Sect. 2.5); and (**e**) efforts in standardization (Sect. 2.6)

2.2.1 Analog Front-Ends

The analog front-end is in charge of generating, modulating, filtering, and amplifying signals at terahertz frequencies with multi-gigahertz bandwidths. There are mainly three approaches to build terahertz front-ends:

- In the **electronic approach**, the limits of the devices and designs used in microwave and millimeter-wave frequency systems are pushed towards the terahertz band. As of today, CMOS-based technologies have been used to build systems at frequencies approaching 300 GHz, with transmit power approaching

1 mW [41]. To increase the power output to tens of milliwatts, amplifiers based on III–V semiconductor materials can be utilized [42]. III–V semiconductor materials can also be utilized to build Schottky diodes used in frequency-multiplying chains with transmit power in the order of up to a few hundred milliwatts [43]. While this power figure might seem low, this is in fact a world record. In general terms, electronic terahertz devices can provide high transmission power, but at the cost of non-line arities and phase noise.

- In the **photonics approach**, the limits of devices and architectures used in optical communication systems are pushed down in frequency towards the terahertz band. There are mostly three fundamental techniques to do so. First, quantum cascade lasers (QCLs) are designed to emit terahertz photons, by leveraging intra-band transitions in III–V semiconductor materials [44]. When cooled to cryogenic temperatures, these can emit powers approaching 10 mW. Second, photoconductive antennas [45], which consist of a metallic terahertz antenna with a photoconductor material at the gap, can be utilized to downcovert an optical pulsed signal to terahertz frequencies. Such devices can be commonly found in time-domain spectroscopy platforms, with output powers approaching 1 mW. Last but not least, the use of frequency difference generation is becoming increasingly popular [46]. In this case, a photomixer is used to multiply a modulated optical laser with a second non-modulated laser at a second wavelength shifted from the first laser by the target terahertz carrier signal. While the generated terahertz power can still not meet that of electronic frequency-multiplied sources, the higher linearity, the much lower phase noise, and the possibility to leverage optical modulators make this a promising future path.
- In the **plasmonic approach**, the goal is to leverage the properties of plasma waves and surface plasmon polariton waves (SPP) to build devices that intrinsically operate at terahertz frequencies. Among others, plasma waves –or oscillations of electrical charges in a material– at terahertz frequencies can be excited in the two-dimensional electron gas (2DEG) channel of a high-electron-mobility transistor (HEMT), due to the so-called Dyakonov-Shur instability. Such a transistor can be built with III–V semiconductor materials [47] and/or with two-dimensional materials such as graphene [48]. Besides signal generation, the same structure can provide high-speed modulation of the generated plasma wave in frequency and amplitude [49]. Moreover, again by leveraging graphene, direct phase modulation can be achieved by means of a tunable plasmonic waveguide [50]. In broad terms, the size of plasmonic devices is very small, in the order of hundreds of nanometers to a few micrometers, which leads to a relatively very low generated power (approaching 1 μW). However, their very small footprint supports both their embedding in nanomachines to enable the nanoscale applications of the terahertz band as well as their integration in larger numbers to build highly functional on-chip arrays [51]. Compared to the electronic and photonic approaches, the plasmonic approach has been much less explored, resulting in a high-risk high- reward opportunity.

2.2.2 Antenna Systems

As in any wireless communication system, antennas are needed to convert on-chip electrical currents into free-space propagating electromagnetic waves at the transmitter and perform the reciprocal function at the receiver. Moreover, antennas can also be found along the channel, acting as reflecting surfaces. Fundamentally, classical antenna theory remains valid at terahertz frequencies, but there are some caveats. First, the very small wavelength of terahertz radiation leads to very small terahertz antennas. For example, a resonant dipole antenna at 1 THz is approximately 150 μm and has the conventional doughnut-shaped radiation diagram. Correspondingly, in reception, the very small size of the dipole results in a very small effective area or aperture, leading to very high spreading losses (more in Sect. 2.3). Making the antenna larger so as to increase its effective area automatically leads to a more directional radiation pattern. This is why directional antennas are commonly used at terahertz frequencies. For example, as of today, horn antennas or even small dish antennas with gains ranging from 20 dBi to 55 dBi are commercially available.

Another way to achieve high gain is by building antenna arrays. Antenna arrays offer the advantage of being able to program the radiation diagram by controlling at least the phase if not also the amplitude at every antenna. The very small size of individual terahertz antennas allows in fact their dense integration in very small footprints. Besides the radiating elements, an array needs to integrate at least also the control elements (e.g., phase shifters/time delays and amplitude controllers). Moreover, if the antenna needs to support multiple input multiple output (MIMO) communications, it will require multiple front-ends, up to one per antenna, but usually one per sub-group of antennas [52]. As of today, designs with up to 16 streams each feeding either one or a small group of antennas have been demonstrated when following an electronic approach [53].

In addition to antennas, lenses can be utilized to control the radiated terahertz signals. Lenses can be utilized to focus the signal at a distance or to generate different types of wavefronts, such as non-diffracting Bessel-beams [54, 55] and self-accelerating Airy beams [56]. These types of wavefronts are particularly relevant in the near-field, as we will discuss in Sect. 2.3. Moreover, besides dielectric lenses, these functionalities can also be implemented by means of metasurfaces. Metasurfaces are arrays of meta-atoms, or custom-designed electromagnetic elements whose size is much smaller than the wavelength [57]. As with antenna arrays, programmable metasurfaces outperform fixed dielectric lenses as they can be tuned to implement different functionalities by inputting specific magnitude and phase distributions.

When adopting the plasmonic approach to build analog front-ends, plasmonic nano-antennas are needed [37]. The main advantage of plasmonic nano-antennas is that they are significantly smaller than the wavelength. While this leads to a smaller effective area, it also opens the door to their very dense integration [58]. Moreover, due to the small size of the plasmonic front-end itself, arrays of integrated plasmonic

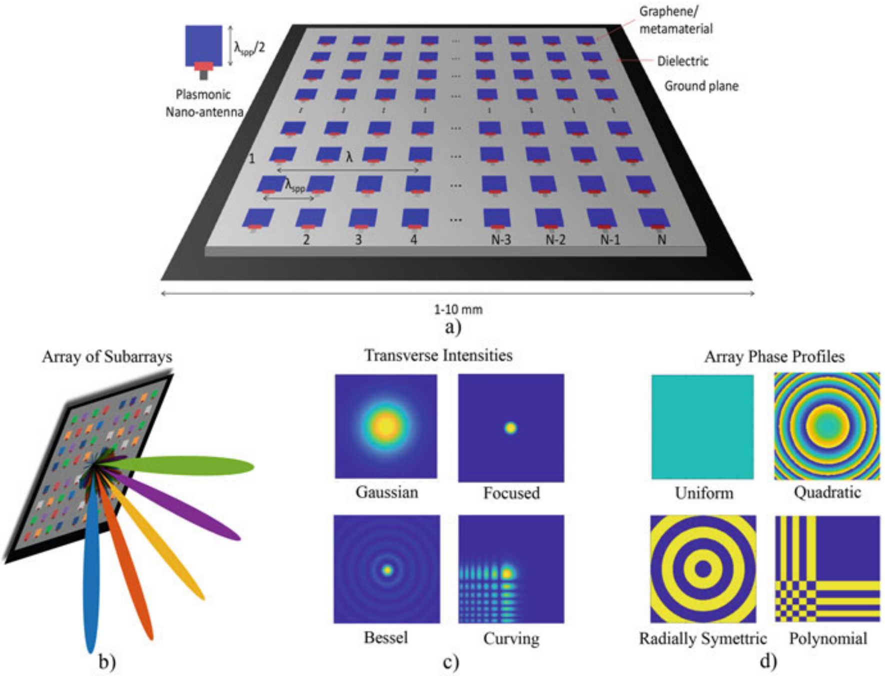


Fig. 5 The plasmonic array powered through individual, nano-sized front ends. (**a**) The plasmonic array can be densely packed to mimic a deterministic metasurface response that can allow the creation of: (**b**) simultaneous spatial streams through an array of sub-arrays architecture; (**c**) complex transverse intensities for both far-field and near-field operations thanks to the ability to create complex phase profiles; (**d**) examples of phase profile that can be created at the array surface, which allow the custom beam shaping possibilities as discussed in Sect. 2.4

front-ends and plasmonic nano-antennas can be made [51]. The fact their size is smaller than the wavelength and that each array element provides its own power and independent phase and amplitude control results in the possibility to perform not only beamforming as with metallic antenna arrays, but also wavefront engineering as with programmable metasurfaces. As mentioned at the beginning of this section, all these structures can be designed to operate in transmission, reception or reflection [59] (Fig. 5).

2.2.3 Digital Back-Ends

The main motivation to move to terahertz frequencies is its enormous bandwidth. Accordingly, a digital signal processing (DSP) engine able to exploit such bandwidth is needed. Common to optical (wired and wireless) systems, a major bottleneck in the DSP engine is posed by the digital-to-analog and analog-to-digital converters (DACs and ADCs, respectively). As of today, data converters

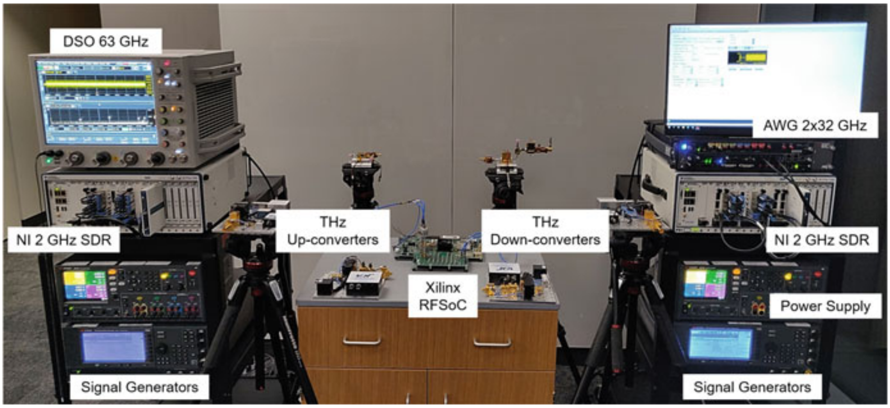


Fig. 6 The TeraNova testbed at Northeastern University: an integrated testbed for ultra-broadband communication networks in the THz band. The platform consists of several transmit and receive nodes with analog front-ends able to operate at frequencies up to 1.05 THz (i.e., true THz), antennas with gains ranging from 20 dBi up to 55 dBi, and DSP engines able to process up to 32 GHz of bandwidth per channel [61]

with sampling frequencies of up to 256 Gigasamples-per-second (GSaps) can be found in commercial laboratory-grade equipment. These data converters can thus operate with bandwidths approaching 100 GHz. However, their size, cost, and thermal requirements limit their application to very specific setups, far from what a handheld device could afford. Alternatively, highly parallelized DSP engines are being developed. In particular, separate sub-channels can be independently processed by much slower data converters and (orthogonally) multiplexed in frequency. Following this approach, we have recently demonstrated what as of today is the fastest software-defined-radio platform for wireless communications, able to process in real-time 8 GHz of bandwidth, by multiplexing four 2-GHz-wide channels in frequency [60].

Conclusion As of today, there are multiple demonstrated solutions relating to analog front-ends, antenna systems, and digital back-ends. By leveraging them, multiple terahertz experimental testbeds (e.g., see Fig. 6) have been built and are being utilized to study the terahertz channel, test new physical layers, and identify the needed solutions to build actual terahertz networks. Therefore, we can state that the terahertz technology gap is almost closed!

2.3 Terahertz Propagation and Channel Modeling

The terahertz band can provide users with a very large transmission bandwidth, at the cost of very high propagation losses. By properly understanding their origin, techniques to minimize such losses can be adopted. In this section, we first review the basics of terahertz propagation, and then highlight recent results relating to different meaningful application scenarios.

2.3.1 Propagation Basics

There are three main phenomena affecting the propagation of terahertz waves [62]:

- **Molecular absorption**: At terahertz frequencies, different molecules, but particularly water vapor, internally vibrate when excited with terahertz radiation. As a result, the electromagnetic energy of the wave is converted into kinetic energy in the molecules. The molecular absorption loss accounts for this phenomenon. Interestingly, however, molecular absorption is highly frequency selective: at some frequencies, the absorption can exceed hundreds of dBs, whereas, at some other frequencies, absorption is negligible. The position of such frequencies is fixed by the structure of the molecules, but their width and intensity depend on different parameters, including the number of molecules found along the path, the temperature, and the pressure. Mathematically, the molecular absorption losses, *Aabs*, can be computed from the Beer-Lambert loss as:

$$A_{abs}\left(f,d\right)=e^{i,g}\sum k^{i,g}(f)d, \tag{1}$$

where f and d stand for frequency an distance, respectively, and $k^{\mathrm{i,g}}$ is the absorption coefficient for the isotopologue i of gas g, given by:

$$k^{i,g}(f)=\frac{p}{p_0}\frac{T_{STP}}{T}Q^{i,g}\sigma^{i,g}(f), \tag{2}$$

where p stands for pressure, p_0 is the reference pressure (1 atm), T_{STP} is the standard-pressure temperature (273.15 K), $Q^{i,g}$ is the molecular volumetric density, and $\sigma^{i,g}$ is absorption cross-section. The latter is modeled as

$$\sigma^{i,g}(f)=S^{i,g}G^{i,g}(f), \tag{3}$$

where $S^{i,g}$ is the line intensity and $G^{i,g}$ is the line shape. The line shape depends on the type of system. For example, from [62], for systems whose pressure is above 0.1 atm, the line shape is governed by collisions among molecules of the same type and an appropriate line shape for low energy photons (i.e., terahertz waves) is the Van Vleck-Weiskopf shape, but this would be different for example in high-altitude systems (like in non-terrestrial networks). Utilizing these equations and the parameters given in the HITRAN (HIgh resolution TRANsmission molecular absorption) database [63], we illustrate the molecular absorption across frequency in Fig. 7 for different altitudes, where clearly the absorption lines can be seen. In the majority of cases, users will operate in the absorption-free windows. Even

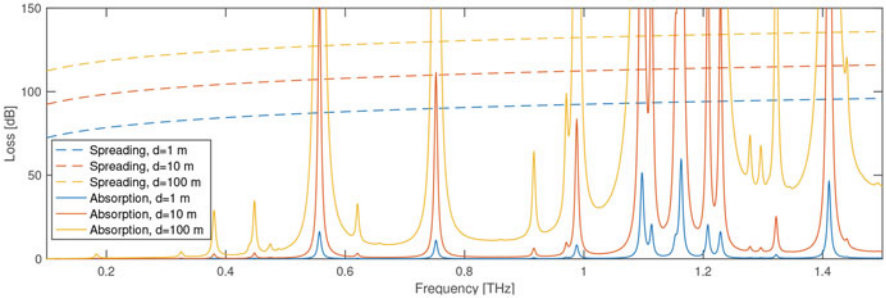


Fig. 7 Spreading loss (dashed lines) and molecular absorption loss (solid lines) for frequencies ranging from 0.1 to 1.1 THz and three different distances (1, 10 and 100 m)

in that case, users will experience a distance-dependent bandwidth resulting from the broadening of the absorption peaks as more molecules are encountered over an increasing distance. This is a phenomenon that can be leveraged for spatial multiplexing [58, 64] or for physical-layer security purposes [65, 66]. Finally, it is relevant to note that the radiation absorbed by vibrating molecules is ultimately re-emitted out of phase resulting in what is referred to as the molecular absorption noise or emissivity of the channel [67]. Different from the traditional independent additive white Gaussian noise, molecular absorption noise is additive and Gaussian, but colored and correlated with the transmitted signal.

- **Spreading losses**: As discussed in Sect. 2.2, the very small wavelength of terahertz signals leads to very small antennas, unless directional antenna designs are adopted. The spreading losses, also known as free-space path losses, account for the fraction of the transmitted power that the receiver can intercept when both transmitter and receiver use ideal omnidirectional antennas. Mathematically, the spreading losses A_{spr}, are given by

$$A_{spr}\left(f, d\right) = \left(\frac{\lambda}{4\pi d}\right)^2, \tag{4}$$

where λ is the signal wavelength. It is relevant to note that, strictly speaking, this is not a channel effect, but an antenna effect. Together with the very high molecular absorption losses at some specific frequencies mentioned before, this is the reason why traditionally people have associated the terahertz band with extreme losses and, thus, very short communication distances. Certainly, this is not the case as, by properly selecting the frequency band of operation and adopting directional antennas at the transmitter and/or the receiver, the losses can be easily avoided.

- **Blockage**: Besides interacting with gaseous molecules, terahertz radiation interacts with different types of materials (in liquid or solid state). For example, terahertz radiation is reflected by metals and other conducting materials. Such reflections might be in the specular direction, or diffusely scatted if the surface roughness of the obstacle is comparable to the wavelength. In addition, terahertz radiation is partially absorbed by wood, concrete, and different types of plastics. The combination of these phenomena with the fact that directional beams are used to compensate for the aforementioned spreading losses leads to blockage of the signals. This is perhaps the real bottleneck from the physics propagation perspective. For example, all the long-range terahertz links demonstrated to date (e.g., over to 2 km [28]), require perfect line of sight. This is what motivates the adoption of antenna structures that operate in reflection, as those discussed in Sect. 2.2.

2.3.2 Indoor and Outdoor Scenarios

With these properties and their mathematical models at hand, one can predict the actual terahertz channel in diverse scenarios. This can be done through analytical models, full-wave electromagnetic simulations or ray-tracing solutions. However, there are many elements in the propagation environment that can change, which would require rerunning the models as many times as needed. This is not practical, and this is why stochastic channel models are needed. Those channel models can be again derived from synthetic channel realizations or, preferably, from experimental data sets.

Thanks to the progressive closure of the terahertz technology gap discussed in Sect. 2.2, in the last 5 years there have been multiple experimental campaigns aimed at measuring different relevant indoor and outdoor scenarios [68–70]. Some of the key aspects to highlight are as follows.

- In indoor scenarios, multi-path propagation (including even double reflections) is likely and highly contributes to the total channel response. Such multi-path is usually generated by walls or obstacles near the line-of-sight path, mostly because of the directional antennas needed at the transmitter and the receiver. As a result, typical delay spread values are in the order of a few nanoseconds.
- The situation is similar outdoors but, in this case, weather plays a critical role. Contrary to popular belief, terahertz waves can propagate in rain and snow, but these introduce both additional losses (from 3 dB over 50 m in rain to 10 dB over the same distance with snow [71]), and also lead to a different power delay profile. Ultimately, these phenomena motivate the adoption of highly-adaptive physical layer solutions, such as those that we discuss in Sect. 2.4.

2.3.3 Extreme Communication Scenarios

As mentioned in the introduction, the applications of terahertz communication and sensing systems can expand multiple scales, from the nanoscale all the way to satellite networks. In the last decade, there have been several works aimed at studying the terahertz channel mostly in two meaningful application scenarios, namely, wireless networks on chip and intra-body wireless nanosensor networks. On-chip, the main challenge for terahertz wave propagation is not absorption (the humidity level within chips is extremely low) or spreading losses (the target link distances are the millimeter scale mostly), but multi-path propagation generated by the different material layers in common CMOS integrated-circuit fabrication processes as well as by other components found on-chip, such as the processing cores themselves [72]. Once again, it is that multi-path and the resulting delay spread that guides the development of the physical layer.

When it comes to intra-body terahertz systems, the key aspect to take into account is that the human body is not a mere obstacle or not even a layered material, but a non-uniform assembly of different types of cells and other elements in between [73]. For example, red blood cells are 10–20 μm in diameter, which is much smaller than the wavelength of the lower end of the terahertz band, but start to become very noticeable as we move towards true terahertz frequencies (1 THz and beyond). Moreover, different types of cells exhibit different absorption coefficients. What is worse, different from the absorption from gaseous molecules, which can freely vibrate, molecules in liquids and solids cannot freely move, but they experience friction, and this friction leads to a temperature increase or, broadly speaking, photothermal effects [74]. The study of such photothermal effects is critical not only in the context of nanoscale applications but also to provide guidance to the regulation of terahertz systems, as we will discuss in Sect. 2.6.

On the other extreme, the study of terahertz communication and sensing systems for space applications is becoming increasingly popular. The same fundamental properties apply, but the scenario is significantly different [75]. For example, properly choosing the operation frequency band is critical in the design of up and down-links (between the ground users and satellites) as such signals need to traverse the atmosphere. The volumetric density of different molecules, the temperature, and the pressure change through the atmosphere make the modeling of absorption more challenging. Above the atmosphere, not only there is no absorption, but the atmosphere itself can serve as an excellent filter to prevent eavesdropping or jamming from the ground, again if the right frequency is selected. However, the presence of space debris leads to again, multi-path propagation and in the worst case blockage in inter-satellite links [76]. It is worth mentioning, however, that despite the challenges, the terahertz channel is still more favorable than the optical wireless channel in space, as the much smaller wavelength of optical signals leads to a higher likelihood for blockage and reduced misalignment tolerance [77].

Conclusion In light of not only physical models but extensive experimental channel measurement campaigns, it is clear that terahertz waves propagate much better than what was traditionally believed. Nevertheless, the peculiarities of the terahertz channel such as the distance-dependent bandwidth due to molecular absorption or the high impact of blockage require the development of tailored communication and networking techniques.

2.4 Ultrabroadband Communication Techniques

With the device capabilities at hand and the channel peculiarities in mind, physical layer solutions that can effectively utilize the terahertz band are needed. Many of the existing solutions implemented at lower frequencies can be simply scaled up, but there are new opportunities that only appear at terahertz frequencies. This is what the research community is pursuing.

2.4.1 Ultra-massive MIMO Communications

As discussed in Sect. 2.2, terahertz antennas can be very small, i.e., millimetric and even sub-millimetric at true terahertz frequencies. As a result, large numbers of antennas can be integrated in small footprints. When following the electronic approach, the size and power requirements of the analog front-ends hampers the possibility of utilizing one front-end per antenna element and, thus, limits the maximum number of parallel streams that can be realized. Instead, each analog front-end in the array controls a sub-group of antennas, in what is known as the array of sub-arrays architecture [52]. Effectively, an array of sub-arrays is equivalent to an array of directional antennas, where the beam for each directional antenna can be tuned [78].

Beyond this architecture, when adopting the plasmonic approach, each element in the array integrates a plasmonic source, a modulator, and the actual antenna. This is what opens the door to ultra-massive multiple-input multiple-output (MIMO) systems [58]. Ultra-massive MIMO is not just MIMO or massive MIMO with more antennas, but a collection of technologies and techniques that can be leveraged to maximize the achievable data rate over the terahertz band. In a graphene-based plasmonic array, elements can be virtually grouped, continuously or in an interleaved fashion, to generate different types of overlapping sub-arrays. Those sub-arrays can be utilized for beamforming or for spatial multiplexing, as with traditional MIMO or massive MIMO systems.

More excitingly, by leveraging the tunability of graphene, individual antennas or different sub-arrays can be tuned to operate at different frequency sub-bands within the terahertz range, without physically modifying the array. This also opens the door to frequency multiplexing. This is particularly meaningful because, in the majority of cases, highly directional terahertz links do not support many parallel orthogonal channels and, thus, introducing control on the frequency can be leveraged for multiplexing purposes. Beyond this, as we explain in the next section, there are other opportunities that arise relating to the manipulation of the signal wavefront.

2.4.2 Wavefront Engineering in the Near Field

Whether densely-integrated antenna arrays or fixed directional antennas, the radiating structures used in macroscale terahertz applications are going to be much larger than the wavelength. At this point, it is relevant to note that the far field of a radiating structure depends on the antenna size as $2D^2\ \lambda$, where D is the largest antenna dimension [79]. Accordingly, for example, the far field of a 10-cm antenna array at 130 GHz starts at 8.67 m. The far field of the same antenna size at 1 THz does not start until 66.67 m. As a result, traditional beamforming strategies, which imply aplane wave assumption with a uniform phase and where the spreading effect results in a Gaussian intensity [79], including those proposed for terahertz systems, are inaccurate [59].

Rather than being a problem, this opens the door to more exciting opportunities that are only possible in the near field. One of the near-field solutions that has been relatively well-explored is beamfocusing [80, 81]. With this technique, the phase at the different elements of an array are set to emulate the phase distribution of a dielectric lens that concentrates the signal strength on a given point in the near field. The main challenge associated with this technique is that the focusing occurs at a point and, thus, any type of mobility would trigger a recalculation of the beamfocusing weights (as opposed to the ability to communicate along a beam or direction in the far field).

Beyond beamfocusing, non-diffraction wavefronts can be utilized in the near field, including:

- **Bessel beams** can focus the signal strength along a line or direction in the near field [54, 59, 82], matching what beamforming can do in the near field. In addition, Bessel beams exhibit an interesting self-healing property: even if the center spot of a Bessel beam is blocked by an obstacle, the Bessel beam can regenerate itself [83].
- Similarly, **Airy beams** can focus the signal strength along a curving line of propagation [56]. This property can be leveraged, for example, to bend around obstacles or corners.

In both cases, mobility has a lower impact than beamfocusing as only when the receiver moves off the beam, the beam needs to be recomputed. In addition, both techniques offer a practical way to overcome the presence of obstacles (see Fig. 8), which we discussed in Sect. 2.3 is perhaps the main remaining challenge that cannot be addressed just with higher gain or careful frequency resource allocation.

Moreover, on top of Gaussian, Bessel and Airy beams, one can introduce and modulate orbital angular momentum (OAM). Specifically, a beam that is said to have OAM manifests a spiral phase in the transverse direction, resulting in a helical wavefront and a phase singularity (a zero-intensity vortex) in the center. Overlapping beams that follow helical modes define an orthogonal basis (see Fig. 8). This can be leveraged in different ways: different streams can be sent along different OAM modes, each one with its own amplitude and/or phase modulation; or one stream can be sent by encoding different symbols in different OAM modes.

Finally, it is relevant to note that these techniques are not fundamentally new (they have been leveraged in optics mostly for sensing applications), but for the first time, they become both practical and advantageous for communications when we move to terahertz frequencies. Moreover, while usually these are implemented by

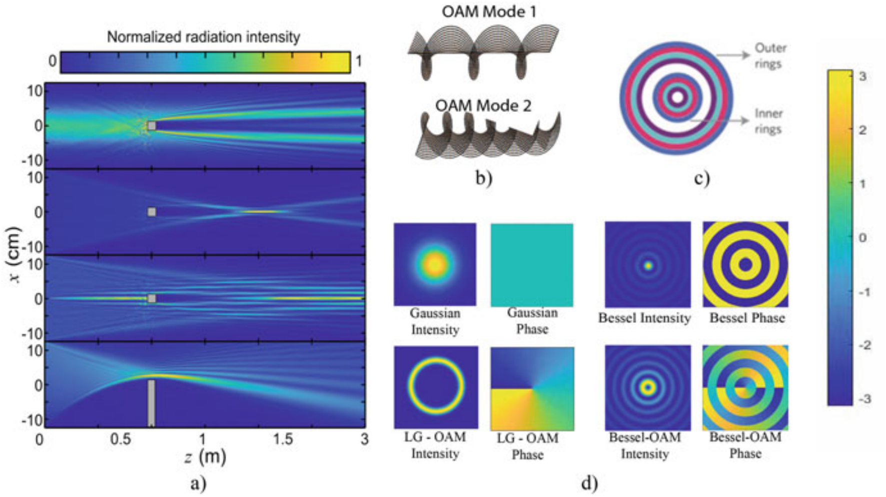


Fig. 8 Wavefront engineering and several options: (**a**) The response of different beams to obstacles, in which Bessel and Airy beams (third and fourth) have superior healing capability. (**b**) The introduction of OAM modes allows simultaneous streams that propagate in concentric rings, an example of which is seen in (**c**) the outer rings correspond to higher OAM modes; (**d**) The spiral mode of OAM can be introduced on top of the regular phase profile required to create such beams. Color bar shows the phase in radians

means of different type of lenses or phase plates (e.g., axicons to generate Bessel beams or spiral phase plates to generate OAM modes), all of them can be generated by an array of elements with phase/delay and amplitude control, such as those discussed in Sect. 2.2.

2.4.3 Ultrabroadband Waveforms, Modulation and Coding

Together with the spatial design and control of terahertz signals (wavefronts), we need to design and control terahertz signals in time (waveforms).

As the starting point, traditional modulations have been upscaled and tested at terahertz frequencies. These include both fundamental single carrier modulations (e.g., M-PSK, M-QAM), multi-carrier modulations (e.g., OFDM) as well as more sophisticated modulations that have become popular in 5G systems, such as DSF-Spread-OFDM and OTFS [84–86]. While these techniques work, they do not particularly exploit or benefit from any of the peculiarities of the terahertz band channel.

Focusing then on modulations that can only work or perform particularly well because of the terahertz channel, we need to distinguish between the short-range and the medium- to long-range terahertz channel. Over short distances (i.e., below 1 m), molecular absorption is negligible. As a result, the terahertz band comes across as an almost 10-THz-wide band. In this scenario, the transmission of femtosecond-long pulses, whose bandwidth is in the order of a few THz, can be leveraged. For example, in [87], in time-spread on-off keying (TS-OOK), pulses are transmitted by following an asymmetric on-off keying modulation and spread in time. This technique is has minimal complexity, the hardware to support has been long tested in time-domain terahertz spectroscopy platforms, and offers the flexibility to support either Tbps single links or very large number of parallel links at multi-Gbps each.

As we increase the communication distance, molecular absorption breaks down the bandwidth into multiple disjoint windows. Pulse-based modulations can be tailored to operate in different windows [88, 89]. Beyond that, hierarchical bandwidth modulations [90, 91] allow the transmission of multiple streams within a single directional transmission beam to users that are at different distances and, thus, not only experience a different signal-to-noise ratio (SNR), but also different bandwidth. Similar to conventional hierarchical or concatenated modulations, different users can decode the same stream with different modulation orders depending on their SNR. However, differently from those, near users can benefit from the larger bandwidth by refreshing the high-order modulation more frequently, while the channel will filter out those modulations to the far away user. Aimed at overcoming the hardware limitations and not only the channel, an optimized M-APSK scheme is proposed in [92] for reliable, low-PAPR, and spectrally efficient ultra-broadband THz communication. The modulation design is based on arranging the constellation points in a number of concentric rings, similar to PSK. The number of points in each ring proportionally varies with its amplitude, maintaining the same minimum distance within adjacent points. Thus, there are fewer points in an inner circle compared to the outer.

Besides increasing the data rates, there are other techniques that are being explored. For example, direct-sequence spread spectrum (DSSS) has been proposed [93] as a way to securely share the spectrum not only with other communication users but also with passive users of the terahertz spectrum (more in Sect. 2.6). Similarly, chirp spread spectrum (CSS), a waveform that is commonly used both in radar sensing applications as well as in communications over highly frequency selective channels, has also been proposed as a way to communicate at terahertz frequencies even when partially covering or even crossing an absorption line [94]. All these techniques have been experimentally tested and demonstrated using existing terahertz communication testbeds [61].

When it comes to error control coding, research is much more scarce. In the context of pulse-based terahertz communications in nano-scale, low-weight coding schemes [95] have been proposed as a way to neither detect nor correct but prevent errors. By reducing the coding weight, i.e., the number of logical ones over the number of logical zeros, both molecular absorption noise and interference can be reduced, which leads to a lower number of channel errors. At the macroscale, no error control technique has been proposed tailored to the peculiarities of the terahertz channel, but for now there is either no error control, or just a simple adaption of common techniques in 5G, such as LDPC codes or polar codes. Potentially, techniques such as Guessing Random Additive Noise Decoding (GRAND) [96], in which the information contained in the noise is leveraged to minimize the number of errors, can be leveraged to beneficially exploit the impact of molecular absorption.

Conclusion While communication solutions developed for lower frequency systems, including the state of the art in 5G, can be *ported* to the terahertz band, these can at most fight against the properties of the terahertz channel. Instead, the physical layer of terahertz networks should leverage such properties to enable capabilities that are not possible at other frequencies. Joint wavefront and waveform engineering in the near field is one such approach.

2.5 Ultradirectional Networking Protocols

For the time being, the majority of work at terahertz frequencies has been focused on the hardware, the channel, and the physical layer. Now that those are on track, it is time to define the higher layers in the protocol stack of terahertz systems.

2.5.1 Neighbor Discovery and Medium Access Control

The need for high-gain directional antenna systems at the transmitter and the receiver introduces many challenges for initial link establishment. At lower frequencies, including in millimeter-wave systems, this problem is partially solved by allowing one of the users to operate in quasi-omnidirectional mode during the neighbor discovery phase. At terahertz frequencies, this is not an option either because antennas do not support that or, if they do, additional changes are needed to be able to close the link.

If the antennas have a fixed gain or beamwidth, the only way to locate other users is by sweeping the space by means of mechanical or electronic beam-steering. To ensure that a group of space-sweeping users can find each other and can do so in the shortest time, there are different techniques that can be followed. For example, in [97], it is shown how rather than waiting for a maximum in the received power from an intended user to determine its location, one can monitor the variations in the measured electric field and map those to the different regions of the antenna radiation diagram, automatically determining the relative orientation between the users. Machine-learning techniques can be leveraged also to identify the turning patterns of users in neighbor discovery mode.

If the antennas can dynamically adjust their beamwidth, other options become available. While simply widening the beam will not suffice, simultaneously reducing the bandwidth can be used to reduce the noise at the receiver, which ultimately plays a role as critical as the transmission power, despite many times not being considered as a system parameter. Of course, all these techniques will need to capture the fact that users might be in the near-field or in the far-field, and in fact transition between the two regimes, with all the implications that this might have.

Even when knowing the location of the different users, medium access control (MAC) can become challenging. The problem is not competing for the channel resources, but synchronizing with the receiver. In traditional MAC protocols, the transmitter pushes data to the receiver either directly or after securing that no other transmitter is going to access the channel. However, even if that is done, there is no guarantee that the receiver is facing or aligned with the transmitter. In [98], it was shown that a modified polling or receiver-initiated MAC protocol outperforms the more conventional MAC protocols found in wireless standards, such as Carrier-Sense-Multiple-Access (CSMA), even when tailored to the physical layer of terahertz networks.

Other alternatives that have been explored include leveraging the more than likely presence of multiple radio access technologies available to the user to coordinate the terahertz data exchange. For example, in [99], a microwave link is utilized as the control plane of the data plane at terahertz frequencies. A similar solution was later proposed in [100]. The main challenge with these solutions is the fact that, while multiple radios exist in the same device, current chip architectures and communication standards do not facilitate their coordination. This is something that certainly can change if shown to be the best approach.

2.5.2 Multi-hop Relaying, Routing and End-to-End Transport

As mentioned in Sect. 2.3, multi-kilometer-long links at terahertz frequencies have been demonstrated. However, these are only possible in strict line-of-sight propagation. To further increase the communication distance or to ensure connectivity even when a direct path between the transmitter and the receiver does not exist, relaying becomes critical. In this direction, in [101], a mathematical framework was developed to study the optimal relaying distance that maximizes the network throughput, by taking into account the multiple cross-layer effects between the channel, the antenna, and the physical, link, and network layers. These include the unique trade-offs between bandwidth, beamwidth, latency, and throughput with distance.

Whether between users with the help of relays or in mobile adhoc networks at terahertz frequencies, routing becomes critical. Similarly to other network functionalities, one cannot simply reuse existing techniques because those do not capture the behavior of terahertz systems. For example, as discussed before, even when knowing the exact location of all users, geographical routing techniques do not suffice because, as important as the location, is the direction in which the different users are changing, and this changes in time. Certainly, if the network is orchestrated by a common protocol, there is going to be a certain pattern in the connectivity map through time-based on the different users turning requirements. Such patterns can eventually be understood and leveraged by machine-learning techniques, and then guide the routing decision at every hop, including forwarding the information to the best next hop, deflecting it to a second or third-based alternative, or reflecting it back to the previous node [102].

On top of all these solutions, there is a need for a mechanism to ensure end-to-end reliability [103]. At the very least, the timings in TCP will need to be intelligently adapted to accommodate the potentially very **disparate** latencies between aligned links that can make the most of the physical layer and links in which MAC, relaying and routing drastically reduce the throughput. In parallel to the development of real-time terahertz communication platforms needed to test innovative networking solutions [60], simulation tools can be leveraged. For example, TeraSim [104] is an ns-3 extension that incorporates the terahertz hardware and channel peculiarities and implements early physical and link layer solutions, compatible with existing models and protocols in ns-3.

Conclusion While the physical layer of terahertz communications can meet or exceed the Tbps requirement stated in the introduction of this chapter, the throughput experienced by the user can be orders of magnitude lower unless we totally rethink the higher layers in the protocol stack. This is where the next big contributions to the field are needed to make the terahertz band ready for 6G in 2028.

2.6 Policy and Standardization

For a new technology to go beyond the research lab and have an impact in society, we need to be able to legally use it and we need to agree on how to use it. In this section, we summarize the main issues relating to terahertz spectrum policy and standardization.

2.6.1 Spectrum Policy Above 100 GHz

Based solely on the physics of the channel, one would think that, indeed, the terahertz band channel can provide users with unprecedentedly large bandwidths, comparable only to those in optical systems. However, besides the laws of physics, wireless systems need to obey international spectrum policy laws set by the International Telecommunications Union (ITU), and then further tailored by the government of each country. For example, in the United States, that is the task of the Federal Communications Commission (FCC).

As of today, all the spectrum up to 275 GHz has been allocated, meaning, the types of services that can operate at different frequencies have been established. While many times the terahertz band has been described as no man's land, the reality is that the scientific community has had an interest in the terahertz band over several decades. As a result, in World Radio Conference (WRC) in 2001, the periodic venue in which the ITU agenda is set, multiple frequencies above 100 GHz were exclusively allocated to sensing applications. While one could think that this is not a problem, there is after all a huge bandwidth above 100 GHz, the reality is that current passive sensing allocations manage to break down and split the terahertz in multiple sub-bands. For example, between 100 and 200 GHz, the largest contiguous band that the communication users can access is only 12.5 GHz. This is significant but for example not more than the 14 GHz unlicensed band between 57 and 71 GHz offers. Between 200 and 300 GHz, the largest communication window is 23 GHz, larger, but still far from the tens to hundreds of GHz that the channel supports.

The frequencies that the scientific sensing community chose are not random, but matched to interesting physical phenomena, such as resonances by water vapor and other molecules. Therefore, those cannot be changed. However, the sensors for such phenomena are either found in radio telescopes or on Earth-orbiting satellites. Despite that, no emissions are allowed at those frequencies, nowhere at any time. While such a strict ruling was possibly the only conceivable way to ensure the protection of such frequency bands in 2001, today we have technologies and techniques to dynamically share the spectrum above 100 GHz. For example, in [105], we designed, built and tested a dual-band system able to dynamically switch between the 120–140 GHz and 210–240 GHz bands in real-time so as to avoid interference to sensing satellites automatically being tracked. This technology is available today and this is just one example of many coexistence and spectrum sharing alternatives above 100 GHz [106].

2.6.2 Standardization Activities

While many times we relate terahertz systems exclusively to 6G and, thus, to the 3GPP efforts, the reality is that there is already a standard for wireless point-to-point links in the 252–325 GHz frequency range, namely, the IEEE 802.15.3d, which was first approved in 2017 [107]. This standard captures many of the early solutions developed in the field and envisioned many diverse applications, from on-chip communications to back-haul links. The first revision of this standard has just been approved, supporting now additional frequency bands above in the spectrum as well as additional modulations, including some of those discussed in Sect. 2.4.

While standards are often a good starting point to understanding current practical solutions, they are usually not the best solutions in the field, as scientific and engineering principles meet patents and business preferences by the major key stakeholders in the field.

Conclusion Beyond publications, the broader research community, starting with academic groups, has the opportunity and even moral duty to influence both spectrum policy and standardization. While this is true at any frequency, this is especially the case for the fastly evolving terahertz spectrum.

2.7 Conclusion

Terahertz communication and sensing systems have experienced a major transformation in the last decade. Today, many of the bottlenecks that discouraged the exploration of this frequency band, including the lack of device technologies and doubtful channel models, have been overcome. Innovative communication algorithms and networking protocols are needed to not only accommodate the peculiarities of the terahertz hardware and channel but to exploit them in ways that we can increase the data rate, reliability, throughout, and connectivity of terahertz networks. These need to be supported by spectrum policies and embraced by the upcoming standards. *How far are we?* Close enough to see systems in the 100–300 GHz range in 6G.

3 Free Space Optical (FSO) in 6G

Zabih Ghassemlooy • Stanislav Zvanovec • Shivani Rajendra Teli • Asghar Gholami

3.1 Introduction

The fifth generation (5G) wireless networks are being rolled out and deployed globally since 2020, exhibiting the following key characteristics: high data rate (20 Gbps peak and 100 Mbps per user), high area data capacity (10 Mbps/m^2), high spectrum and energy efficiency (by three and ten times, respectively), low latency (1 ms end-to-end), high mobility (up to 500 km/h), and high density (1 million devices per km^2). To achieve the 5G requirements, several key technologies such as the millimeter wave (mmWave), massive multiple-input multiple-output (MIMO), and ultra-dense network (UDN) have been proposed [108]. Since there are currently over 11 billion mobile connections worldwide, which exceeds the current world population of 8 billion, and is growing at around 30–40% per year, 5G will not be able to satisfy all future needs beyond the year 2030. Therefore, the focus is on the next-generation wireless networks, namely sixth generation (6G), which is expected to provide extremely high capacity, almost 100% coverage area by integrating terrestrial networks with underwater and space links; high location accuracy and update rate; and enhanced time and phase synchronization accuracy compared to the 5G network to meet the needs of end users as well as the industry [109]. In urban areas, 6G networks are expected to use small area cells to provide connections between users and devices using appliances, devices, displays, and lamp posts. These smaller cells will overlap seamlessly with classical cloud services, requiring dedicated machine learning and artificial intelligence as well as dynamic resource allocation on demand to offer new services or expand the capacity of existing networks. Moreover, centralized or cloud radio access network (C-RAN) fronthaul architectures would be considered in order to address interoperability and seamless connectivity between different solutions and technologies [110].

There are, however, several issues including (i) the scarcity of licensed spectrum resources, which cannot easily be resolved by existing radio frequency (RF) wireless systems, which are almost saturated by licensed spectrum and do not provide enough room to grow. As a result, wireless connectivity is less stable and less effective, resulting in slower data transfer rates; (ii) the 5G networks offer limited coverage in remote rural areas and on some roads, where alternative communication networks are required to complement radio frequency-based wireless networks for cost-effective, seamless, and ubiquitous service provision. In such cases, balloon and unmanned aerial vehicle communication links [111] are critical; (iii) maritime communication networks providing ships with high-quality communication links; (iv) virtual reality (VR), and a mix of VR and augmented reality [112], where mmWave, terahertz (THz), fiber, and optical frequency bands could be adopted; (v) healthcare, where it can provide high data rates, secure and high performance, ultra-low latency, and high reliability thus enabling remote surgeries via robotics, automation, and artificial intelligence [111]. Although it is practical to locate the available RF spectrum and use RF-based wireless technologies for mobile access networks, it is challenging to meet the very high data rate of Tbits/s of 6G due to the relatively narrow bandwidth [109, 113]; and (vi) growing concerns about security in RF communication, where RF provides a blanket cover, under which anyone could listen, therefore, is highly important in 5G and 6G networks [114, 115]. As a way to achieve a secure communication system, a cryptographic approach has traditionally been used at higher layers of the protocol layer of the network to achieve this goal [116]. As an alternative, physical layer security (PLS) focuses on the channel environment with random properties such as noise and fading. Consequently, as a result, the security of the network cannot be compromised by an eavesdropper who uses a more advanced decryption technique. When the legitimate channel's quality is higher than the wiretapping channel's quality, then PLS allows for secure communication to occur. In PLS, the primary goal is to increase the level of secrecy in the system as much as possible. The secrecy rate is simply the difference between the capacity of legitimate links and the capacity of wiretaps.

The potential enabling technologies that could be adopted in 6G RANs include the upper band of mmWave, THz, reconfigurable intelligent surface, directional antenna, intelligent radio, optical, etc. Unlike RF carrier where spectrum usage is limited and restricted, optical bands do not require any spectrum licensing and therefore, is an attractive prospect for high bandwidth and capacity applications. The optical bands (mostly the infrared) is already a key enabler for global communications networks, where continents are connected via optical fiber links. They also support large-scale connectivity in the Internet of Things (IoT) and the Internet-of-Everything and form the backbone of modern communication networks providing very high-speed connections to urban areas and increasingly also to homes [117]. Therefore, extending the optical fiber connectivity to include the free space transmission for the last mile access and non-terrestrial networks seems a natural step. Optical wireless communication (OWC), which uses the optical carrier to transfer information, offers extremely high bandwidth, ease of deployment, unlicensed spectrum allocation, reduced power consumption (~1/2 of RF), reduced size (~1/10 of the RF antenna diameter) and improved channel security [118–120]. Figure 9 shows the electromagnetic spectrum including the three main optical bands of ultraviolet, infrared, and visible light that are utilized in OWC.

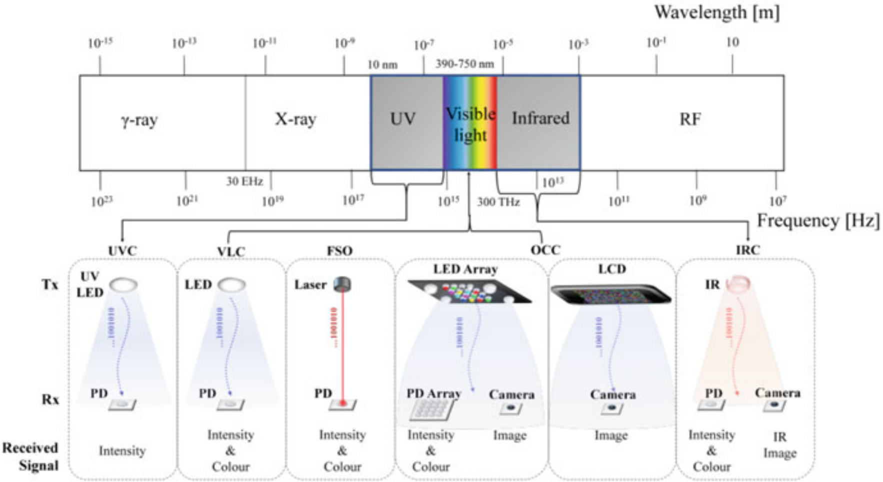


Fig. 9 Electromagnetic spectrum and OWC bands

This chapter provides highlights in prospective technologies of optical wireless communication focusing on free space optical communication (FSO) and optical camera communication (OCC) within the context of next-generation wireless networks. Section 3.2 outlines the FSO links and their applications including (i) 5G Fronthaul links in rural and urban areas; (ii) 5G backhaul and space communication; (iii) offshore and underwater communication; (iv) high-speed trains; and (v) other applications, as well as the key issues and areas for future work. Section 3.3 covers the transmission of radio signals over FSO, and the last section is all about the emerging camera-based communication technology.

3.2 *Free Space Optical Links*

The OWC can be used both indoors and outdoors with a range of coverage ranging from less than meter to several kilometers [121–123]. OWC technologies include FSO communication, visible light communication (VLC), optical camera communication, and light-based wireless networking for transmitting data at high speeds over the visible, ultraviolet, and infrared spectrums, which is best known as LiFi [124]. In terms of the end users, LiFi is like WiFi. VLC utilizes the visible band (i.e., unlicensed) to offer high transmission rates. When compared with well-established optical fiber-based infrared communication, VLC is highly appealing since it utilizes existing light emitting diode (LED)-based lighting fixtures in both indoor and outdoor environments. Note, for higher transmission speeds (e.g., greater than 10 Gbps), where large bandwidths are needed, RF-based wireless systems are not the best option because they are too complex to implement and are too costly. In outdoor applications, FSOs with transmission windows of 850, 1300, and 1550 nm are most used, along with 10 mm fiber [125], which has gained much attention as a promising solution for the problem of spectrum scarcity that wireless systems are currently facing [126]. It is possible to deploy FSOs quickly, with higher bandwidth capacities, built-in security features (a narrow beam of light, which is very difficult to detect and therefore difficult to intercept) [127], with lower power consumption than traditional microwave links, low installation and maintenance costs, and a reduced level of inter-channel interference due to the narrow laser beams that offer high spatial selectivity [128, 129].

In the past decade, research and development activities have increased in the field of OWC, especially about FSO and VLC wireless technologies that focus on high data rates and short to long transmission distances, as well as the development of new commercial products. Most works reporting on FSO utilize the simple, low complexity, and cost-effective intensity modulation and direct detection (IM/DD) technique with low power usage and relatively simple deployment compared with the coherent systems but have limited reception sensitivity and a limited bit rate of a few gigabits per second. In addition, IM/DD leads to the underutilized spectrum in OWC [130]. To increase the data rates, there are several options including (i) multilevel modulation schemes [131]; (ii) multi-carrier modulations such as orthogonal frequency division multiplexing (OFDM) [132]. It is important to note, however, that in (i) and (ii) the photodetector's higher sensitivity and linearity are particularly important for higher modulation orders; (iii) hybrid wavelength division multiplexing and OFDM [133]; (iv) higher-order modulation schemes and channel coding, known in the literature as coded modulation [134]; and (v) coherent detection, which supports two-dimensional modulation, increases receiver (Rx) sensitivity and improves system robustness against channel-induced impairments at the expense of increased system complexity, which may hinder the practical implementation of FSO communication in some scenarios, such as aerial or underwater applications [118, 123]. A further concern with coherent detection relates to its computational complexity, including local oscillator locking, in-phase/quadrature compensation, frequency offset estimation, and carrier phase estimation [135]. Thus, it is necessary to combine a simplified structure with enhanced Rx sensitivity in order to achieve a balanced trade-off between IM/DD and coherent detection. Note that FSO links in both front-haul and back-haul can carry Gbps capacity over several kilometers by employing advanced adaptive optics, fast steering mirrors, and digital signal processors.

In indoor environments, the randomness of propagation can be exploited in order to enhance the quality of service and therefore reduce the requirement for line-of-sight paths in OWC links. A potential option is intelligent reflecting surfaces (IRS), which consist of a variety of components (or surfaces) for reflecting waves and controlling their phase via bias circuits and trajectory controls [136, 137]. A smart controller can be used to switch on and off photodetectors or varactor diodes connected between the reflecting elements, thereby modifying the patch's wave response. Note, IRS is also known as reconfigurable intelligent surfaces (RIS), intelligent wall, software-controlled meta-surfaces, large intelligent surfaces, smart

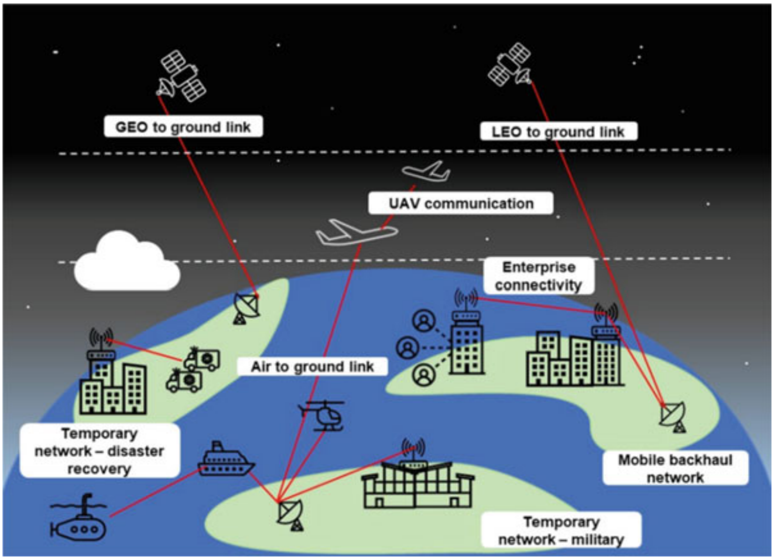


Fig. 10 FSO application areas

reflector arrays, software-defined surface, etc., and differ from massive-MIMO, relay-based links, beamforming, and backscatter communications. For point-to-multipoint optical IRS-based FSO links using micro-mirror array and phased array as well as under the combined effect of atmospheric turbulence and pointing error have been reported [138, 139].

3.2.1 FSO Applications

The FSO technology has been used in a wide range of applications, ranging from underwater to deep-space communications, see Fig. 10.

The followings are a few of the key areas in which this technology can be applied.

5G Fronthaul Links in Rural and Urban Areas

There is a major problem with bandwidth bottlenecks within the access network, which has prevented the end users from having sufficient bandwidth to meet their needs [140]. There have been several technologies developed for this purpose (wired and wireless), including Ethernet passive optical networks (EPON), where the bandwidth bottleneck is moved towards the fronthaul (last mile) access network [141, 142]. There is, however, a problem with wireless-based networks in such scenarios due to their limited bandwidth, which is more acute in rural areas [143]. FSO provides reliable network services for fronthaul and enterprise connections in rural and urban areas by replacing expensive wired communication infrastructure (e.g., optical fiber). As part of a hybrid optical wireless solution for 5G and 6G, C-RAN fronthaul mmWave with multiple bands (20–300 GHz), FSO, and radio over fiber/FSO, along with MIMO, are all potential options [144–146]. In addition, parallel FSO links-based C-RAN can be interoperable with passive optical networks and thereby provide a unique solution for the implementation of a self-contained fiber infrastructure, which can be reconfigured using software-defined networking and network function virtualization, thereby providing adaptive beamforming techniques and massive MIMO.

5G Backhaul

Since 5G requires many cell sites (mostly small cells) with a distance between them and the base station in the hundreds of meters rather than kilometers to permit high rates in cellular networks, it will lead to high-cost backhauling as a function of cell densification. The main possibilities for backhauling are wired (e.g., fiber optics) and wireless (RF and optical) links. Compared with both wired solutions, which are costly to deploy, and RF (microwave and mmWave), which lack sufficient bandwidth, FSO is seen to be a viable option for the backhaul that can meet both the capacity and cost requirements [113]. In addition, by combining FSO with unmanned aerial vehicles and balloons, non-terrestrial high-speed wireless backhaul connections as well as fixed meshed networks could be covered, while adhering to the standards of 6G networks [114]. Note that operators will have to throttle speeds if they do not have high-speed links in their fronthaul and backhaul. Since FSO reduces the distance between site cells and the C-RAN because of the density of site cells, it is a promising solution for mobile backhaul [147].

Space Communication

FSO communication has become a hot topic in space communication in the last decade, which includes ground-to-satellite communications, satellite-to-ground communications, satellite-to-satellite communications, and satellite-to-airborne communications (unmanned aerial vehicles or balloons), where FSO replaces existing radio frequency communication links. Since there are multiple transmission windows available for selection in the atmosphere, it is crucial to carefully consider the wavelength selected for transmission. One of the four main atmospheric transmission windows with a low amount of absorption by the atmosphere is the short-wave infrared wavelength band between 0.8 and 1.7 m, which is very popular due to the relative maturity and commercial availability of devices. However, FSO links operating in this band face experience significant attenuation under light to moderate fog, thus potentially reducing link availability [122, 123].

In addition, with the emergence of advanced high-speed Internet of the sky by NASA [148] as well as the planned massive deployment of low Earth orbit satellites in the next 10 years [149], a global-level laser-based satellite link with relays is becoming increasingly feasible for data communication, remote sensing, and Earth observation. Additionally, (i) with the use of coherent FSO systems, the link capacities can be increased by 10 to 100 times when compared with the capacity of systems based on RF signals [150]; (ii) by using optical transceivers, satellite communication systems can be designed to reduce their size, weight, and power specifications. For example, a lighter and smaller spacecraft will lead to a lower launch cost or perhaps more room for scientific instruments, while a lighter spacecraft will consume less energy; and (iii) the use of integrated photonics can boost space communications by reducing the payload size. Note that a team of international researchers demonstrated a record data rate of up to 1 Tbits/s over a transmission distance of 53 km using a single wavelength FSO system [151].

Off-Shore and Underwater Communication

Increasing levels of activity in offshore and underwater environments require high-speed, reliable, and robust communication links. This includes ship-to-ship, ship-to-shore, ship-to-satellite, submarine-to-space communication, terrestrial-air-underwater, submarines, autonomous underwater vehicles, underwater wireless sensor networks, and diver communications [152–154]. The RF-based links are more suitable for (i) long-rage ship-to-ship and ship-to-shore communications; and (ii) very short-range underwater communications. In salty conductive water with very high absorption using RF technology with large-size antennas, and high transmit power will only provide low data rates [155]. While underwater links based on widely used acoustic waves at frequencies of tens of kHz can provide long-distance communication link spans up to several tens of kilometres but at a higher latency and lower data rates [156]. Consequently, underwater FSO is a potential solution, which offers higher data throughput, lower latency, high security, and low power usage [157].

In underwater environments, scattering and absorption are both important factors that reduce the light intensity and time dispersion, resulting in a reduced transmission span of tens of meters and a limited bandwidth of a few GHz. A wavelength-dependent attenuation effect is caused by absorption, and seawater exhibits a relatively low absorption coefficient in the blue and green wavelength bands at frequencies between 450 and 550 nm, which is used for the development of underwater FSOs. The scattering of light by small particles suspended in water deflects the light beam from its original path and is a wavelength-dependent phenomenon. In the case of multiple scattering, time dispersion is induced, resulting in a limited bandwidth. The presence of turbulence is another factor that adversely affects underwater FSOs, further limiting the performance of the link. A typical underwater FSO channel is modeled mainly using Monte Carlo simulations, which take scattering and absorption into account as well as turbulence effects [158, 159]. Further, underwater FSO systems are highly suitable for large-scale underwater wireless applications as sensor networks due to their low cost and relatively simple implementation characteristics. The future 6G network is expected to offer full integration among terrestrial, air, and underwater communication entities, thus the need for an integrated end-to-end system, which may use fiber-wireless-based unmanned aerial vehicles assisted hybrid FSO and underwater optical communication with a decode-and-forward relaying protocol [160].

High-Speed Trains

There is no doubt that transportation is one of the key factors contributing to climate change, with air travel having the highest specific impact on climate change, followed by railways and buses [158]. The most environmentally friendly mode of transportation is by train as it produces 80% less greenhouse gas emissions per kilometer than cars, transports more people, and reduces noise pollution [159]. Providing passengers with high-speed broadband Internet access for video on demand, voice over Internet protocol, streaming, videoconferencing, etc. is crucial to making rail travel more attractive. A broadband connection between trains and infrastructure is also important for the safety of trains and for the exchange of communication-related information between trains and infrastructure (i.e., trackside equipment, trains, signaling systems, and people), which is needed by the network managers and railway operators. In the railway environment, there are several issues related to RF-based wireless technologies, including (i) Faraday cage, which results in a high attenuation of signals; (ii) vibration; (iii) ground infrastructures; (iv) pylons, tunnels, and overhead cables that can result in loss of synchronization, disconnections, and packet losses; (v) coupling/decoupling train carriages; and (vi) bandwidth bottleneck.

By using FSO for ground-to-train communications, the acquisition, tracking, and pointing mechanism provides a large coverage area for the track-side base station as well as uninterrupted internet service with minimal handover time [161]. A high-speed image sensor-based ATP mechanism [162], rotating transceivers with a wide beam located on the ground and the train [163] have been used to increase the coverage area, reliability, and complexity of the FSO system. For seamless handovers and high bandwidth communications, two FSO coverage models were proposed, namely single wavelength and dual-wavelength [164], whereas a sectorised base station model, see Fig. 11, was investigated for ground-to-train communication [165], which provides a stable high-speed link between the train and the ground base station under harsh weather conditions and mitigates the effect of geometric loss.

Other Applications

The FSO technology can also be used in several other applications. This includes (i) *wireless power transfer* – Providing suitable power for the charging of batteries in remote devices, such as IOT devices and wireless sensor networks as part of the

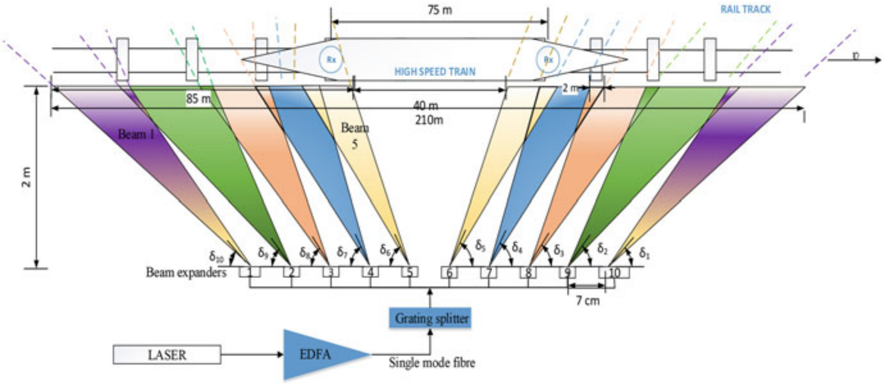


Fig. 11 The laser/fiber array covers 170 m (85 m on either side). The distance between the two BSs is 210 m [165]

future 6G network; (ii) *robotic systems* - It is envisaged that FSO-supported 6G will be used for communication between robots, and for communication between robots and servers; (iii) *inter- and Intra-chip communication* – Replace all data wiring links within and between chips with 3D domains that provide very high-speed, low-latency communication; (iv) *disaster and emergency relief network* – FSO systems can be quickly deployed to provide fast and reliable communication links, where the existing communications networks are no longer operational; and (v) *hybrid FSO/RF communications* – To achieve 99.999% link availability under all weather conditions; overcome the limitations of both RF and FSO technologies; reduce power consumption and costs by only using FSO or RF link at any given time depending on the weather conditions; link alignment with the auto-tracking system; and reduced installation cost and complexity.

3.2.2 Key Issues and Areas for Future Work

Atmospheric Channel

The FSO channel consists of terrestrial and space sections, where both optical signal attenuation and intensity fluctuation are considered. Atmospheric attenuation due to fog, haze, rain, snow, smoke, and so on affects the intensity of the transmission signal. In order to optimize FSO systems, channel modeling is essential, which assists in the design of the system in order to enhance performance. There are many traditional channel models that are based on physical theories and mathematical statistics to describe certain channel effects and atmospheric randomness [166]. E.g., attenuation due to both absorption and scattering processes can modelled using exponential and its compound function [122, 123]. Whereas, turbulence, which causes fluctuation in intensity, phase, and angle of arrival, has been extensively modelled and reported in the literature, where intensity fluctuation is the most prevalently studied in FSO [122, 123].

In traditional channel modeling, simplified assumptions have been made in order to avoid backscattering and depolarization effects due to the fact that certain complex partial differential equations cannot be solved numerically. Both wireless and optical communication systems have successfully utilized autoencoder-based end-to-end learning as a powerful and prospective solution for global optimization [167]. Recent developments in deep learning have made it possible to obtain accurate channel models in an efficient and accurate manner using data-driven methods, which have low complexity without iterative operation, are simpler to model, and are differentiable, which makes them compatible with an end-to-end system [168, 169]. It is important to note that free space channels exhibit unique properties and effects as they relate to the interaction between the optical carrier signal and the atmosphere. Therefore, deep learning algorithms previously applied to fiber and wireless RF channel modeling cannot be directly applied to FSO channels, thus network structures, loss functions, learning rates, and so on, should be further investigated [170].

The uncertainty of the channel condition in terms of turbulence, fog, etc., has been a major issue in the development and deployment of FSO links. Generally, the performance of FSO is mostly attenuated by the fog condition. The presence of aerosols and particles triggers the absorption and scattering of the propagating optical signals. In addition, several mitigation techniques have been proposed in the literature to reduce the effect of turbulence including (i) the use of aperture averaging; (ii) using the mid-wave (3–5 μm)) and long-wave infrared (8–14 μm) bands, where attenuation experienced is significantly less but devices are costly and not readily available; (iii) like maximum likelihood sequence estimation [171]; (iv) error-correcting codes [172]; (v) cooperative relaying [173]; (vi) transmitter (Tx) power control [174]; (vii) adaptive modulation [175]; (viii), spatial and time diversities such as adaptive optics, equal gain, selection combining, and the maximum ratio [176, 177] have been reported in the literature; and (ix) multiple beam transmission/TX diversity, relay-based transmission, and hybrid RF/FSO have been proposed [178]. Note, in case of thick fog, high-power lasers together with special mitigation techniques can potentially offer a better link availability. At higher data rates i.e., beyond a few Gbps, FSO with coherent transceivers offers the best option compared with IM/DD, thus allowing for increased sensitivity and spectral efficiency. Even though FSO links with coherent detection avoid the data rate limitation due to the chromatic dispersion in IM-DD systems turbulence-induced amplitude, and phase distortions (i.e., beam wondering and dispersion), as well as channel losses, cannot be ignored, which leads to inter-symbol interference [179]. In addition, substantial improvements are still needed in terms of power consumption, complexity, and cost in order for FSO with coherent detection to be a successful commercial product.

By using adaptive optics, spatial diversity, enhanced modulation schemes, and coding, some of the issues may be addressed. However, these techniques usually involve a high level of system complexity, as well as a fast response time [180–183]. Furthermore, link-level optimization is also capable of significantly reducing the overall system cost. It is possible to meet both power and cost requirements in short-range links by using low transmit optical power without avoiding optical amplification. In IRS-based FSO hybrid RF-FSO systems with multiple reflecting elements, channel modeling considering both turbulence and pointing error needs further detailed studies as well as experimental verification, which needs to be compared with traditional IM/DD-based links.

Channel Coding and Modulation

In order to improve the bit error rate (BER) performance of the link, forward error correction (FEC) is an effective technique to correct the errors caused by attenuation and turbulence due to intensity fluctuations and fading of the received signal. There are two types of FEC, linear and non-linear block codes, and they are treated in the same manner in both wired and wireless links according to their characteristics. In terms of block codes, Reed-Solomon codes and low-density parity checks (LDPCs) are the most widely used [184, 185]. There are several types of non-linear block codes, such as convolution codes, turbo codes, and Trellis coded modulation, all of which provide improved and worse BER performance under mild and strong channel conditions, respectively [186]. There has been a great deal of research done on concatenated codes in traditional communication systems, but they are rarely used in FSO systems [187]. Concatenated codes have the potential to become a promising method of improving the reliability of coded modulation FSO, therefore, the need for further research on this for use in FSO systems.

Modulation includes (i) intensity and phase-based schemes such as based-band pulse modulations (pulse time modulation schemes such as pulse position modulation, pulse interval modulation, etc.) [122]; on-off keying [188]; high-level modulation (M-ary phase shift keying and quadrature amplitude modulation schemes, etc.) [189]. Hence, spectral efficiency depends on modulation level M. High-level modulations can effectively ensure high efficiency, however, compared with multi-level pulse time modulation schemes they lack power efficiency [122]; and (ii) orbital angular momentum modulation and its higher-order derivatives, which is used for long-range and deep space FSO communication [190–192]. It is possible to compensate for distortion with FEC as well as use adaptive optics in order to ensure tracking and pointing. In addition, it can also be used in conjunction with other modulation schemes to increase the throughput of the data. To address the power, complexity, and performance requirements of FSO systems, coding/modulation schemes with soft decisions, hard decisions, or both as well as advanced digital signal processing chipsets are an important area of further research. It is important to note that power consumption is of utmost importance in certain applications, such as space, where very low-rate FEC codes are adopted in FSO links to ensure link availability regardless of weather conditions.

Probabilistic Constellation Shaping (PCS)

Under additive white Gaussian noise channel conditions, PCS is intended to reduce the modulation gap to the theoretical channel capacity limit defined by Shannon [193]. The additional gain of 1.53 dB can be critical in optical amplifier-free FSO links in order to meet power and cost objectives. Note, PCS has already been adopted in access networks [194], underwater VLC [195], and FSO links [196]. Changing the distribution of constellation points can improve the performance of the FSO link using coherent detection by increasing the signal-to-noise ratio [197]. Further, geometric shaping has been adopted to improve the spectral efficiency of the system by adjusting the positions of constellation points [198, 199]. Developing new shaping strategies together with efficient algorithms for the FSO channel are areas of future research. There is a trade-off between the transmission performance and the performance of the system, i.e., shaping gain, channel conditions, etc., thus the need for optimization of PCS schemes for FSO application.

Hybrid Radio-FSO Transmission

To enhance the reliability of the FSO link in the backhaul and space communications, a hybrid approach based on an RF link such as mmWave [131] with a comparable data rate is often proposed and investigated along with the FSO link, most preferably the next-generation coherent systems [200] to provide link availability under all weather conditions. In contrast to the FSO link, the RF link is less susceptible to fog, smoke, pointing errors, and turbulence-induced fading. However, the RF link is more sensitive to rain and small-scale fading [201].

Furthermore, the use of buffered and non-buffered relay-based parallel hybrid RF-FSO systems can provide increased coverage over greater distances [202]. In hybrid RF-FSO systems, the channel state information (CSI) of the FSO link needs to be monitored frequently and a feedback signal must be sent back to turn on the RF link to always ensure link availability. Of course, this can lead to frequent channel switching (i.e., the ping-pong effect) and therefore greatly reduce the transmission efficiency [203]. In airborne FSO links, this is a problem since the CSIs of both RF and FSO channels can change frequently and rapidly. Thus, the use of parallel data transmission over both links a maximal ratio combing scheme at the Rx to maximize the channel spectrum utilization and combat the fading [204]. Note, a Turbo-RF/FSO system can achieve a maximum coding gain of 17 dB using RF-FSO hybrid channels [205]. In contrast to LDPC, Turbo codes usually have the disadvantage of a higher level of decoding complexity, thus the trade-off between performance and complexity.

Pointing Error

Building vibrations, sways, and thermal expansions, among other factors, will result in optical beam displacement along with horizontal and vertical axes, resulting in intolerable signal fading and reducing link performance significantly. As a result, a high precision pointing system is essential when it comes to long-distance FSO links in order to improve acquisition probabilities and shorten acquisition times. Note that, the ratio of channel capacity to the bandwidth is maximum when there is no pointing error and decreases with the pointing error. Several techniques have been proposed to address the pointing error problem including (i) using a telescope to increase the Rx's field of view; (ii) diversity, i.e., multiple optical TXs and Rxs; (iii) a robust control and optimization strategy, e.g., optical transmission beams with large divergence angles help to compensate for reducing the pointing error impact on the link performance at the cost of increased beam divergence/geometric loss [122]; and (iv) advanced algorithms particularly for moving platforms, which is the most cost-effective solution [206, 207].

Security

To achieve a secure communication system, cryptographic is commonly used at higher layers of the network protocol, whereas PLS [208] focuses on the channel environment with random properties such as noise and fading increasing the system's secrecy rate so that even multiple eavesdroppers with more advanced decryption techniques cannot compromise the network's security [209].

3.3 Radio over FSO

The 5G mobile network considers the use of mmWave frequency bands with frequencies up to 86 GHz [114]. In 5G, new radio access network (RAN) concepts, such as C-RAN, in which network resources are pooled in a centralized baseband unit (BBU) pool, are considered [210]. The remote radio heads are then usually connected with the central BBU pool by an optical fiber representing the fronthaul network. It is the capacity of the fronthaul network interface that is the bottleneck to providing sufficient high data rate connections between the radio sites and the core networks. This issue can be solved using an analog transmission over an optical fronthaul infrastructure, known as radio over fiber (RoF) [211]. This technique enables the transmission of the radio signal through an optical medium in the original form, and thus, there is no need for any processing in the remote antenna site except filtering and amplification prior to the signal being radiated by the antenna, see Fig. 12. For example, the mobile fronthaul based on the analog RoF can then be scaled for up to 96 GHz transmission [212] covering all considered 5G frequency bands. The utilization of FSO technology introduces a higher level of flexibility in terms of avoiding burying fiber in dense urban areas or overcoming both natural and artificial obstacles [213].

Two experimental campaigns have been carried out by using two MZMs to double the incoming frequency to desired 25 GHz band and modulate data with

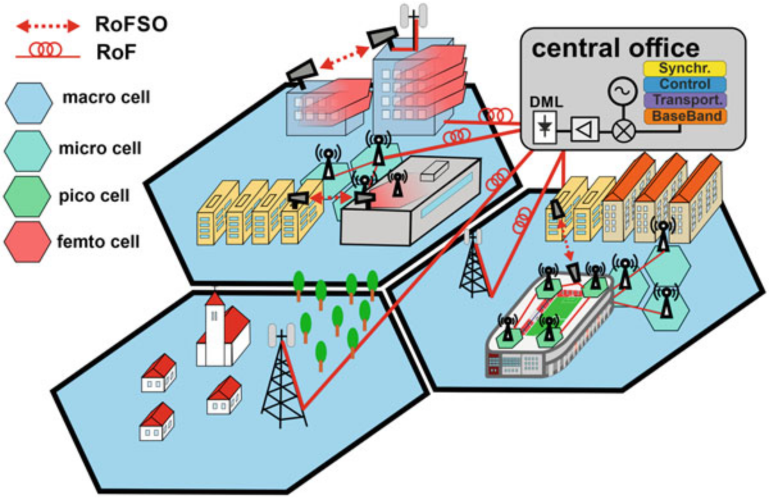


Fig. 12 Example of RoFSO implementation within 5G

a fiber, FSO and mmWave antenna seamless transmission in [214] and [215]. Both works used long term evolution (LTE) wireless standards to evaluate the transmitted signal quality with the maximal bandwidth of 20 MHz and 64-quadrature amplitude modulation (QAM). The main focus was put on the impact of atmospheric turbulence, which affects the signal in FSO channel with a maximal experimental indoor FSO range of 40 m. The measured system performance in [216, 217] was consequently used for simulated extended wireless ranges to provide maximal coverage. Although successful transmission of the mmWave signal over up to 50 km of optical fiber, 40 m of FSO, and 40 m of seamless antenna channel was demonstrated, the used LTE transmission frames with a 20 MHz bandwidth seemed to be insufficient for the high-capacity RAN. The beating of the two optical tones to generate the mmWave signal at 95 GHz was also used in [218] for an RF wireless transmission. The authors exploited a dual-drive MZM to generate the two-tone optical signal while a phase MZM was used for data modulation. The proposed system was afterward tested over 20 km of single mode fiber (SMF) and then the mmWave up-conversion was realized at a self-fabricated high-frequency optical-to-electrical converter. A successful 5 m long wireless transmission of orthogonal frequency division multiplexing (OFDM) 32-QAM signal with 3 GHz bandwidth at 95 GHz was demonstrated with the error vector magnitude (EVM) as low as 12.1%. A high throughput, i.e., 1.4 Gb/s, has been demonstrated for the analog RoF 5G fronthaul transmission with a 9 m long antenna seamless link at 25.5 GHz by using multicore fibers as the optical medium in [219, 220].

Also, a low-frequency directly modulated laser (DML) can be used for the system using photonic frequency doubling by beating of the optical sidebands at the direct photodetection, as outlined in [221]. The usage of DML represents a less complex and cost-effective solution by substituting the continuous wave (CW) laser and one MZM for data modulation. There was described the photonic signal generation of 40 GHz mmWave and transmission over 10 km of SMF and 1.5 m long FSO channels by using the combination of DML and MZM. Up to 1 GHz signal bandwidth and up to 64-QAM modulation format with the maximal bit rate of 2 Gb/s were received in the system and evaluated in terms of EVM. The combination of DML and MZM was used also for the experimental comparison of remote and local photonic mmWave signal generation [222]. In order to reveal the impact of either local or remote photonic generation at 40 GHz, a quadrature phase shift keying (QPSK) signal was employed by the authors with a maximal bandwidth of 250 MHz. The local and remote generation in fact relied on the location of the MZM, which was biased at the null point and was responsible for the optical carrier suppression. In the experiment, the MZM was placed behind or in the front of the optical distribution network, consisting of up to 25 km of SMF. It has been revealed that remote generation leads to a higher frequency response than a local generation. The low-frequency DML was used as well as a low-cost solution for intermediate frequency (IF) over fiber (IFoF) transmission in [223], where authors experimentally demonstrated end-to-end fiber wireless fronthaul. The proposed approach introduced a broadband IFoF transmission combined with narrow band IFoF by using the low-frequency DML and classic analog RoF approach utilizing CW laser and MZM for 28 GHz RoF transmission with a seamless 10 m long antenna channel. The total system capacity was 34.2 Gb/s over 20-km broadband IFoF, 1 km narrowband IFoF, 500 m RoF and 10 m wireless mmWave links. For the system performance evaluation, the authors used a 5G new radio (NR) signal with QPSK and 64-QAM modulations while all recorded EVM magnitudes after the wireless link satisfied the 8% criterion. The most comprehensive comparison between the classical analog RoF and the photonic frequency doubling system with CS mmWave transmission for mobile optical fronthaul using 5G NR signals has been experimentally demonstrated in [224].

3.3.1 Hybrid Wireless Systems Combining OWC and RF

Hybrid wireless systems can integrate two or more technologies (e.g., FSO/RF, VLC/RF, VLC/FSO, VLC/WIFI, and many other combinations) into a hybrid network and exploit their advantages. Hybrid systems can play a key role in link reliability, wireless connectivity, and interference reduction [124]. The performance of OWC technologies can be affected by many factors including turbulence, dense fog, or a pointing error in the case of FSO. These factors can significantly influence the reliability and performance of an FSO system. One possible solution to improve link performance is a combination of RF and FSO technologies. By combining these technologies, seamless connections, boasting a large range, reliability, and bandwidth can be established in current networks [225]. Analyses of the AF dual-hop RF/FSO system combining a Rayleigh distributed RF link and the FSO part simulated as a Gamma-Gamma turbulence channel were introduced for the first time in [226]. Extending the previous work considered the pointing error in the FSO link [227]. The dual-hop system outage performance, where RF was modeled as Rayleigh fading and FSO as M-distributed fading, was investigated in [228]. The effect of turbulence and the pointing error on the channel capacity of the RF/FSO system with a Nakagami-m distributed RF link was studied in [229]. However, similar work for a Decode and Forward (DF) scheme was performed in [230]. Heterodyne detection with variable gain and a fixed relay scheme was considered in [231] and achieved a higher capacity of about 1.5 b/s/Hz compared to IM/DD. For data rate improvement, the combination of the dual-hop MIMO-RF/FSO system with a Gamma-Gamma distributed link was presented in [124]. The multi-user hybrid DF-based system was analyzed in [232] with the derivation of the symbol error probability of each user. The idea of extending the previous work about an optimal power allocation scheme for a multi-user system was proposed in [233] to optimize overall system performance.

In the case of VLC/RF hybrid technologies, they are used to mitigate line of sight path LOS blockages, inter-cell interference, and handover issues, which results in finding a solution to distribute users among the RF and VLC access points to improve system performance with acceptable fairness of the system [234]. Access point assignment was first studied in [235] and found an optimal load balance between one RF access point and one VLC access point. A method where all traffic is at first assigned to a VLC network, followed by users receiving a lower data rate than the defined RF thresholds are re-allocated to the RF access point [236]. Dynamical user distribution, based on channel conditions for multiple VLC and RF access points, was proposed in [237]. This implementation improved system performance by about 40%, compared to a single VLC or RF network. The implementation of centralized and distributed algorithms for resource allocation among both types of access points was analyzed in [238].

To limit the number of handovers and their hard implementation in VLC, in [239], a dynamic load balance algorithm, which assigns quasi-static users to VLC access points and moves users to RF access points was proposed. Based on previous work, two types of load-balancing algorithms were proposed in [240]. A joint optimization algorithm achieves more than a 1.5 times higher data rate than a separate optimization algorithm. However, with significantly higher computation complexity, it reaches even more than 1000 times the separate optimization algorithm. A different approach is proposed in [241] based on users' statistical information on channel blockage: the users which are influenced by channel blockages should switch to RF access points. To decrease optimization complexity, a load balance fuzzy logic-based system was proposed in [242]. A user scores an access point based on several conditions (throughput, SNR, and interference) and then, based on the score, decides whether to connect to RF or VLC access points. Recently, the hybrid FSO/VLC communication system has gained wide popularity due to its properties, such as high data rate, security, and relatively low interference.

A cascaded FSO-VLC system consisting of multiple VLC access points using a DF relaying scheme was proposed in [243]. The FSO link is characterized by path-loss, pointing error, and atmospheric conditions while the SNR for both links is statistically characterized, considering the randomness of end-user's positions for indoor and outdoor environments. The achieved results provide a compelling solution for current broadcasting systems. Following the extension of the previous work on parallel RF/FSO, an outdoor link was proposed in [244]. The effect of the outdoor parallel link significantly improves system performance, especially in very strong turbulence conditions where outage probability can be improved approximately 58 times. The performance of the hybrid DF relaying VLC/FSO/VLC system is derived in terms of a closed-form expression for the outage probability [245] for the VLC modeled as the Lambertian emission model and the FSO link as a Gamma-Gamma channel under the impact of turbulence, semi-angle, and FOV of a detector. The first experimental application of the hybrid FSO/VLC system was presented for space-air-ground-ocean-integrated communication in [246]. A simple network mechanism for identification, user mobility control, and network routing is designed for the interconnection of VLC access points. The system was designed to transfer data rate 450 Mb/s over a 1 m long OFDM-based VLC interconnection and a 960 Mb/s on and off keying (OOK)-based FSO over 430 m without a turbulence condition. A hybrid FSO/VLC system with m-CAP, suitable for the last mile and last meter access networks interconnected by an SMF offering both security at the physical layer and lower installation costs, was presented in [247]. We showed that comparing 2-CAP and 10-CAP achieved an approximately 43% improvement in the data rate under the clear atmosphere condition. Further, for a 500 m-long 10-CAP hybrid FSO/VLC link under turbulence, the drops in Rb were 8.9% and 30% for C $_{n}{}^{2}$ of 3.9×10^{-17} m$^{-2/3}$ and 2.4×10^{-15} m$^{-2/3}$, respectively, compared to the link with no turbulence.

3.4 The VLC-OCC Technology for 6G

6G networks are expected to provide superior coverage by integrating terrestrial networks with space and underwater links, given that existing wireless systems cannot provide high-speed data rates for non-terrestrial applications [248]. The OWC technology utilizing visible light as transmission media is a high-speed communication technique with an unlicensed frequency range of 400–800 THz and can be adopted as an alternative approach to solving these problems. As a complement to RF systems, VLC uses an unlicensed frequency band of 400–800 THz ((10,000 times larger than the RF bandwidth) for high-speed wireless communication, which is safe for human eyes and can also be used for indoor positioning and sensing. This technology is available for most indoor and to some extent outdoor (e.g., vehicular communication [249], and underwater communications [250]) applications. For short-range VLCs the first IEEE standard 802.15.7 was introduced in 2011, whereas a topic interest group for IEEE 802.11 was set up in 2016.

In VLC systems the most widely used light source is the white LED, which can be generated by (i) blue emitters with a phosphor layer with a typical bandwidth of a few tens of MHz; and (ii) red-green-blue (RGB) and RGB-yellow emitters, which is preferable due to the higher modulation bandwidth and wavelength division multiplexing capability [123, 128]. LEDs offer several benefits over existing lighting infrastructures, such as lower power consumption, longer life expectancy, higher energy efficiency, reduced maintenance, lower heat generation characteristics, and fast switching speed (orders of magnitude higher). Laser diodes can also be used in VLC systems to achieve much higher data rates (orders of magnitude higher than LEDs) over long transmission spans due to their large modulation bandwidth but will require precise alignments between the Tx and the Rx [128]. In VLC systems both photodetectors (PDs) and image sensor-based receivers could be used for high and very low-speed applications [123, 128]. The latter is very attractive in IoT applications with requirements of (i) low data rates, less complexity, and cost-effectiveness; (ii) resource availability; (iii) quality of service and reliability; (iv) transmission range; (v) safety and (vi) low power consumption. As the IoT paradigm opens the doors to innovations, which contribute to interactions between objects and humans, it enables the realization of smart cities, infrastructures, and services for enhancing the quality of life and improving better utilization of resources.

Over the past few years, we have seen smart devices with built-in high-resolution complementary metal-oxide-semiconductor (CMOS) cameras [251]. These CMOS cameras can capture high-resolution videos with a resolution of at least 1280 × 720 pixels and a capture rate of 30 frames per second (fps) [252, 253], which is more than adequate for low-speed applications. Due to the large scale and increasing availability of mobile phones, smartphone VLC can be attractive, as nearly all mobile users effectively carry and regularly use camera-based optical Rxs. Not only smartphones but also most new generation smart devices have built-in CMOS cameras, providing the ability to capture photos and videos as well as being used for data communications (low speed), indoor localization, and range findings [252, 254]. The smartphone or camera-based VLC has been studied within the framework of OWC [255] and is considered a candidate for IEEE 802.15.7rl standard and is referred to as optical camera communications (OCC) [256]. OCC represents an extension of VLC with the advantage of no additional hardware to establish device-to-device communications at low data rate R_b and indoor positioning [257]. Unlike conventional VLCs, which employ PDs as the Rx, in OCC a mobile phone CMOS camera is used to capture two-dimension (2D) data in the form of image sequences, thus being able to transmit more information compared to the PD-based VLCs. The OCC technology is making remarkable progress in the key application as part of the fourth industrial revolution i.e., IoT, smart vehicles, etc. [258].

Image processing on the Rx side also provides OCC with advantages to classifying shapes and estimating the distance-based depth perception from the vision of cameras. However, in OCC, R_b is limited by the frame-rate f_R of the ISs, i.e., the camera. R_b can be increased by using higher frame rate cameras, which are very costly, and the camera capture speed, which is defined as the physical parameter of the sensor (electronics) and the graphics processor speed in the hardware domain. The f_R of the camera is generally confined to either 30 fps or 60 fps, except for some slow-motion capable cameras with f_R of 120–240 fps. Some specialized high-speed cameras are available with f_R ranging from 1000 to 21,000 fps [259]. However, these

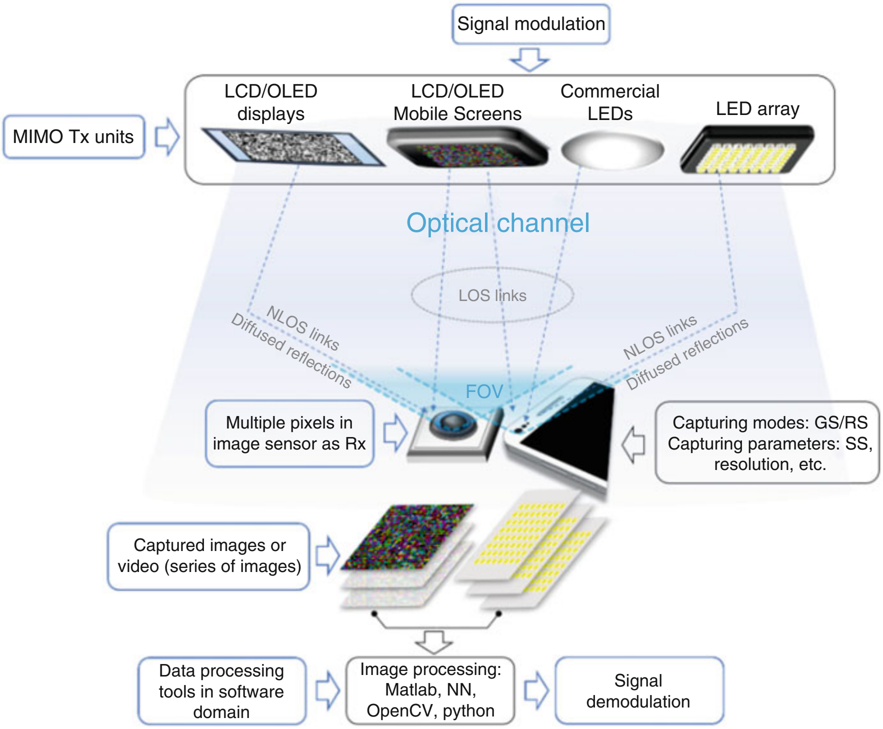


Fig. 13 Overview of OCC technology

cameras are not suitable for use in mobile devices and are less likely implemented for OCC applications. Therefore, OCC is further extended to offer massive MIMO capabilities in order to increase R_b using light emitting diodes (LEDs) and PD arrays in the form of multiple pixels in ISs for IoT applications in both indoor and outdoor environments [260].

Figure 13 shows a general overview of the OCC scheme. Recent studies in this area have outlined the use of liquid crystal displays (LCDs) with multiple embedded neo-pixels, LCD or organic light emitting diode (OLED)-based mobile phone screens [251], and LED arrays [253] together with commercial LEDs installed as the MIMO Tx units in indoor as well as outdoor infrastructures. The Tx units in MIMO can be modulated using the simple OOK [261] data format or complex modulation schemes such as color intensity modulation (CIM) MIMO [262] and under-sampled phase shift on-off keying) [263] to improve the data throughput. The modulated signals can be captured on the Rx side via the LOS [261] or non-line-of-sight- based on diffuse reflections [264]. In [251, 261, 264], a range of CMOS technology-based cameras used in the mobile phone (front/rear camera), digital single-lens reflex cameras with higher capture speeds (ranging from 50 to 1000 fps), and surveillance cameras were reported. These CMOS cameras can capture images or record videos in global shutter (GS) and rolling shutter (RS)-based capturing modes at different shutter speeds (SS) and resolutions.

As previously mentioned, the image sensor (IS) captures the intensity-modulated Tx in the form of a video recording or image frames at particular fps. Therefore, to perform an OCC reception and data evaluation, image processing tools in the software domain (i.e., MATLAB, OpenCV, and Python) [265, 266] should be employed on the Rx side. Image processing is of paramount importance in OCC for the demodulation of the received data in the form of captured image frames. Therefore, it is necessary to have robust and reliable image-processing algorithms and schemes. In recent years, neural network (NN) has attracted much attention to solving complex problems related to image recognition using an intelligent machine-learning technique. NN can be adopted for identifying objects' shapes in images, transcribing speech into text, matching classified items, and selecting relevant results of a search [267]. Along with image processing, an artificial NN equalizer, when used in VLC, can achieve high R_b by reducing the influence of intersymbol interference [268, 269]. NN in the form of trained neurons also plays an important role in motion detection (MD) over the existing indoor OCC links as was demonstrated in one of our works [270].

The camera Rx consists of an imaging lens, an image sensor along with Bayer filter, and an internal image processor which is a 2D array of nanoscopic PDs that detects information in the form of an image, see Fig. 14. A lens is a device that converges or diverges light beams. The light is projected by passing through the focal point of the lens on the sensor surface. In imaging, convex lenses are used to create a real image on the image plane of the camera, which is located at the image sensor plane. However, for the purpose of variable magnification and image correction, usually, a complex set of convex and concave lenses are used in a lens system. The magnification factor of the lens can be expressed as:

$$k = \frac{f}{f - d_L} \tag{5}$$

where d_L is the distance of the object from the lens and f is the focal length of the lens.

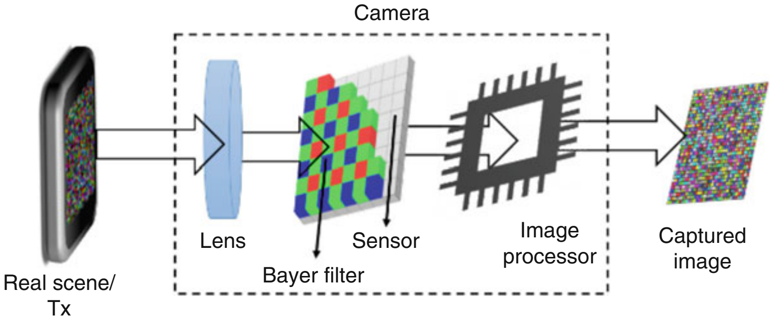


Fig. 14 Schematic of a digital camera

Bayer filter is attached over the sensor in order to make each pixel sensitive to either of the primary colors (RGB). Note, the sensor only captures monochromatic images without a Bayer filter [271]. The internal image processor performs image processing by demosaicing method to form a colored output image which is the data image for OCC post-processing [272].

The maximum achievable R_b in OCC can be obtained as [273]:

$$R_b = \frac{1}{8} WH \left(\log_x x_G\right) f_R \tag{6}$$

where $W \times H$ is the size of IS in terms of pixels, x_G is the grayscale signal obtained from each pixel and f_R is the camera capture speed in fps. Note, 1/8 is a rate reduction factor for three dimensions (3D) formats. For example, considering a 1000 fps QVGA (320 × 240 pixels) 256 grayscale IS, the maximum achievable R_b is 76.8 Mbps [273]. However, this is impractical R_b as each of the pixels in the IS should represent a unique data transmission which is not practical due to the long and varying transmission distance L between the camera and the Tx.

Moreover, IS resolution (IS_{Res}) with respect to dimensions (either horizontal or vertical) can be calculated as twice the size of FOV over the size of the smallest feature of the camera IS_{Res}. To make an accurate measurement of the captured image, a minimum of two pixels per smallest feature is considered. Considering numerical calculation, if the FOV covers 200 mm and the smallest feature needed capturing is 2 mm, the required IS_{Res} is 200 pixels. Therefore, a camera with a resolution of 640 × 480 pixels would be effective because 200 pixels is less than the smallest dimension, which is 480 pixels. For smartphone cameras, the pixel density, usually referred to as pixel per inch (PPI), is given as:

$$PPI = \frac{\sqrt{W^2 + H^2}}{Diagonal\ screen\ size} \tag{7}$$

The higher the PPI, the more details can be found within the image. The exposure time sets the amount of light that reaches the IS, which determines how light or dark an image will appear. Note, too much light captured will overexpose (too bright/no details) images, thus resulting in the blooming effect, while less light results in underexposed (dark/grainy/less details) images. The blooming effect means that the number of photons reaching the detector exceeds its maximum capacity, and the excess photons will either spill and merge to adjacent pixels or are not counted, thus leading to non-precise intensities [274].

The camera’s exposure is based mainly on three camera settings: aperture, ISO, and SS [273]. A camera’s SS is typically measured in fractions of a second. Slow SSs allow more light to be collected and are used for low-light and night photography, while fast SSs help to freeze motion [275]. Aperture also called as f_{stop}, controls the amount of light being captured through the lens as well as controls the depth of field, which is the portion of a scene that appears to be sharp. For a very small aperture, the depth of field is large, while for a large aperture, the depth of field is small. In photography, the aperture is expressed by *F-number* (focal ratio), which represents the ratio of the diameter of the lens aperture to the length of the lens [275]. Higher ISO (i.e., the sensitivity of the camera) means faster light absorbed by the sensor, but at the cost of increased noise level [275].

3.4.1 GS and RS Mechanisms

There are two types of camera sensors, i.e., a charge-coupled device (CCD) and a CMOS sensor. CMOS technology is a mature technology used in a wide range of devices such as solid-state memories, CPUs, and ISs. In CMOS technology, metal-oxide semiconductor field effect transistors (MOSFET) are used for switching. The applications of CMOS cameras range from professional photographing cameras, low-cost cameras, and surveillance to industrial high-speed machine vision. CCD is not installed within mobile devices (smartphones) due to its larger analog-to-digital converter (ADC) as compared to CMOS. The advantages of using CMOS over CCD cameras include low power consumption, faster readout, more programmability, and low cost. Therefore, CMOS is widely implemented in smartphones due to its compact ADC design. The major difference between CCD and CMOS sensors is in the capturing modes, i.e., the shutter mode. CCD sensors use GS while most CMOS sensors use RS capturing modes, as illustrated in Fig. 15a, b. As in [276], CMOS sensors can be used in both RS and GS capturing modes based on the system requirements. Accordingly, CMOS is preferred for communication applications in the literature [277–279].

A GS camera allows its sensor to be exposed to light once, i.e., it can hold either the ON or OFF state of an LED in a single frame, see Fig. 15b. The CMOS-RS camera sequentially integrates light on all pixels and then it operates similarly as a scanning function. Unlike a GS camera, in RS capturing mode, the sensor scans row-by-row of pixels (line-wise) of the entire image, which therefore introduces a sequential readout technique. This scan process is tied to the system clock and is limited by the sampling rate of the ADC. The pixel sensors within the camera continuously integrate the light that falls on their surface and then each row of pixels is exposed simultaneously at the exposure time. In RS cameras, the readout time ensures that there is no overlapping of the rows of pixels and allows multiple exposures in a single captured image. The latter enables multiple LED states to be achieved at the same time in a single frame as each row is exposed once to the light. Therefore, the captured image of the switched LED is composed of a set of black and white stripes.

On the other hand, due to a single large ADC, GS in CCD takes longer capturing and processing time for the whole frame, which is reduced in CMOS sensors due to smaller ADC. Due to compact ADC, faster processing, and advancements of electronic fabrication, CMOS sensors are preferably installed and utilized. Therefore, CMOS is employed for high-speed cameras with a capture rate beyond 1000 fps, which cannot be accomplished using the CCD sensor due to its slower ADC.

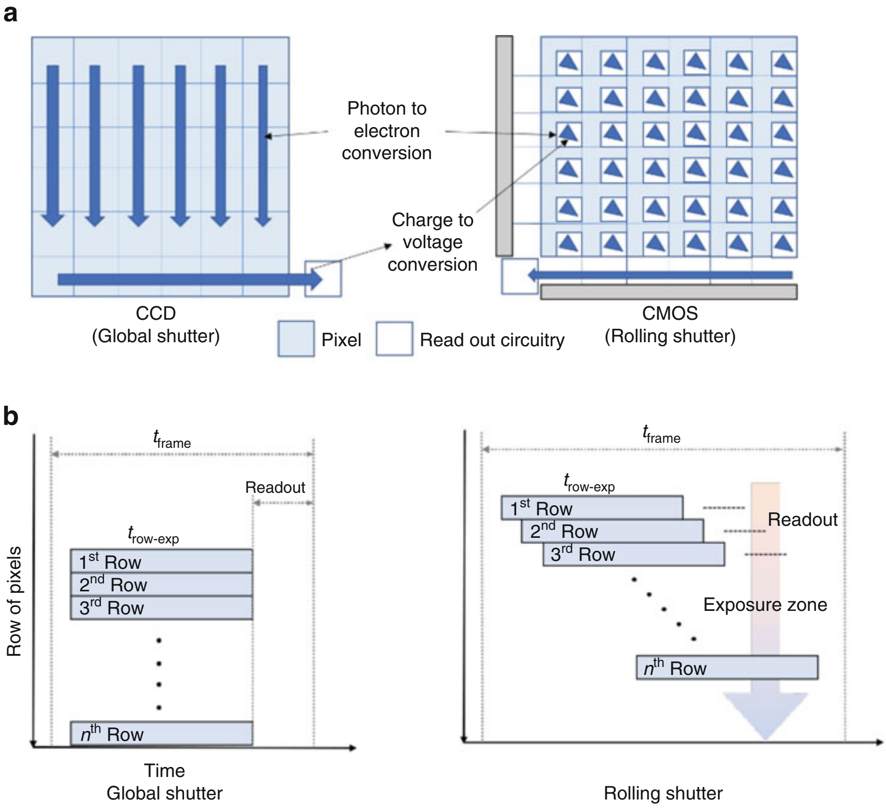


Fig. 15 (**a**) CCD and CMOS ISs, and (**b**) GS and RS mechanisms

Using the RS effect of a CMOS IS for VLC can be promising to provide flicker-free OCC links as well as to increase the R_b [280]. This enables multiple LED states (ON/OFF) to be captured at the same time. Therefore, for an LED that flickers ON and OFF according to the modulated binary bit stream, the captured image contains a bunch of black and white stripes. The stripes' widths depend on the modulation frequencies, and the number of strips depends on the distance between the camera and LED [280].

Despite the fact that the OCC technology for indoor and outdoor IoT environments offer a significant improvement in system performance and link reliability, there are still many challenges including (i) MIMO-OCC that uses large size commercially available display Tx and multi-channel modulations to enhance link span and data rate, respectively. In addition, NN-assisted motion detection based on pattern recognition algorithms can be developed in Python- or Android-based applications for smartphone to provide real-time device control in smart environments; (ii) fiber optic lighting-based OCC using commercially available plastic optical fiber to improve coupling efficiency to enhance the emission and hence the link spans; (iii) long-range outdoor links utilizing the already available lighting infrastructure such as street and traffic lights and surveillance cameras. Within this, efficient detection of regions of interest and positioning algorithms can be developed to establish communication links between moving vehicles and the infrastructure; and (iv) hybrid VLC to overcome the problem of bandwidth efficiency using multiple wavelengths, as well as utilizing equalization and different modulation formats such as OFDM to improve the reception success rate and the BER performance for low- and high-speed VLC links, respectively.

3.5 Conclusion

As potential and complementary candidates to the RF wireless systems in 5G and 6G wireless networks, this chapter discussed the emerging optical wireless communication technologies, focusing on free space optics and optical camera communication. It is expected that optical wireless technology will be used in applications where cable-based systems are too costly to deploy, and RF-based systems cannot be deployed due to spectrum shortages and congestion. In addition, this technology can be deployed in certain applications to free up the RF spectrum for use in applications where the light-based system can be used. It is anticipated that optical wireless technology utilizing the light emitting diodes and cameras of smart devices will have a significant impact on IoT applications where it will provide data communication, sensing, localization, and transfer of images.

References

1. Ngo HQ, Ashikhmin A, Yang H, Larsson EG, Marzetta TL (2017) Cell-free massive MIMO versus small cells. IEEE Trans Wirel Commun 16(3):1834–1850
2. Nayebi E, Ashikhmin A, Marzetta TL, Yang H, Rao BD (2017) Precoding and power optimization in cell-free massive MIMO systems. IEEE Trans Wirel Commun 16(7):4445–4459
3. Buzzi S, D'Andrea C (2017) Cell-free massive MIMO: user-centric approach. IEEE Wirel Commun Lett 6:706–709
4. Interdonato G, Björnson E, Ngo HQ, Frenger P, Larsson EG (2019) Ubiquitous cell-free massive MIMO communications. J Wirel Commun Netw 2019(1):197–209
5. Demirhan U, Alkhateeb A (2022) Enabling cell-free massive MIMO systems with wireless millimeter wave fronthaul. IEEE Trans Wirel Commun 21(11):9482–9496
6. Parida P, Dhillon HS, Molisch AF (2018) Downlink performance analysis of cell-free massive MIMO with finite fronthaul capacity. In: IEEE 88th vehicular technology conference (VTC-Fall), Chicago, IL, USA, vol 2018, pp 1–6
7. Masoumi H, Emadi MJ (2020) Performance analysis of cell-free massive MIMO system with limited fronthaul capacity and hardware impairments. IEEE Trans Wirel Commun 19(2):1038–1053
8. Guenach M, Gorji AA, Bourdoux A (2021) Joint power control and access point scheduling in fronthaul-constrained uplink cell-free massive MIMO systems. IEEE Trans Commun 69(4):2709–2722

9. Buzzi S, D'Andrea C, Zappone A, D'Elia C (2020) User-centric 5G cellular networks: resource allocation and comparison with the cell-free massive MIMO approach. IEEE Trans Wirel Commun 19(2):1250–1264
10. Attarifar M, Abbasfar A, Lozano A (2019) Modified conjugate beamforming for cell-free massive MIMO. IEEE Wirel Commun Lett 8(2):616–619
11. Buzzi S, D'Andrea C, D'Elia C (2018) User-centric cell-free massive MIMO with interference cancellation and local ZF downlink precoding. In: Proceedings of 2018 15th international symposium on wireless communication systems (ISWCS), pp 1–5
12. Björnson E, Sanguinetti L (2020) Making cell-free massive MIMO competitive with MMSE processing and centralized implementation. IEEE Trans Wirel Commun 19(1):77–90
13. Björnson E, Sanguinetti L (2020) Scalable cell-free massive MIMO systems. IEEE Trans Commun
14. Boldi M et al (2011) Coordinated multipoint (CoMP) systems. In: Mobile and wireless communications for IMT-advanced and beyond
15. Liu J et al (2018) Initial access, mobility, and user-centric multi-beam operation in 5G new radio. IEEE Commun Mag 56(3):35–41
16. Tugfe Demir Ö, Björnson E, Sanguinetti L (2021) Foundations of user-centric cell-free massive MIMO. Found Trends Signal Process 14
17. Björnson E, Hoydis J, Sanguinetti L (2017) Massive MIMO networks: spectral, energy, and hardware efficiency. Found Trends Signal Process 11(3–4):154–655
18. Series M (2015) IMT vision–framework and overall objectives of the future development of IMT for 2020 and beyond. Recommendation ITU 2083
19. Interdonato G, Buzzi S (2022) The promising marriage of mobile edge computing and cell-free massive MIMO. In: Proceedings of IEEE international conference on communications (ICC)
20. D'Andrea C, Interdonato G, Buzzi S (2021) User-centric handover in mmWave cell-free massive MIMO with user mobility. In: 29th European signal processing conference (EUSIPCO), pp 1–5
21. Piesiewicz R, Kleine-Ostmann T, Krumbholz N, Mittleman D, Koch M, Schoebel J, Kurner T (2007) Short-range ultra-broadband terahertz communications: concepts and perspectives. IEEE Antennas Propag Mag 49(6):24–39
22. Federici J, Moeller L (2010) Review of terahertz and subterahertz wireless communications. J Appl Phys 107(11):6
23. Song H-J, Nagatsuma T (2011) Present and future of terahertz communications. IEEE Trans Terahertz Sci Technol 1(1):256–263
24. Rappaport TS, Xing Y, Kanhere O, Ju S, Madanayake A, Mandal S, Alkhateeb A, Trichopoulos GC (2019) Wireless communications and applications above 100 ghz: opportunities and challenges for 6g and beyond. IEEE Access 7:78 729–78 757
25. Akyildiz IF, Han C, Hu Z, Nie S, Jornet JM (2022) Terahertz band communication: an old problem revisited and research directions for the next decade. IEEE Trans Commun 70(6):4250–4285
26. He D, Guan K, Fricke A, Ai B, He R, Zhong Z, Kasamatsu A, Hosako I, Kürner T (2017) Stochastic channel modeling for kiosk applications in the terahertz band. IEEE Trans Terahertz Sci Technol 7(5):502–513
27. Saad W, Bennis M, Chen M (2019) A vision of 6G wireless systems: applications, trends, technologies, and open research problems. IEEE Netw 34(3):134–142
28. Sen P, Siles JV, Thawdar N, Jornet JM (2022) Multi-kilometre and multi-gigabit-per-second sub-terahertz communications for wireless backhaul applications. Nat Electron:1–12
29. Boulogeorgos A-AA, Alexiou A, Merkle T, Schubert C, Elschner R, Katsiotis A, Stavrianos P, Kritharidis D, Chartsias P-K, Kokkoniemi J et al (2018) Terahertz technologies to deliver optical network quality of experience in wireless systems beyond 5G. IEEE Commun Mag 56(6):144–151
30. Suen JY, Fang MT, Denny SP, Lubin PM (2015) Modeling of terabit geostationary terahertz satellite links from globally dry locations. IEEE Trans Terahertz Sci Technol 5(2):299–313

31. Alqaraghuli AJ, Abdellatif H, Jornet JM (2021) Performance analysis of a dual terahertz/ka band communication system for satellite mega-constellations. In: 2021 IEEE 22nd international symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM). IEEE, pp 316–322
32. Karasik BS, Sergeev AV, Prober DE (2011) Nanobolometers for THz photon detection. IEEE Trans Terahertz Sci Technol 1(1):97–111
33. Smye S, Chamberlain J, Fitzgerald A, Berry E (2001) The interaction between terahertz radiation and biological tissue. Phys Med Biol 46(9):R101
34. Zhang Z, Zhang Y, Zhao G, Zhang C (2007) Terahertz time-domain spectroscopy for explosive imaging. Optik 118(7):325–329
35. Siegel PH (2007) THz instruments for space. IEEE Trans Antennas Propag 55(11):2957–2965
36. Jornet JM, Akyildiz IF (2014) Graphene-based plasmonic nano-transceiver for terahertz band communication. In: The 8th European conference on antennas and propagation (EuCAP 2014). IEEE, pp 492–496
37. Jornet JM, Akyildiz IF (2013) Graphene-based plasmonic nano-antenna for terahertz band communication in nanonetworks. IEEE J Select Areas Commun 31(12):685–694
38. Akyildiz IF, Jornet JM (2010) Electromagnetic wireless nanosensor networks. Nano Commun Netw 1(1):3–19
39. Akyildiz IF, Jornet JM (2010) The internet of nano-things. IEEE Wirel Commun 17(6):58–63
40. Abadal S, Alarcón E, Cabellos-Aparicio A, Lemme MC, Nemirovsky M (2013) Graphene-enabled wireless communication for massive multicore architectures. IEEE Commun Mag 51(11):137–143
41. Sengupta K, Nagatsuma T, Mittleman DM (2018) Terahertz integrated electronic and hybrid electronic–photonic systems. Nat Electron 1(12):622–635
42. Deal WR, Leong K, Zamora A, Gorospe B, Nguyen K, Mei XB (2017) A 660 GHz up-converter for THz communications. In: Compound semiconductor integrated circuit symposium (CSICS), 2017 IEEE. IEEE, pp 1–4
43. Siles JV, Cooper KB, Lee C, Lin RH, Chattopadhyay G, Mehdi I (2018) A new generation of room-temperature frequency-multiplied sources with up to 10 higher output power in the 160-GHz–1.6-THz range. IEEE Trans Terahertz Sci Technol 8(6):596–604
44. Williams BS, Kumar S, Hu Q, Reno JL (2006) High-power terahertz quantum cascade lasers. In: In 2006 conference on lasers and electro-optics and 2006 quantum electronics and laser science conference. IEEE, pp 1–2
45. Berry CW, Hashemi MR, Jarrahi M (2014) Generation of high power pulsed terahertz radiation using a plasmonic photoconductive emitter array with logarithmic spiral antennas. Appl Phys Lett 104(8):081122
46. Sung M, Moon S-R, Kim E-S, Cho S, Lee JK, Cho S-H, Kawanishi T, Song H-J (2021) Design considerations of photonic THz communications for 6g networks. IEEE Wirel Commun 28(5):185–191
47. Dyakonov M, Shur M (1993) Shallow water analogy for a ballistic field effect transistor: new mechanism of plasma wave generation by dc current. Phys Rev Lett 71(15):2465
48. Crabb J, Cantos-Roman X, Jornet JM, Aizin GR (2021) Hydrodynamic theory of the Dyakonov-Shur instability in graphene transistors. Phys Rev B 104(15):155440
49. Crabb J, Cantos-Roman X, Aizin GR, Jornet JM (2022) Amplitude and frequency modulation with an on-chip graphene-based plasmonic terahertz nanogenerator. IEEE Trans Nanotechnol 21:539–546
50. Singh PK, Aizin G, Thawdar N, Medley M, Jornet JM (2016) Graphene-based plasmonic phase modulator for terahertz-band communication. In: 2016 10th European conference on antennas and propagation (EuCAP). IEEE, pp 1–5
51. Singh A, Andrello M, Thawdar N, Jornet JM (2020) Design and operation of a graphene-based plasmonic nano-antenna array for communication in the terahertz band. IEEE J Select Areas Commun 38(9):2104–2117

52. Lin C, Li GYL (2016) Terahertz communications: an array-of-subarrays solution. IEEE Commun Mag 54(12):124–131
53. Abu-Surra S, Choi W, Choi S, Seok E, Kim D, Sharma N, Advani S, Loseu V, Bae K, Na I et al (2021) End-to-end 140 GHz wireless link demonstration with fully-digital beamformed system. In: 2021 IEEE international conference on communications workshops (ICC workshops). IEEE, pp 1–6
54. Durnin J (1987) Exact solutions for nondiffracting beams. I. The scalar theory. JOSA A 4(4):651–654
55. Durnin J, Miceli J, Eberly JH (1988) Comparison of Bessel and Gaussian beams. Opt Lett 13(2):79–80
56. Siviloglou G et al (2007) Observation of accelerating airy beams. Phys Rev Lett 99(21):213901
57. Tao H, Strikwerda A, Fan K, Padilla W, Zhang X, Averitt R (2009) Reconfigurable terahertz metamaterials. Phys Rev Lett 103(14):147401
58. Akyildiz IF, Jornet JM (2016) Realizing ultra-massive mimo (1024 × 1024) communication in the (0.06–10) terahertz band. Nano Commun Netw 8:46–54
59. Singh A, Alqaraghuli AJ, Jornet JM (2022) Wavefront engineering at terahertz frequencies through intelligent reflecting surfaces. In: 2022 IEEE 23rd international workshop on signal processing advances in wireless communication (SPAWC). IEEE, pp 1–5
60. Abdellatif H, Ariyarathna V, Petrushkevich S, Madanayake A, Jornet JM (2022) Real-time ultra-broadband software-defined radio platform for terahertz communications. In: Proceedings of the IEEE conference on computer communications workshops (INFOCOM) – demo tracks. IEEE
61. Sen P, Ariyarathna V, Madanayake A, Jornet JM (2021) A versatile experimental testbed for ultrabroadband communication networks above 100 GHz. Comput Netw 193:108092
62. Jornet JM, Akyildiz IF (2011) Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the terahertz band. IEEE Trans Wirel Commun 10(10):3211–3221
63. Gordon I, Rothman L, Hargreaves R, Hashemi R, Karlovets E, Skinner F, Conway E, Hill C, Kochanov R, Tan Y et al (2022) The hitran2020 molecular spectroscopic database. J Quant Spectrosc Radiat Transf 277:107949
64. Han C, Jornet JM, Akyildiz I (2018) Ultra-massive mimo channel modeling for graphene-enabled terahertz-band communications. In: 2018 IEEE 87th vehicular technology conference (VTC spring). IEEE, pp 1–5
65. Petrov V, Moltchanov D, Jornet JM, Koucheryavy Y (2019) Exploiting multipath terahertz communications for physical layer security in beyond 5G networks. In: IEEE INFOCOM 2019 – IEEE conference on computer communications workshops (INFOCOM WKSHPS). IEEE, pp 865–872
66. Fang Z, Guerboukha H, Shrestha R, Hornbuckle M, Amarasinghe Y, Mittleman DM (2022) Secure communication channels using atmosphere-limited line-of-sight terahertz links. IEEE Trans Terahertz Sci Technol 12(4):363–369
67. Kokkoniemi J, Lehtomäki J, Juntti M (2016) A discussion on molecular absorption noise in the terahertz band. Nano Commun Netw 8:35–45
68. Ju S, Xing Y, Kanhere O, Rappaport TS (2021) Millimeter wave and sub-terahertz spatial statistical channel model for an indoor office building. IEEE J Select Areas Commun 39(6):1561–1575
69. Xing Y, Rappaport TS (2021) Propagation measurements and path loss models for sub-THz in urban microcells. arXiv preprint arXiv:2103.01151
70. Priebe S, Kurner T (2013) Stochastic modeling of THz indoor radio channels. IEEE Trans Wirel Commun 12(9):4445–4455
71. Sen P, Hall J, Polese M, Petrov V, Bodet D, Restuccia F, Melodia T, Jornet JM (2022) Terahertz communications can work in rain and snow: impact of adverse weather conditions on channels at 140 GHz. In: Proceedings of the 6th ACM workshop on millimeter-wave and terahertz networks and sensing systems, pp 13–18

72. Abadal S, Han C, Jornet JM (2019) Wave propagation and channel modeling in chip-scale wireless communications: a survey from millimeter-wave to terahertz and optics. IEEE Access 8:278–293
73. Elayan H, Shubair RM, Jornet JM, Johari P (2017) Terahertz channel model and link budget analysis for intrabody nanoscale communication. IEEE Trans Nanobioscience 16(6):491–503
74. Elayan H, Johari P, Shubair RM, Jornet JM (2017) Photothermal modeling and analysis of intrabody terahertz nanoscale communication. IEEE Trans Nanobioscience 16(8):755–763
75. Kokkoniemi J, Jornet JM, Petrov V, Koucheryavy Y, Juntti M (2021) Channel modeling and performance analysis of airplane-satellite terahertz band communications. IEEE Trans Veh Technol 70(3):2047–2061
76. Sümen G, Kurt GK, Görçin A (2022) A novel LFM waveform for terahertz-band joint radar and communications over inter-satellite links. In: GLOBECOM 2022–2022 IEEE global communications conference. IEEE, pp 6439–6444
77. Alqaraghuli AJ, Siles JV, Jornet JM (2023) The road to high data rates in space: terahertz vs. optical wireless communication. IEEE Aerosp Electron Syst Mag
78. Bodet DM, Jornet JM (2022) Impact of antenna element directivity and reflection-interference on line-of-sight multiple input multiple output terahertz systems. In: 2022 3rd URSI Atlantic and Asia Pacific radio science meeting (AT-AP-RASC). IEEE, pp 1–4
79. Balanis CA (2016) Antenna theory: analysis and design. Wiley
80. Monnai Y, Altmann K, Jansen C, Hillmer H, Koch M, Shinoda H (2013) Terahertz beam steering and variable focusing using programmable diffraction gratings. Opt Express 21(2):2347–2354
81. Fu X, Yang F, Liu C, Wu X, Cui TJ (2020) Terahertz beam steering technologies: from phased arrays to field-programmable metasurfaces. Adv Opt Mater 8(3):1900628
82. Li S, Wang J (2017) Adaptive free-space optical communications through turbulence using self-healing Bessel beams. Sci Rep 7(1):43233
83. Chu X (2012) Analytical study on the self-healing property of Bessel beam. Eur Phys J D 66:1–5
84. Sen P, Pados DA, Batalama SN, Einarsson E, Bird JP, Jornet JM (2020) The Teranova platform: an integrated testbed for ultra-broadband wireless communications at true terahertz frequencies. Comput Netw 179:107370
85. Hermelo MF, Shih P-TB, Steeg M, Ng'oma A, Stöhr A (2017) Spectral efficient 64-QAM-OFDM terahertz communication link. Opt Express 25(16):19360–19370
86. Tarboush S, Sarieddeen H, Alouini M-S, Al-Naffouri TY (2022) Single-versus multicarrier terahertz-band communications: a comparative study. IEEE Open J Commun Soc 3:1466–1486
87. J. M. Jornet and I. F. Akyildiz, "Femtosecond-long pulse-based modulation for terahertz band communication in nanonetworks," IEEE Trans Commun, vol. 62, no. 5, pp. 1742–1754, May 2014
88. Boronin P, Petrov V, Moltchanov D, Koucheryavy Y, Jornet JM (2014) Capacity and throughput analysis of nanoscale machine communication through transparency windows in the terahertz band. Nano Commun Netw 5(3):72–82
89. Han C, Akyildiz IF (2014) Distance-aware multi-carrier (DAMC) modulation in terahertz band communication. In: 2014 IEEE international conference on communications (ICC). IEEE, pp 5461–5467
90. Hossain Z, Jornet JM (2019) Hierarchical bandwidth modulation for ultra-broadband terahertz communications. In: ICC 2019–2019 IEEE international conference on communications (ICC). IEEE, pp 1–7
91. Bodet D, Sen P, Hossain Z, Thawdar N, Jornet JM (2022) Hierarchical bandwidth modulations for ultra-broadband communications in the terahertz band. IEEE Trans Wirel Commun
92. Sen P, Ariyarathna V, Jornet JM (2022) An optimized M-ary amplitude phase shift keying scheme for ultrabroadband terahertz communication. In: 2022 IEEE 19th annual consumer communications & networking conference (CCNC). IEEE, pp 661–666

93. Bosso C, Sen P, Cantos-Roman X, Parisi C, Thawdar N, Jornet JM (2021) Ultrabroadband spread spectrum techniques for secure dynamic spectrum sharing above 100 GHz between active and passive users. In: 2021 IEEE international symposium on dynamic spectrum access networks (DySPAN). IEEE, pp 45–52
94. Sen P, Pandey H, Jornet JM (2020) Ultra-broadband chirp spread spectrum communication in the terahertz band. In: Next-generation spectroscopic technologies XIII, vol 11390. SPIE, pp 7–18
95. Jornet JM, Akyildiz IF (2011) Low-weight channel coding for interference mitigation in electromagnetic nanonetworks in the terahertz band. In: 2011 IEEE international conference on communications (ICC). IEEE, pp 1–6
96. Sarieddeen H, Médard M, Duffy KR (2022) Grand for fading channels using pseudo-soft information. In: GLOBECOM 2022–2022 IEEE global communications conference. IEEE, pp 3502–3507
97. Xia Q, Jornet JM (2019) Expedited neighbor discovery in directional terahertz communication networks enhanced by antenna side-lobe information. IEEE Trans Veh Technol 68(8):7804–7814
98. Xia Q, Hossain Z, Medley M, Jornet JM (2019) A link-layer synchronization and medium access control protocol for terahertz-band communication networks. IEEE Trans Mob Comput 20(1):2–18
99. Yao X-W, Jornet JM (2016) TAB-MAC: assisted beamforming mac protocol for terahertz communication networks. Nano Commun Netw 9:36–42
100. Zhang X, Han C, Wang X (2021) Dual-radio-assisted (DRA) mac protocols for distributed terahertz networks. IEEE Open J Vehic Technol 2:111–124
101. Xia Q, Jornet JM (2021) Multi-hop relaying distribution strategies for terahertz-band communication networks: a cross-layer analysis. IEEE Trans Wirel Commun 21(7):5075–5089
102. Xia Q, Jornet JM (2020) Routing protocol design for directional and buffer-limited terahertz communication networks. In: ICC 2020–2020 IEEE international conference on communications (ICC). IEEE, pp 1–7
103. Polese M, Jornet JM, Melodia T, Zorzi M (2020) Toward end-to-end, full-stack 6G terahertz networks. IEEE Commun Mag 58(11):48–54
104. Hossain Z, Xia Q, Jornet JM (2018) Terasim: an ns-3 extension to simulate terahertz-band communication networks. Nano Commun Netw 17:36–44
105. Polese M, Ariyarathna V, Sen P, Siles JV, Restuccia F, Melodia T, Jornet JM (2022) Dynamic spectrum sharing between active and passive users above 100 GHz. Commun Eng 1(1):6
106. Polese M, Cantos-Roman X, Singh A, Marcus MJ, Maccarone TJ, Melodia T, Jornet JM (2021) Coexistence and spectrum sharing above 100 GHz. arXiv preprint arXiv:2110.15187
107. Petrov V, Kurner T, Hosako I (2020) IEEE 802.15. 3d: first standardization efforts for sub-terahertz band communications toward 6G. IEEE Commun Mag 58(11):28–33
108. Wang CX, Haider F, Gao XQ et al (2014) Cellular architecture and key technologies for 5G wireless communication networks. IEEE Commun Mag 52:122–130
109. Giordani M et al (2020) Toward 6G networks: use cases and technologies. IEEE Commun Mag 58(3):55–61
110. Chanclou P, Neto LA, Grzybowski K, Tayq Z, Saliou F, Genay N (2018) Mobile fronthaul architecture and technologies: a RAN equipment assessment [invited]. IEEE/OSA J Opt Commun Netw 10(1):A1–A7. https://doi.org/10.1364/JOCN.10.0000A1
111. Giordani M, Polese M, Mezzavilla M, Rangan S, Zorzi M (2019) Towards 6G networks: use cases and technologies. arXiv:1903.12216
112. Piran M, Suh D (August 2019) Learning-driven wireless communications, towards 6G. arXiv:1908.07335
113. Alzenad M et al (2018) FSO-based vertical backhaul/fronthaul framework for 5G+ wireless networks. IEEE Commun Mag 56(1):218–224
114. Gao Y, Hu S, Tang W, Li Y, Sun Y, Huang D, Cheng S, Li X (2018) Physical layer security in 5g based large scale social networks: opportunities and challenges. IEEE Access 6:26350–26357

115. Wu Y, Khisti A, Xiao C, Caire G, Wong K, Gao X (2018) A survey of physical layer security techniques for 5G wireless networks and challenges ahead. IEEE J Sel Areas Commun 36(4):679–695
116. Osseiran A et al (2014) Scenarios for the 5G mobile and wireless communications: the vision of the METIS project. IEEE Commun Mag 52(5):26–35
117. Guo H, Wang Y, Liu J, Kato N (2022) Super-Broadband Optical Access Networks (OANs) in 6G: vision, architecture, and key technologies. IEEE Wirel Commun. https://doi.org/10.1109/MWC.007.2100550
118. Arnon S, Barry JR, Karagiannidis GK, Schober R, Uysal M (2012) Advanced optical wireless communication systems. Cambridge University Press, New York
119. Chan VWS (2006) Free-space optical communications. J Lightwave Tech 24(12):4750–4762
120. Nirmalathas A, Song T, Edirisinghe S, Wang K, Lim C, Wong E, Ranaweera C, Alameh K (2021) Indoor optical wireless access networks – recent progress [invited]. J Opt Commun Netw 13:A178–A186
121. Carruthers JB, Kahn JM (2000) Angle diversity for nondirected wireless infrared communication. IEEE Trans Commun 48(6):960–969
122. Ghassemlooy Z, Popoola WO (2010) Chapter 17: Terrestrial free-space optical communications. InTech, pp 356–392
123. Uysal M, Capsoni C, Ghassemlooy Z, Boucouvalas AC, Udvary E (2016) Optical wireless communications: an emerging technology. Springer, Berlin/Heidelberg, pp 1–634
124. Chowdhury MZ, Hossan MT, Islam A, Jang YM (2018) A comparative survey of optical wireless technologies: architectures and applications. IEEE Access 6:9819–10220
125. Song S, Liu Y, Xu T, Liao S, Guo L (2021) Channel prediction for intelligent FSO transmission system. Opt Express 29:27882–27899
126. Garlinska M, Pregowska A, Gutowska I, Osial M, Szczepanski J (2021) Experimental study of the free space optics communication system operating in the 8–12 mm spectral range. Electronics 10:875
127. Abdulameer LF, Sripati U, Kulkarni M (2013) Enhancement of security for free space optics based on reconfigurable chaotic technique. In: International conference on communication and electronics system design, pp 87600O–87600O-7
128. Ghassemlooy Z, Popoola W, Rajbhandari S (2019) Optical wireless communications: system and channel modelling with MATLAB. CRC Press Taylor and Francis Group, London
129. H. Kaushal and G. Kaddoum, “Optical communication in space: challenges and mitigation techniques,” IEEE Commun Surv Tuts, vol. 19, no. 1, pp. 57–96, First quarter 2017
130. Yao H, Ni X, Chen C, Li B, Zhang X, Liu Y, Tong S, Liu Z, Jiang H (2018) Performance of M-PAM FSO communication systems in atmospheric turbulence based on APD detector. Opt Express 26:23819–23830
131. Liu Y, Chen K, Song S, Pan Y, Liu Y, Guo L (2023) Reinforcement learning aided geometric shaping and self-canceling coherent detection for a PAM4 FSO communication system. J Opt Commun Netw 15(1):16–28. https://doi.org/10.1364/JOCN.468220
132. Amphawan A, Chaudhary S, Chan VWS (2014) 2 X 20 Gbps – 40 GHz OFDM Ro-FSO transmission with mode division multiplexing. J Eur Opt Soc Rapid Publ 9:14041. https://doi.org/10.2971/jeos.2014.14041
133. Mehtab S et al (2021) Performance evaluation of a 4 × 20-Gbps OFDM-based FSO link incorporating hybrid W-MDM techniques. Front Phys 9. https://doi.org/10.3389/fphy.2021.746779
134. Chatzidiamantis ND, Lioumpas AS, Karagiannidis GK, Arnon S (2011) Adaptive subcarrier PSK intensity modulation in free space optical systems. IEEE Trans Commun 59(5):1368–1377
135. Debbah M et al (2020) Smart radio environments empowered by reconfigurable intelligent surfaces: how it works, state of research, and the road ahead. IEEE J Sel Areas Commun 38(11):2450–2525. https://doi.org/10.1109/jsac.2020.3007211
136. Basar E et al (2019) Wireless communications through reconfigurable intelligent surfaces. IEEE Access 7:116753–116773. https://doi.org/10.1109/access.2019.2935192

137. Qingqing W et al (2021) Intelligent reflecting surface-aided wireless communications: a tutorial. IEEE Trans Commun 69(5):3313–3351. https://doi.org/10.1109/tcomm.2021.3051897
138. Wang H, Zhang Z, Zhu B, Dang J, Wu L (2020) Two new approaches to optical IRSS: schemes and comparative analysis. arXiv:2012.15398
139. Yang L, Guo W, da Costa DB, Alouini M-S (2020) Free-space optical communication with reconfigurable intelligent surfaces. arXiv:2012.00547
140. Zhang Q, Yang Y, Guo C, Zhou X, Yao Y, Lau APT, Lu C (2019) Algorithms for blind separation and estimation of transmitter and receiver IQ imbalances. J Lightwave Technol 37:2201–2208
141. Rahman AK, Anuar MS, Aljunid SA, Junita MN (2008) Study of rain attenuation consequence in free space optic transmission. In: Proceedings of the 2nd Malaysia conference on photonics telecommunication technologies (NCTT-MCP '08). IEEE, Putrajaya, pp 64–70
142. Hayle ST et al (2021) Integration of fiber and FSO network with fault-protection for optical access network. Opt Commun 484:126676
143. Kramer G, Pesavento G (2002) Ethernet passive optical network (EPON): building a next-generation optical access network. Commun Mag IEEE 40:66–73
144. Leitgeb E, Gebhart M, Birnbacher U (2005) Optical networks, last mile access and applications. J Opt Fiber Commun Res 2:56–85
145. Dhasarathan V, Singh M, Malhotra J (2020) Development of high-speed FSO transmission link for the implementation of 5G and internet of things. Wirel Netw 26:2403–2412
146. Pham AT, Trinh PV, Mai VV, Dang NT, Truong C-T (2015) Proceedings of the Opto-Electronics and Communications Conference (OECC). Hybrid free-space optics/millimeter-wave architecture for 5G cellular backhaul networks (Shanghai, 2015)
147. Jurado-Navas A, Tatarczak A, Lu X, Vegas Olmos JJ, Garrido-Balsells JM, Tafur Monroy I (2015) 850-nm hybrid fiber/free-space optical communications using orbital angular momentum modes. Opt Express 23:33721–33732
148. Hume A (2017) Nasa taking first steps toward high-speed space internet [Online]. Available: https://www.nasa.gov/feature/goddard/2017/nasa-taking-first-steps-toward-high-speed-space-internet
149. Daehnick C et al (2020) Large LEO satellite constellations: will it be different this time? [Online]. Available: https://www.mckinsey.com/industries/aerospace-and-defense/our-insights/large-leo-satellite-constellations-will-it-be-different-this-time
150. Feng X, Jiang H, Wu Z, Wang T, Jiang H, Gao S (2019) 60 Gbit/s coherent wavelength-division multiplexing free-space optical modulating retro-reflector in a turbulence-tunable atmospheric cell. Opt Commun 448:111–115
151. European project VERTIGO reaches key milestone by achieving record free-space laser communication transmission speed. Thales Group, ThalesAlenia Space, 3 October 2022
152. Khan MTA, Shemis MA, Alkhazraji E, Ragheb AM, Esmail MA, Fathallah HA, Alshebeili SA, Khan MZM (2017) Optical wireless communication at 100 Gb/s using L-band quantum-dash laser. In: Conference on lasers and electro-optics Pacific Rim. Optical Society of America, Washington, DC, pp 1–3
153. Sharma V, Kumar N (2014) Improved analysis of 2.5 Gbps-inter-satellite link (ISL) in inter-satellite optical wireless communication (ISOWC) system. Opt Commun 286:99–102
154. Fried DL (1967) Scintillation of a ground-to-space laser illuminator. J Opt Soc Am 57(8):980–983
155. Pompili D, Akyildiz IF (2009) Overview of networking protocols for underwater wireless communications. IEEE Commun Mag 47:97–102
156. Zhang Z et al (2019) 6G wireless networks: vision, requirements, architecture, and key technologies. IEEE Veh Technol Mag 14(3):28–41
157. Zeng Z, Fu S, Zhang H, Dong Y, Cheng J (2017) A survey of underwater optical wireless communications. IEEE Commun Surv Tut 19(1):204–238
158. Vali Z, Gholami A, Ghassemlooy Z, Michelson DG, Omoomi M, Noori H (2017) Modeling turbulence in underwater wireless optical communications based on Monte Carlo simulation. J Opt Soc Am A 34:1187–1193

159. Gabriel C, Khalighi MA, Bourennane S, Léon P, Rigaud V (2013) Monte Carlo-based channel characterization for underwater optical communication systems. J Opt Commun Netw 5:1–12
160. Borken-Kleefeld J, Berntsen T, Fuglestvedt J (2010) Specific climate impact of passenger and freight transport. Environ Sci Technol 44:5700–5706
161. https://www.saveatrain.com/blog/why-choosing-to-travel-by-train-is-environmentally-friendly/. Accessed June 2022
162. Singh P, Bohara VA, Srivastava A (12, 2022., Art no 7355712) On the optimization of integrated terrestrial-air-underwater architecture using optical wireless communication for future 6G network. IEEE Photon J 14(6):1. https://doi.org/10.1109/JPHOT.2022.3210481
163. Urabe H et al (2012) High data rate ground-to-train FreeSpace optical communication system. Opt Eng Spec Sect Free-Space Laser Commun 51(3)
164. Mori K et al (2014) Fast handover mechanism for high data rate ground-to-train free-space optical communication system. IEEE Globecom Workshops (GC Wkshps), Austin, pp 499–504
165. Fan Q, Taheri M, Ansari N, Feng J, Rojas-Cessa R, Zhou M, Zhang T (2018) Reducing the impact of handovers in ground-to-train free space optical communications. IEEE Trans Veh Technol 67(2):1292–1301
166. Fathi-Kazerooni S et al (2018) Optimal positioning of ground base stations in free-space optical communications for high-speed trains. IEEE Trans Intell Transp Syst 19(6):1940–1949
167. Paudel R, Ghassemlooy Z, Le-Minh H, Rajbhandari S (2013) Modelling of free space optical link for ground-to-train communications using a Gaussian source. IET Optoelectron 7:1–8
168. Dang S, Amin O, Shihada B, Alouini M-S (2020) What should 6G be. Nat Electron 3(1):20–29
169. Andrews LC, Phillips RL (1998) Laser propagation through random media. SPIE Optical Engineering Press, Bellingham, pp 135–159
170. O'Shea, T. and Hoydis, J. Introduction to deep learning for the physical layer," IEEE Trans Cogn Commun Netw, vol. 3, no. 4, pp. 563–575, Dec. 2017, doi: https://doi.org/10.1109/TCCN.2017.2758370
171. Karanov B et al (2018) End-to-end deep learning of optical fiber communications. J Lightwave Technol 36(20):4843–4855. https://doi.org/10.1109/JLT.2018.2865109
172. Wang D et al (2020) Data-driven optical fiber channel modeling: a deep learning approach. J Lightwave Technol 38(17):4730–4743. https://doi.org/10.1109/JLT.2020.2993271
173. Chen W et al (2022) Deep learning-based channel modeling for free space optical communications. J Lightwave Technol. https://doi.org/10.1109/JLT.2022.3213519
174. Zhu X, Kahn JM (2003) Performance bounds for coded free-space optical communications through atmospheric turbulence channels. IEEE Trans Commun 51(8):1233–1239
175. Uysal M, Navidpour SM, Li J (2004) Error rate performance of coded free-space optical links over strong turbulence channels. IEEE Commun Lett 8(10):635–637
176. Wilson SG, Brandt-Pearce M, Cao Q, Baedke M (2005) Optical repetition MIMO transmission with multipulse PPM. IEEE J Sel Areas Commun 23(9):1901–1909
177. Barsimantov O, Nikulin VV (2011) Adaptive optimization of a free space laser communication system under dynamic link attenuation. J Opt Commun Netw 3(3):215
178. Karimi M, Uysal M (2012) Novel adaptive transmission algorithms for free-space optical links. IEEE Trans Commun 60(12):3808–3815. https://doi.org/10.1109/TCOMM.2012.091012.110550
179. Razavi M, Shapiro JH (2005) Wireless optical communications via diversity reception and optical preamplification. IEEE Trans Wirel Commun 4(3):975–983. https://doi.org/10.1109/TWC.2005.847102
180. Tsiftsis TA, Sandalidis HG, Karagiannidis GK, Uysal M (2009) Optical wireless links with spatial diversity over strong atmospheric turbulence channels. IEEE Trans Wirel Commun 8(2): 951–957. https://doi.org/10.1109/TWC.2009.071318

181. Abadi MM, Ghassemlooy Z, Zvanovec S, Bhatnagar MR, Wu Y (2017) Hard switching in hybrid FSO/RF link: investigating data rate and link availability. In: 2017 IEEE international conference on communications workshops (ICC workshops), 21–25 May 2017, pp 463–468, https://doi.org/10.1109/ICCW.2017.7962701
182. Huang Y, Huang H, Chen H, Alvarado JC, Fontaine NK, Mazur M, Zhang Q, Ryf R, Amezcua-Correa R, Song Y, Li Y, Min W (2022) Free-space optics communications employing elliptical-aperture multimode diversity reception under anisotropic turbulence. J Lightwave Technol 40:1502–1508
183. Chen J, Huang Y, Cai R, Zheng A, Yu Z, Wang T, Liu Z, Gao S (2020) Free-space communication turbulence compensation by optical phase conjugation. IEEE Photon J 12:7905611
184. Yi X, Shen C, Yue P, Wang Y, Ao Q, Zhao P (2019) Performance analysis for a mixed RF and multihop FSO communication system in 5G C-RAN. J Opt Commun Netw 11:452–464
185. Boluda-Ruiz R, Garcia-Zambrana A, Castillo-Vazquez C, Castillo-Vazquez B, Hranilovic S (2018) Amplify-and-forward strategy using MRC reception over FSO channels with pointing errors. J Opt Commun Netw 10:545–552
186. Koh YT, Davidson F (1989) Interleaved concatenated coding for the turbulent atmospheric direct detection optical communication channel. IEEE Trans Commun 37(6):648–651
187. Schmalen L, Aref V, Cho J et al (2014) Next generation error correcting codes for lightwave systems[C]. European Conference on Optical Communication (ECOC), pp 1–3
188. Zongcheng Z (2003) Principle and application of error correction coding. Electronics Industry Press, pp 237–238
189. Huixi L, Qingjun Z, Liying T (2012) Performance of concatenated codes over an APD-based optical system with estimated channel-state information. j russ laser res 33(1):1–9
190. Gappmair W, Hranilovic S, Leitgeb E (2011) OOK performance for terrestrial FSO links in turbulent atmosphere with pointing errors modeled by Hoyt distributions. Commun Letts 15(8):875–877
191. Kiasaleh K (2006) Performance of coherent DPSK freespace optical communication systems in K-distributed turbulence. IEEE Trans Commun 54(4):604–607
192. Fang J, Bi M, Xiao S, Yang G, Liu L, Zhang Y, Hu W (2018) Polarcoded MIMO FSO communication system over gamma-gamma turbulence channel with spatially correlated fading. J Opt Commun Netw 10:915–923
193. Anguita JA, Joaquin H, Djordjevic IB (2014) Coherent multimode OAM superpositions for multidimensional modulation. IEEE Photon J 6(2):1–12
194. Ming L, Milorad C, Yuzuru T (2014) Evaluation of channel capacities of OAM-based FSO link with real-time wavefront correction by adaptive optics. Opt Express 22(25):31337–31346
195. Cho J, Winzer PJ (2019) Probabilistic constellation shaping for optical fiber communications. J Lightw Technol 37(6):1590–1607
196. Chen Y-W et al (2020) Asynchronous multi-service fiber-wireless integrated network using UFMC and PS for flexible 5G applications. Proc Opt Fiber Commun Conf:Paper M2H.4
197. Zou P, Hu F, Li G, Chi N (2020) Optimized QAM order with probabilistic shaping for the nonlinear underwater VLC channel. Proc Opt Fiber Commun Conf:Paper Th1K.7
198. Guiomar FP et al (2020) Adaptive probabilistic shaped modulation for high capacity free-space optical links. J Lightw Technol 38(23):6529–6541
199. Buchali F, Steiner F, Böcherer G, Schmalen L, Schulte P, Idler W (2016) Rate adaptation and reach increase by probabilistically shaped 64-QAM: an experimental demonstration. J Lightw Technol 34(7):1599–1609
200. Matsumine T, Yankov MP, Forchhammer S (2022) Geometric constellation shaping for concatenated two-level multi-level codes. J Lightwave Technol 40:5557–5566
201. Giordani M, Zorzi M (2020) Satellite communication at millimeter waves: a key enabler of the 6G era. In: Proceedings of international conference on computer network communication (ICNC), Big Island, pp 383–388

202. Fernandes MA, Monteiro PP, Guiomar FP (2022) Free-space terabit optical interconnects. J. Lightw Technol 40(5):1519–1526
203. Nag B, Das A, Ansari IS et al (2018) Performance analysis of hybrid FSO systems using FSO/RF-FSO link adaptation. IEEE Photon J 10(3):1–17
204. M. Najafi, V. Jamali, and R. Schober, “Optimal relay selection for the parallel hybrid RF/FSO relay channel: non-buffer-aided and bufferaided designs,” IEEE Trans Commun, vol. 65, no. 7, pp. 2794–2810, Jul. 2017
205. Odeyemi KO, Owolawi PA (2019) Selection combining hybrid FSO/RF systems over generalized induced-fading channels. Opt Commun 433:159–167
206. Wu S, Li S, Lin Y, Zhou H (2022) Performance analysis of hybrid FSO/RF transmission assisted airborne free-space optical communication system. J Commun Inf Netw 7(3):252–258. https://doi.org/10.23919/JCIN.2022.9906939
207. Tapse H, Borah DK (2011) Hybrid optical/RF channel performance analysis for turbo codes. IEEE Trans Commun 59(5):1389–1399
208. Peng C et al (2022) Pointing-error correction of optical communication terminals on motion platforms based on a K-nearest neighbor algorithm. IEEE Photon J. https://doi.org/10.1109/JPHOT.2022.3220400
209. Peng C et al (2021) Modeling and correction of pointing errors in gimbalstype optical communication terminals on motion platforms. IEEE Photon J 13(3):1–15. https://doi.org/10.1109/JPHOT.2021.3078227
210. Pattanayak DR, Dwivedi VK, Karwal V, Yadav PK, Singh G (2022) Physical layer security analysis of multi-hop hybrid RF/FSO system in presence of multiple eavesdroppers. IEEE Photon J. https://doi.org/10.1109/JPHOT.2022.3226351
211. Lee J et al (2018) Spectrum for 5G: global status, challenges, and enabling technologies. IEEE Commun Mag 56(3):12–18
212. Habibi MA et al (2019) A comprehensive survey of RAN architectures toward 5G mobile communication system. IEEE Access 7:70371–70421
213. Rommel S et al (2020) Towards a scaleable 5G fronthaul: analog radio-over-fiber and space division multiplexing. J Lightwave Technol 38:5412–5422
214. Dat PT et al (2019) Seamless convergence of fiber and wireless systems for 5G and beyond networks. J Lightwave Technol 37(2):592–605
215. Andrews L, Phillips R (2005) Laser beam propagation through random media. SPIE Press, Bellingham
216. Nguyen D, Bohata J, Komanec M, Zvánovec S, Ortega B, Ghassemlooy Z (2019) Seamless 25 GHz transmission of LTE 4/16/64-QAM signals over hybrid SMF/FSO and wireless link. J Light Technol 37(24):6040–6047
217. Nguyen D-N et al (2019) M-QAM transmission over hybrid microwave photonic links at the K-band. Opt Express 27(23):33745–33756
218. Dat PT, Umezawa T, Kanno A, Yamamoto N, Kawanishi T (2021) Seamless fiber–wireless system in W-band using optical phase modulation and self-homodyne receiver. IEEE Photon Technol Lett 33(20):1159–1162
219. Rommel S et al (2020) Towards a scaleable 5G fronthaul: analog radio-over-fiber and space division multiplexing. J Light Technol 38(19):5412–5422
220. López JD, Cardona et al (2021) Power-over-fiber in a 10 km long multicore fiber link within a 5G fronthaul scenario. Opt Lett 46:5348–5351
221. Vallejo L, Ortega B, Nguyen DN, Bohata J, Almenar V, Zvánovec S (2020) Usability of a 5G fronthaul based on a DML and external modulation for M-QAM transmission over photonically generated 40 GHz. IEEE Access 8:223730–223742
222. Vallejo L et al (2021) On the 40 GHz remote versus local photonic generation for DML-based C-RAN optical fronthaul. J Light Technol 39(21):6712–6723
223. Kao HY, Ishimura S, Tanaka K, Nishimura K, Inohara R (2021) End-to-end demonstration of fiber-wireless fronthaul networks using a hybrid multi-IF-over-fiber and radio-over-Fiber system. IEEE Photon J 13(4):1–6

224. Bohata J et al (2021) Experimental comparison of DSB and CS-DSB mmW formats over a hybrid fiber and FSO fronthaul network for 5G. Opt Express 29:27768–27782
225. Varshney N, Puri P (2017) Performance analysis of decode-and-forward-based mixed MIMORF/FSO cooperative systems with source mobility and imperfect CSI. J Lightwave Technol 35(11):2070–2077
226. Lee E, Park J, Han D, Yoon G (2011) Performance analysis of the asymmetric dual-hop relay transmission with mixed RF/FSO links. IEEE Photon Technol Lett 23(21):1642–1644
227. Ansari IS, Yilmaz F, Alouini M (2013) Impact of pointing errors on the performance of mixed RF/FSO dual-hop transmission systems. IEEE Wirel Commun Lett 2(3):351–354
228. Samimi H, Uysal M (2013) End-to-end performance of mixed RF/FSO transmission systems. IEEE/OSA J Opt Commun Netw 5(11):1139–1144
229. Zhang J, Dai L, Zhang Y, Wang Z (2015) Unified performance analysis of mixed radio frequency/free-space optical dual-hop transmission systems. J Lightw Technol 33(11):2286–2293
230. Anees S, Bhatnagar MR (2015) Performance evaluation of decode-and-forward dual-hop asymmetric radio frequency-free space optical communication system. IET Optoelectron 9(5):232–240
231. Zedini E, Ansari IS, Alouini M (2015) Performance analysis of mixed Nakagami-m and Gamma Gamma dual-hop FSO transmission systems. IEEE Photon J 7(1):1–20
232. Miridakis NI, Matthaiou M, Karagiannidis GK (2014) Multiuser relaying over mixed RF/FSO links. IEEE Trans Commun 62(5):1634–1645
233. Salhab AM, Al-Qahtani FS, Radaydeh RM, Zummo SA, Alnuweiri H (2016) Power allocation and performance of multiuser mixed RF/FSO relay networks with opportunistic scheduling and outdated channel information. J Lightw Technol 34(13):3259–3272
234. Obeed M, Salhab AM, Alouini M-S, Zummo SA (2019) On optimizing VLC networks for downlink multi-user transmission: a survey. IEEE Commun Surv Tutor PP(c):1–1
235. Stefan I, Haas H (2014) Hybrid visible light and radio frequency communication systems. In: 2014 IEEE 80th vehicular technology conference (VTC2014-Fall), pp 1–5
236. Wang Y, Basnayaka DA, Haas H (2015) Dynamic load balancing for hybrid Li-Fi and RF indoor networks. In: 2015 IEEE international conference on communication workshop (ICCW), pp 1422–1427
237. Basnayaka DA, Haas H (2015) Hybrid RF and VLC systems: improving user data rate performance of VLC systems. In: 2015 IEEE 81st vehicular technology conference (VTC Spring), pp 1–5
238. Li X, Zhang R, Hanzo L (2015) Cooperative load balancing in hybrid visible light communications and WiFi. IEEE Trans Commun 63(4):1319–1329
239. Wang Y, Haas H (2015) Dynamic load balancing with handover in hybrid Li-Fi and Wi-Fi networks. J Lightw Technol 33(22):4671–4682
240. Wang Y, Basnayaka DA, Wu X, Haas H (2017) Optimization of load balancing in hybrid LiFi/RF networks. IEEE Trans Commun 65(4):1708–1720
241. Wu X, Haas H (2017) Access point assignment in hybrid LiFi and WiFi networks in consideration of LiFi channel blockage. In: 2017 IEEE 18th international workshop on signal processing advances in wireless communications (SPAWC), pp 1–5
242. Wang Y, Wu X, Haas H (2016) Fuzzy logic based dynamic handover scheme for indoor Li-Fi and RF hybrid network. In: 2016 IEEE international conference on communications (ICC), pp 1–6
243. Gupta A, Sharma N, Garg P, Alouini M (2017) Cascaded FSO-VLC communication system. IEEE Wirel Commun Lett 6(6):810–813
244. Petkovic M, Cvetkovic A, Narandzic M (2018) Outage probability analysis of RF/FSO-VLC communication relaying system. In: 2018 11th international symposium on communication systems, networks digital signal processing (CSNDSP), pp 1–5
245. Vats A, Aggarwal M, Ahuja S (2019) Outage and error analysis of three hop hybrid VLC/FSO/VLC-based relayed optical wireless communication system. Trans Emerg Telecommun Technol 30(5):e3544

246. Huang Z, Wang Z, Huang M, Li W, Lin T, He P, Ji Y (2017) Hybrid optical wireless network for future SAGO-integrated communication based on FSO/VLC heterogeneous interconnection. IEEE Photon J 9(2)
247. Pesek P, Zvanovec S, Chvojka P, Ghassemlooy Z, Haigh PA (2019) Demonstration of a hybrid FSO/VLC link for the last mile and last meter networks. IEEE Photon J 11(1):1–7
248. Chi N, Zhou Y, Wei Y, Hu F (2020) Visible light communication in 6G: advances, challenges, and prospects. IEEE Veh Technol Mag 15(4):93–102
249. Yahia S, Meraihi Y, Ramdane-Cherif A, Gabis AB, Eldeeb HB (2022) Performance evaluation of vehicular visible light communication based on angle-oriented receiver. Comput Commun 191:500–509
250. Ali MF, Jayakody DNK, Li Y (2022) Recent trends in underwater visible light communication (UVLC) systems. IEEE Access 10:22169–22225
251. Boubezari R, Le Minh H, Ghassemlooy Z, Bouridane A (2016) Smartphone camera based visible light communication. J Lightwave Technol 34(17):4121–4127
252. Teli S, Cahyadi WA, Chung YH (2017) Optical camera communication: motion over camera. IEEE Commun Mag 55(8):156–162
253. Cahyadi WA, Kim YH, Chung YH, Ahn C-J (2016) Mobile phone camera-based indoor visible light communications with rotation compensation. IEEE Photon J 8(2):1–8
254. Hsu K-L, Wu Y-C, Chuang Y-C, Chow C-W, Liu Y, Liao X-L, Lin K-H, Chen Y-Y (2020) CMOS camera based visible light communication (VLC) using grayscale value distribution and machine learning algorithm. Opt Express 28(2):2427–2432
255. Liu W, Xu Z (2020) Some practical constraints and solutions for optical camera communication. Phil Trans R Soc A 378(2169):20190191
256. Jang M (2021) IEEE 802.15 WPAN 15.7 amendment-optical camera communications study group (SG 7a) [Online]. Available: https://www.ieee802.org/15/pub/SG7a.html
257. Le N-T, Jang YM (2016) Performance evaluation of MIMO optical camera communications based rolling shutter image sensor. In: 2016 eighth international conference on ubiquitous and future networks (ICUFN). IEEE, pp 140–144
258. Tuan NM, Phuong TV, Do T-H, Khanh NTV (2021) An highly realistic optical camera communication simulation framework for internet of things applications. In: 2021 21st ACIS international winter conference on software engineering, artificial intelligence, networking and parallel/distributed computing (SNPD-Winter). IEEE, pp 240–242
259. "Photron fastcam SA-Z." [Online]. Available: https://photron.com/fastcam-sa-z/
260. Nguyen H, Nguyen V, Nguyen C, Bui V, Jang Y (2021) Design and implementation of 2DMIMO-based optical camera communication using a light-emitting diode array for long-range monitoring system. Sensors 21(9):3023
261. Nguyen T, Islam A, Hossan T, Jang YM (2017) Current status and performance analysis of optical camera communication technologies for 5G networks. IEEE Access 5: 4574–4594
262. Huang W, Tian P, Xu Z (2016) Design and implementation of a real-time CIM-MIMO optical camera communication system. Opt Express 24(21):24 567–24 579
263. Luo P, Zhang M, Ghassemlooy Z, Le Minh H, Tsai H-M, Tang X, Png LC, Han D (2015) Experimental demonstration of RGB LED-based optical camera communications. IEEE Photon J 7(5):1–12
264. Hassan NB, Ghassemlooy Z, Zvanovec S, Biagi M, Vegni AM, Zhang M, Luo P (2019) Non-line-of-sight MIMO space-time division multiplexing visible light optical camera communications. J Lightwave Technol 37(10):2409–2417
265. Shahjalal M, Hasan M, Chowdhury MZ, Jang YM et al (2019) Smartphone camera-based optical wireless communication system: requirements and implementation challenges. Electronics 8(8):913
266. Duque A, Stanica R, Rivano H, Desportes A (2020) Analytical and simulation tools for optical camera communications. Comput Commun 160:52–62
267. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444

268. Younus OI, Hassan NB, Ghassemlooy Z, Haigh PA, Zvanovec S, Alves LN, Le Minh H (2020) Data rate enhancement in optical camera communications using an artificial neural network equaliser. IEEE Access 8:42 656–42 665
269. Yu K, He J, Huang Z (2020) Decoding scheme based on CNN for mobile optical camera communication. Appl Opt 59(23):7109–7113
270. Teli SR, Cahyadi WA, Chung YH (2018) Trained neurons-based motion detection in optical camera communications. Opt Eng 57(4):040501
271. Tanaka M. Give me four: residual interpolation for image demosaicing and upsampling [Online]. Available: http://www.ok.sc.e.titech.ac.jp/~mtanaka/research.html
272. Li X, Gunturk B, Zhang L (2008) Image demosaicing: a systematic survey. In: Visual communications and image processing 2008, vol 6822. Int Soc Opt Photon, p 68221J
273. Arnon S (2015) Visible light communication. Cambridge University Press
274. Cao F, Gong X, He C, Ngai T (2017) Removing the effect of blooming from potential energy measurement by employing total internal reflection microscopy integrated with video microscopy. J Colloid Interface Sci 503:142–149
275. Mansurov N (2018) Understanding ISO, shutter speed and aperture – a beginner's guide. Photographylife 23
276. Chow C-W, Li Z-Q, Chuang Y-C, Liao X-L, Lin K-H, Chen Y-Y (2019) Decoding CMOS rolling-shutter pattern in translational or rotational motions for VLC. IEEE Photon J 11(2):1–5
277. Takai I, Ito S, Yasutomi K, Kagawa K, Andoh M, Kawahito S (2013) LED and CMOS image sensor based optical wireless communication system for automotive applications. IEEE Photon J 5(5):6 801 418–6 801 418
278. Yamazato T, Takai I, Okada H, Fujii T, Yendo T, Arai S, Andoh M, Harada T, Yasutomi K, Kagawa K et al (2014) Image-sensor-based visible light communication for automotive applications. IEEE Commun Mag 52(7):88–97
279. C. Danakis, M. Afgani, G. Povey, I. Underwood, and H. Haas, "Using a CMOS camera sensor for visible light communication," 2012 IEEE Globecom workshops. IEEE, 2012, pp. 1244–1248
280. Kim BW, Yoo J-H, Jung S-Y (2020) Design of streaming data transmission using rolling shutter camera-based optical camera communications. Electronics 9(10):1561

6G Wireless Architectures

Giacomo Oliveri, Francesco Zardi, Arianna Benoni, Marco Salucci, Andrea Massa, Ernestina Cianca, Debashisha Mishra, Evgenii Vinogradov, and Enrico Natalizio

1 EM Static and Reconfigurable Passive Skins Within the Smart ElectroMagnetic Environment Paradigm

Giacomo Oliveri • Francesco Zardi • Arianna Benoni • Marco Salucci • Andrea Massa

---

G. Oliveri · F. Zardi · A. Benoni · M. Salucci
ELEDIA Research Center (ELEDIA@UniTN – University of Trento), DICAM – Department of Civil, Environmental, and Mechanical Engineering, Trento, Italy
e-mail: giacomo.oliveri@unitn.it; francesco.zardi@unitn.it; arianna.benoni@unitn.it; marco.salucci@unitn.it

A. Massa
ELEDIA Research Center (ELEDIA@UniTN – University of Trento), DICAM – Department of Civil, Environmental, and Mechanical Engineering, Trento, Italy

ELEDIA Research Center (ELEDIA@UESTC – UESTC) School of Electronic Engineering, Chengdu, China

ELEDIA Research Center (ELEDIA@TSINGHUA – Tsinghua University), Beijing, China
e-mail: andrea.massa@unitn.it; andrea.massa@uestc.edu.cn; andrea.massa@tsinghua.edu.cn

E. Cianca (✉)
Center for Teleinfrastructure (CTIF) and CNIT University of Rome Tor Vergata, Rome, Italy
e-mail: cianca@ing.uniroma2.it

D. Mishra
Université de Lorraine, CNRS, Loria, France
e-mail: debashisha.mishra@univ-lorraine.fr

E. Vinogradov · E. Natalizio
Technology Innovation Institute (TII), Abu Dhabi, United Arab Emirates
e-mail: evgenii.vinogradov@tii.ae; enrico.natalizio@tii.ae

V. Loscri et al. (eds.), *The Road towards 6G: Opportunities, Challenges, and Applications*, https://doi.org/10.1007/978-3-031-42567-7_4

1.1 Principles, Opportunities, and Challenges

The transition between subsequent generations of mobile communications standards has traditionally required to deploy upgraded versions of both the network infrastructure (base stations, back-haul links, core transport equipments) as well as of the end-user equipments (e.g., the antennas and wireless modems in phones, laptops, computers, as well as communications-enabled devices) [1, 2]. Such transitions have been made possible by the design and development of more advanced base stations/access points and user terminals, often featuring wider bandwidths, more complex wireless protocols, and/or an increased transmitted power [1, 2]. Unfortunately, such a strategy may not be sufficient to meet the demanding requirements of 6G wireless networks in terms of several key performance indicator including data transfer rate, overall power consumption, and coverage [1, 2]. In fact, the chronic lack of available spectrum is fostering the exploitation of higher carrier frequencies with respect to previous generation systems, including mmWave bands and beyond. Although this choice is known to enable considerable opportunities in terms of achievable data rates and network throughput, the more severe path loss and shadowing effects experienced as the frequency increases can prevent the wireless systems to achieve adequate coverage areas and to reach the desired quality-of-service for all the users within the reach of the base stations / access points.

Starting from these premises, a thorough revision of the fundamental approach to mobile network evolution has recently emerged in the academic and industrial communities. More specifically, the idea that the *propagation environment* can be included within the wireless system design process has fostered a revolutionary vision to the implementation of mobile networks, resulting in the introduction of the "Smart ElectroMagnetic Environment" (SEME) concept [2, 3]. The transformative SEME paradigm primarily originated from the demonstrated possibility that the interaction between wireless waves and buildings/urban structures could be manipulated in order to improve the coverage, to enhance the data rate, to reduce the power consumption and to increase the energy efficiency and reliability of the overall wireless network without the need to install additional base stations [2, 3]. This is made possible by introducing in the environment suitably designed structures such as static passive electromagnetic skins (SP-EMS), reconfigurable passive EMS (RP-EMS, also labeled as reconfigurable intelligent surfaces or RIS), smart repeaters (SR), and integrated access and backhaul (IAB) nodes [3, 4]. It is then expected that 6G networks will benefit from the introduction of SEME concepts and devices enabling the transition of the propagation environment from an uncontrollable element of the communication chain to an integrated and cooperative component of it [2].

The opportunities enabled by the introduction of the SEME vision within communications systems are manifold. As a matter of fact, envisaged applicative scenarios in terms of *6G mobile communications* include

- Coverage improvement and blind spot reduction through static passive EMS. Achieving effective coverage in urban scenarios will become an increasingly challenging task in 6G networks especially when dealing with mmWave carrier frequencies and beyond. The opportunities enabled by anomalous manipulation of waves interacting with buildings coated with EMSs (i.e., controlled focusing, footprint shaping, and non-Snell reflection) has been demonstrated to allow a considerable flexibility in wireless planning and improvement of coverage values without the need to install additional base stations [5, 6]. This is specifically relevant when planning under both quality-of-service and EM field exposure constraints is of interest [7].

- Adaptive wireless coverage. Unlike existing wireless networks, SEME-enabled scenarios involving reconfigurable technologies (e.g., RP-EMSs) will allow the cellular network manager to dynamically modify the urban coverage according to the network needs (e.g., different user densities, time of the day, or temporarily required quality of service in a specific area) [8]. This will be made possible by the capability to set a different "wave reflection configuration" from the environment to properly manipulate the wireless coverage as needed [8].
- Real-time propagation optimization. The presence of dynamically adjustable passive skins in the scenario will enable the wireless system to adapt the *environment response* to the requirements of the users in real time [1, 2, 9]. Such a revolutionary opportunity will pave the way to the constant optimization of the wireless channel features in order to guarantee maximum throughput, minimum fading, mitigated shadowing effects, and improved overall quality [1, 2, 9]. Accordingly, wireless systems leveraging on this opportunity will overcome the limitations of traditional wireless systems, whose performance are unavoidably bounded by the inherent features of the scenario features and propagation properties [1, 2, 9].

It is also worth remarking that the adoption and implementation of SEME technologies is expected to have an impact at the system level in terms of reduction of the overall power consumption. As a matter of fact, the improvement of the wireless propagation features imply the possibility to either reduce the RF power of (or to remove completely) some of the base station/access points without compromising the resulting quality of service [2]. Such an opportunity is specifically relevant considering the converging needs to implement greener communication systems, minimize the carbon footprint, and mitigate potential exposure issues of wireless information networks.

Besides the communications applications, SEME technologies such as EMSs are envisaged to contribute to the evolution and implementation of *integrated sensing and communication* (ISAC) applicative scenarios [10–12]. Early examples of such a opportunities include (i) the possibility to combine the active beamforming of the base station and the passive beamforming of the RP-EMS to maximize the achievable sum-rate of the communication users while satisfying the constraint of beampattern similarity for radar sensing, the restriction of the RIS, and the transmit power budget [11]; (ii) the chance to combine radar and communication functionalities aided by RIS devices in order to maximize the achievable secrecy rate of the system [12]; (iii) the potentiality to share the same frequency bandwidth between a pair of communication transmitter and receiver and a radar through the exploitation of double-RIS-assisted coexistence strategy [10]. However, several other opportunities are envisaged whenever the concurrence of sensing and communications systems in the same bands and location is expected.

The opportunities enabled by SEME concepts are not only limited to EM spectrum and propagation management. As a matter of fact, the physical control enabled by EMSs is not limited in principle to electromagnetic waves, but multi-physics capabilities are feasible and already being studied [13, 14]. As a consequence, multi-physics processing enabled by SEME technologies can be a further opportunity to be exploited within 6G networks, for instance to introduce effective EM insulation devices within the environment (e.g., to minimize inter-cell interferences and improve the network quality) through electromagnetic-thermal conversion materials [14] as well as to yield flexure detection strategies based on mechanically-tunable EMS concepts in order to enable the network to indirectly sense important structural information concerning the wireless infrastructure (e.g., telecommunication tower vibrations) [13].

Despite the previous promising scenarios, several challenges still need to be addressed before the SEME vision can be translated to commercial applications which can be mass-produced at minimum costs at the industrial level. In fact, the implementation of EMS technologies is fundamentally based on the capability to design and fabricate thin planar structures capable of properly tailoring their macro-scale interactions with EM waves (e.g., reflection, transmission, diffraction phenomena) [2, 3, 15–17]. This is usually accomplished by employing regular arrangements of "meta-atoms" with a suitably designed periodicity (typically, half wavelength or below), whose local reflection properties are managed through the configuration of their micro-scale atomic structure [5, 17]. This can be obtained either statically by means of proper geometrical/physical design of each atom constituting the device [5, 17] or through the introduction of electronically-tunable components in each cell of the arising metasurface [8, 9, 18, 19]. Although such a concept has been widely adopted and validated in metamaterial science and engineering in the last decades [17], its customization to the SEME scenarios of interest must take into account the specific constraints of large-scale 6G wireless applications, including

- The need to achieve accurate wave manipulation capabilities and potential dynamic reconfigurability with inexpensive technologies even when large apertures are at hand. In fact, unlike standard metasurface technologies [16, 17, 20], EMSs are expected to be deployed on large-scale indoor and outdoor 6G scenarios. As a consequence, the cost-per-square-meter is expected to be as low as possible when mass produced. This is actually prevented using current RF and mmWave technologies since the designs, the materials, the components, and the production processes are not conceived towards this purpose.
- The resulting solutions must guarantee an effective degree of scalability and modularity [21]. The feasibility of SEME technologies from the manufacturing perspective requires the use of designs that are as modular as possible in order

to reduce the implementation complexity and costs as well as the speed of the deployment. In fact, non-customized or marginally customized EMSs imply simpler and faster planning, production, and installation. Unfortunately, many current design techniques rely on an accurate knowledge of the installation site in order to operate properly, hence demanding a customized approach to planning and deployment [3].

Several of the design and implementations discussed next aim at addressing the above challenges both from the methodological and technological viewpoint.

1.2 Design and Implementations

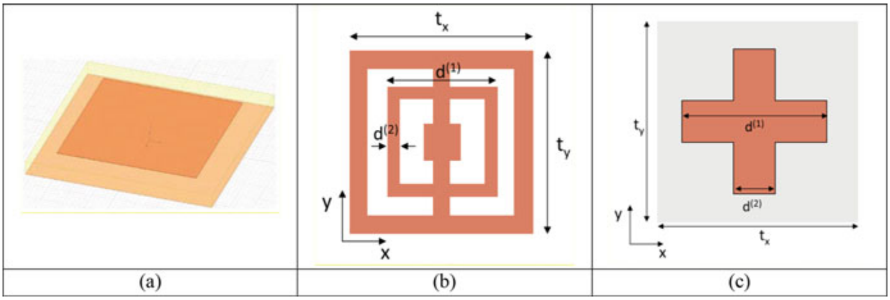
The design and implementation of EMS solutions within the SEME vision illustrated above requires to solve a set of interrelated but heterogeneous problems from the perspective of the available solution descriptors and degrees-of-freedom, the synthesis objectives and desired performance, and the constraints and choices in the technologies at hand. More specifically, the overall design can be split from the functional perspective into the following sub-problems:

- Micro-scale meta-atom design. The synthesis of the meta-atom geometry and architecture capable of implementing the local field manipulation functionality required for the SEME scenario at hand involves both methodological and technological challenges, since it has to account for the actually available materials, stack-up configurations, and electronically-reconfigurable components and in their combination [5, 8, 9, 17, 18].
- Meso-scale EMS synthesis. Once the elementary meta-atom has been designed, the implementation of an EMS (either static or dynamically reconfigurable) requires to define their proper aggregation and resulting architecture taking into account both the arising coverage/wireless control functionalities and the proper modularity constraints [5, 8, 17, 18].
- Macro-scale SEME planning. The deployment of EMSs in an indoor/outdoor SEME scenario demands the definition of the number, location, size, and design of each device taking into account the existing wireless network (i.e., access points / base stations), the environment topology, and the target network performance. This implies the solution of a complex multi-objective problem that accounts also for the implementation and installation costs of such solutions [22].

The previous subdivision also points out the intrinsic multi-scale nature of the SEME design and implementation problems [2]. Such a feature is a natural consequence of the need to address wireless propagation control functionalities at indoor/outdoor communications scale while acting on micro-scale surface descriptors and it is one of the fundamental motivation for the intrinsic complexity of SEME design problems.

With reference to the *micro-scale design problem*, several different strategies have been proposed in the recent literature to achieve wave control in a static or dynamic manner with inexpensive and light planar meta-atoms [5]. In this framework, the basic wireless manipulation functionality required by passive meta-atoms is to enable an accurate control of the *phase* of the locally reflected or transmitted wave, since its magnitude is usually constrained by that of the incident illumination. The fundamental objectives in terms of phase control include to guarantee both (i) a wide phase coverage, so that the achievable phase states cover at least a full 360 [deg] span [5, 17], and (ii) a phase/cell descriptor relation which is as linear as possible, so that the resulting design will be more robust to fabrication tolerances [5, 17]. However, the principle to achieve such a control is considerably different when dealing with SP or RP geometries.

As for the SP-EMS case (Fig. 1), the most commonly adopted strategy to achieve the desired phase control is based on the exploitation of a *parametric* approach in which the geometrical descriptors of the reference meta-atom architecture are modified to yield different phase responses [5, 23, 24]. With reference to PCB technologies in the lowest 6G bands of interest (i.e., sub-6 GHz) and assuming reflecting functionalities, the meta-atom thus usually consist of a cell that includes (i) a metallic groundplane backing the structure, (ii) at least one dielectric substrate, and (iii) at least one printed meta-atom patch [23]. The phase control is thus achieved by properly manipulating this latter element geometry, while the lattice spacing and the substrate are kept uniform across the entire structure. Elementary patches already demonstrated for SEME scenarios in 6G wireless applications are often based on canonical rectangular/square meta-atoms (Fig. 1a). However, several alternative geometries [3] also borrowed from reflectarray engineering [23, 25] may be employed, including layouts enabling wide phase coverages or independent polarization control such as Phoenix (Fig. 1b) and crossed-dipole (Fig. 1c) designs. According to such a paradigm, the resulting SP-EMS consists of a *quasi-periodic* planar layout in which the meta-atom descriptors (e.g., each patch side) smoothly varies across the aperture [5, 6].



(a) (b) (c)

Fig. 1 Examples of potential SP-EMS meta-atoms based on (**a**) rectangular patch architecture [23], (**b**) Phoenix architecture [24], and (**c**) crossed-dipole [24] concepts

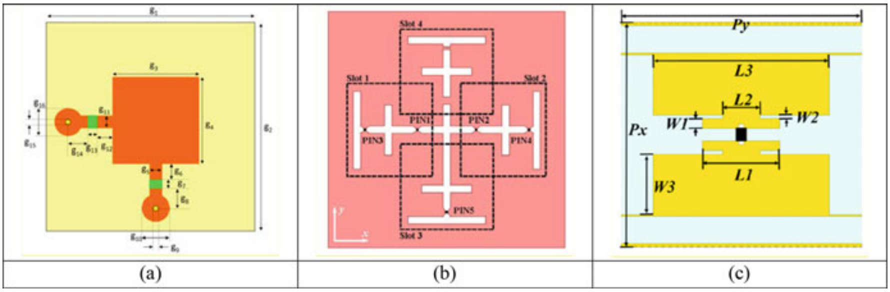


Fig. 2 Top view of recently proposed RP-EMS meta-atoms enabling (**a**) 1-bit control through top-mounted p-i-n diodes [8], (**b**) 2-bit control through the p-i-n installed in the slot-loaded plane [9], and (**c**) 3-bit control through a centrally-placed varactor [18]

A completely different concept is usually employed when designing RP-EMSs meta-atoms [8, 9, 18]. In fact, their implementation typically relies on geometrically identical elements that comprise electronically-reconfigurable discrete components [8, 9, 18]. The dynamic electronic configuration of the component in each meta-atom, which is typically a diode or a varactor, is enabling the local phase control in the RP-EMS [8, 9, 18]. Owing to the constraints associated to the SEME applicative scenario in terms of cost, complexity, and scalability, such an electronic control is usually heavily quantized, with popular approaches featuring 1-bit (Fig. 2a), 2-bit (Fig. 2b), or 3-bit (Fig. 2c) phase control per meta-atom.

Examples of configurations employed to this end shown in Fig. 2 comprise patches with shorting vias that are switchable through surface-mounted p-i-n diodes (Fig. 2a), slot-fed patches in which the slot plane comprises switchable sections (Fig. 2b), or patches including surface-mounted varactors to enable finer phase tuning (Fig. 2c). Several alternative configurations are being studied as well to further reduce costs and improve RP-EMS scalability [3].

It is also worth mentioning that more complex field manipulation functionalities such as frequency filtering, polarization control, and non-linear RF operations are still possible within the same framework [17], although seldom used in current SEME scenarios for 6G communications.

As regards the *meso-scale design problem*, the fundamental objective can be stated as the identification of (i) the unit cell descriptors (i.e., either physical/geometrical for SP-EMSs or electronic for RP-EMSs) that guarantee the desired overall wireless propagation control [5, 6, 8] and (ii) the EMS architecture to achieve such a goal.

As for the first objective, the most basic type of EM manipulation required by EMSs is the "anomalous reflection" (i.e., Non-Snell collimated reflection) [5, 17, 18]. This functionality can be theoretically achieved easily since a closed-form expression for the phase contribution required by each meta-atom in the EMS to yield a certain overall coherent reflection is available provided that the illuminating

EM field is known [5, 17, 18]. As such, coherent reflection is the most commonly employed wireless control function implemented by EMSs [3, 18].

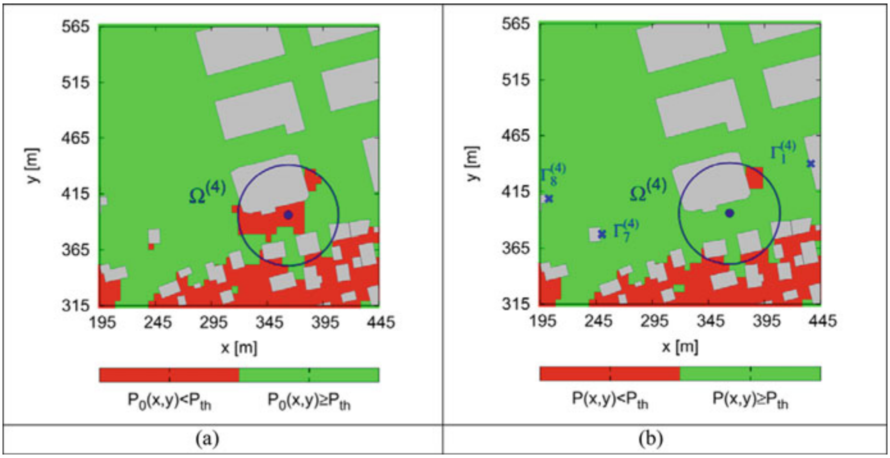
More complex wireless manipulation functionalities have been demonstrated using EMSs, as well [5, 6, 8]. More specifically, the concept of "footprint control" has been shown to enable the improvement of the network coverage also taking into account the user density in the urban scenario at hand [5, 6, 8]. The possibility to tailor the EMS reflection in order to achieve a uniform signal level or quality-of-service in a target area such as a street or a square has been then demonstrated both considering SP-EMS [5, 6] and RP-EMS architectures [8]. Owing to the unavailability of closed form design expressions, suitable non-linear optimization strategies often based on the System-by-Design (SbD) paradigm [26] have been adopted to this end.

As regards the "EMS architecture" design objective, both monolithic [5, 6, 8] and clustered/tiled arrangements [21] have been conceived. The first type of designs consists of regular arrangements of meta-atoms that are synthesized and fabricated in a single step for a specific functionality, hence resulting in a customized and optimal but *non-modular* design, especially when dealing with SP-EMSs [5, 6]. The modularity enabled by the "tiled" approach [21] is thus a fundamental advantage from the practical perspective within the SEME vision, since this can allow a more scalable approach to implementation, fabrication, and installation regardless of the scenario at hand [21]. Advanced design approaches have been already demonstrated in the literature that enable to determine the optimal trade-off implementative solutions assuring a suitable coverage of the areas of interest with the minimum complexity [21]. Owing to the non-trivial associated combinatorial problem, a binary multi-objective optimization method has been employed in this case [21].

Finally, the solution of the *macro-scale design problem* requires the optimal planning of EMSs installed on the building facades to enhance the received signal strength, and thus the wireless coverage and/or the Quality-of-Service in large-scale urban areas [22]. Accordingly, the unique challenges of the associated design problem are related to [22]

- the type of the degrees-of-freedom at hand, that is the presence/absence of the device, its size and orientation, and its configuration for each candidate location for the EMS in the urban scenario at hand;
- the huge scale of the propagation analysis required to assess the quality of a certain guess configuration (i.e., the actual improvement in terms of network-level performance thanks to the presence of the EMSs) also taking into account the several base stations / access points in the scenario, the potential multipath effects and arising multiple incidences on a single EMS, and the needs for statistical accuracy in the coverage predictions;
- the intrinsically complex constraints and objectives in terms of costs, the installation limitations on each building facade, and the overall number of EMSs.

The possibility to formulate and solve the arising planning problem within the SbD framework has been demonstrated [22]. More specifically, a set of optimal tradeoff



(a) (b)

Fig. 3 Demonstration of planning exploiting EMS architectures – Thresholded power map (**a**) without and (**b**) with EMS in urban environments, illustrating the possibility to overcome a blind spot region (highlighted by the blue circle) behind a building. (From [22])

solutions which jointly maximizes the level of power received within "no-coverage" regions and minimize the overall cost and environmental impact has been computed by means of a customized SbD process [22]. The generalization of such an approach to the case of cascaded reflections from multiple EMSs has been recently addressed, as well (Fig. 3).

1.3 Conclusions

The development of 6G mobile networks is expected to require a deep and thorough revision of the fundamental approaches to wireless system conception specifically focusing on the idea that the *propagation environment* can be included within the design process. In this framework, the introduction of the SEME concepts and devices is envisaged to open considerable opportunities both from the performance and from the cost/power consumption perspectives, potentially fostering a new *wireless revolution* in the field. However, several challenges are still being addressed to translate such SEME concepts into practical scenarios. Indeed, the scalability, cost, and proper design and planning of SEME devices such as SP-EMSs and RP-EMSs require careful attention to enable their effective exploitation in 6G networks. Such challenges are the fundamental reason for the considerable interest and advancements that are continuously demonstrated in the recent year within the community at the industrial and academic level [3].

Acknowledgments This work benefited from the networking activities carried out within the Project "MANTLES – Cloaking Metasurfaces for a new Generation of

Intelligent Antenna Systems", (Grant no. 2017BHFZKH) funded by the Italian Ministry of Education, University, and Research within the PRIN2017 Program (CUP: E64I19000560001), the Project "SPEED" (Grant No. 6721001) funded by National Science Foundation of China under the Chang-Jiang Visiting Professorship Program, and the Project "National Centre for HPC, Big Data and Quantum Computing (CN HPC)" funded by the European Union – NextGenerationEU within the PNRR Program (CUP: E63C22000970007).

2 LEO Constellations

Ernestina Cianca

2.1 Introduction

6G is the standard that will see the real integration of terrestrial and not terrestrial nodes, and also the convergence between communication, navigation and sensing. In such a vision, satellite systems will play a key role and LEO constellations will be undoubtedly the main players. After the uncertain start at the end the 90', when the first LEO constellations of satellites have been launched to provide mobile communication services, such as IRIDIUM, GLOBALSTAR, which have been commercially unsuccessful, we are nowadays witnesses of a re- newed interest in using LEO constellations, even with huge number of satellites (tens of thousands). Table 1 shows the main characteristics of currently operative LEO constellations [27]. Some of the constellations provide broadband services to fill the digital divide still present in some areas of the Earth. Some other are used to provide IoT services or Earth Observation. It is also worth highlighting that there is a growing interest to use dedicated or already deployed LEO constellations to provide Positioning, Navigation and Timing (PNT) Services in GNSS-denied environments. Standardization activities are on-going in 3GPP to integrate non terrestrial communication nodes and in particular satellite nodes to the terrestrial networks. In April 2022 the first global standard for terrestrial- satellite communications was published within 3GPP Release 17, specifying the features enabling 4G/5G systems to support a satellite component, with a special focus on LEO constellations. Release 18 of 3GPP has already started working on further improvements of the the performances and to provide new capabilities. However, to really take advantage of the satellite component in the 6G vision, a significant innovation breakthrough in technologies, techniques, and architectures is still needed. Such an innovation should also take into account an issue that could make the space an unsafe place to operate, e.g. the issue of space sustainability.

The renewed interest in the space component has led to an uncontrolled growth of the number of satellites that have been launched. Active satellites and huge amount

Table 1 Summary of current LEO constellations

Constellation	Application	Altitude per orbit (Km)	Total number	Frequency bands downlink
Amazon Kuiper	Broadband Comm	590,610,630	3236	L2–1227.60 MHz
Starlink	Broadband Comm	335.9 340.8 345.6 550 1110 1130 1275 1325	11,943	K-band: 17.8–18-5 GHz 18.8–19.3 GHz 19.7–20.2 GHz V-band: 37.5–42.0 GHz
Oneweb	Broadband Comm	1200	358	Ku/Ka-band:
Astrocast	IoT	600	80	L-band S-band
Hyber	IoT	600	600	UHF: 400,15–401,0 MHz
Myriota	IoT	600	50	VHF: 156–165 MHz UHF: 399–403 MHz
Iridium next	Narrowband Comm	780	66	L-band: 1.616–1.63 GHz K-band: 19.3–19.7 GHz
ICEYE	Earth observation	570	18	X-band: 9,65 GHz

of debris will soon increase the risk to invest in space systems. Therefore, the design of future space systems should have the long-term space sustainability as a key requirement.

This Chapter first provides an overview of the main characteristics of current and near future LEO satellite constellations, in particular highlighting the evolution in the payload and antenna technology, the used frequency bands. Then the state-of-the-art of the standardization activities in 3GPP standards, to integrate the satellite component into the terrestrial network, is presented. The Chapter will then focus on LEO constellation applications and services, with a focus on Internet of Remote Things scenarios [28] and the use of LEO constellations for PNT services, highlighting hot research areas. Finally, the Chapter will present the very urgent issue of sustainable and safe use of the space, which also open novel challenges and research directions.

2.2 Main Characteristics of Current LEO Constellations

As shown in Table 1, most of the constellations have an altitude lower than 600 km. Starlink uses even the so-called Very Low Earth Orbits, around 300 km. Only Oneweb satellites have an altitude higher than 1000 km. A low altitude helps in getting a good SNRs without requiring high transmit power or big antennas, and hence it is a preferred solution both for IoT services, where terminal could have stringent power and dimensions constraints, and for broadband services where high frequency bands are used to get the necessary bandwidth which are subject to high level of attenuation. The reduced distances also helps to achieve a lower latency, which could be crucial in some 5G or 6G future applications. However, such low altitudes also imply:

- the need of a huge number of satellites to guarantee a global coverage.
- high value of the Doppler shifts as satellites moves at very high mean velocity (around 7 km/s). Without compensation, the doppler shift can be as high as 20 or 24 ppm and the Doppler rate can reach 0.09 and 0.27 ppm/s. For comparison, in terrestrial communications, the Doppler shift is mainly related to the UE crystal accuracy which is 10 ppm.
- complex management of the handovers and paging depending from the on-ground coverage.

There are two options for the on-ground coverage, which are shown in Fig. 4:

- **Earth fixed cells** – The cells are fixed to a certain location on Earth from the time where the satellite (these cells belong to), is at a certain elevation angle over the horizon until the same satellite has reached the same elevation angle at the opposite horizon. At that point in time, another satellite takes over and all connected UEs are handed over to a new cell at the new satellite. Only UEs in cells at the satellite coverage border are subject to the handover. Each

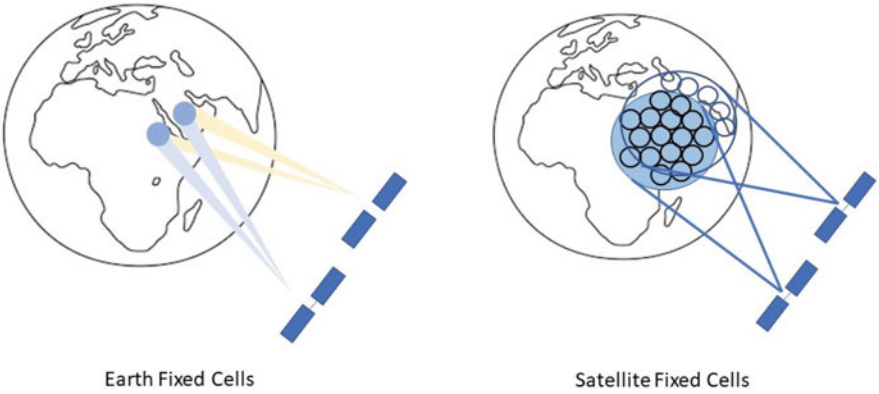


Fig. 4 Earth fixed cells vs Satellite fixed cells coverage

satellite must have the capability to steer beams towards fixed points on Earth and this can be realized through a mechanically steerable beam or a beamforming (BF) technique. A LEO satellite spot-beam can serve a given UE for about 400/500 seconds.
- **Satellite fixed cells (Earth moving cells)** – The cells follow the satellite projection on ground as the satellite antenna system continuously points at the center of the Earth, and gradually move with the speed of the satellite, e.g. 7 km/s. The satellite beams are constantly moving along with the satellite, and a satellite spot-beam is expected to provide coverage for a given UE for just a few seconds.

The choice between the two types of coverage depends from the type of payload and the trade-off between flexibility and complexity. Payloads can be classified in two main classes:

- Bent-pipe payload: they work with the signals in passband and demodulation is not performed on-board;
- Regenerative payload: they work at packet level, thus implementing the demodulation on-board.

The Earth-fixed cells coverage requires on-board beam-steering typically implemented electronically (not mechanically) through beam-forming network (BFN). Regenerative and more complex payloads are needed, which on the other hand provide high flexibility, reconfigurability and modularity for the satellite coverage and resources allocation. LEO constellations experience very different traffic distributions in their antenna Field Of Views (FOVs) as they orbit around the Earth. User distribution may range from uniform over all the FOV to very dense concentrations. In an uneven traffic scenario, only a limited number of the beams are potentially active. In this case, if satellite payload does not provide the needed flexibility in resource allocation, precious satellite resources would be wasted in the non-populated areas and they would not be sufficient to match the throughputs required over more densely populated areas.

The Satellite-fixed cells type of coverage does not require on-board antenna beam-steering capabilities and could be implemented with simple bent-pipe payloads. On the other hand, very frequent handovers (of the order of seconds) that have to be carefully managed, are introduced. Moreover, the Earth coverage is not optimized to the traffic distribution hence channel resources could be wasted. Currently, most of operative constellations are using bent-pipe payloads with non steerable antennas thus implementing this type of coverage. For instance, the current OneWeb constellation consists of bent-pipe satellites with 16 non-steerable, highly elliptical beams. However, it has become clear that the next generation should consist of regenerative payloads, which would be crucial to:

- optimize the use of the limited payload/system space resources
- enable the implementation of ISLs and also inter-layer links, which are crucial to make a more efficient and effective use of the satellite components within the NTN.

The provision of a high degree of flexibility to the payload architecture will translate in wide potential for the satellite component to be an important actor at integrated system level and to pave the way of a key-role in the future 6G integrated connectivity.

2.2.1 Evolution of Frequency Bands

As shown in Table 1, constellations optimized for IoT applications, and hence characterized by low data rates, work mainly at VHF/UHF or maximum L and S band. Downlinks on S-Band would be expected to be able to implement data rates from 100 kbps to 1 Mbps. Larger data rates require the use of higher frequency bands such as Ku-, K-, and Ka-band, which are the frequencies used by Oneweb and Starlink targeting broadband communications. Starlink also plans to use V-band and Q/V band are the next frontier for the feeder links of the so-called High Throughput Satellites, which are meant to provide data rate up to terabit/s. For frequencies higher than 10 GHz, rain is the phenomenon that is responsible for the highest contribution to the total attenuation. To estimate the additional loss due to the atmospheric precipitation, ITU recommendations can be used [29] which have been created using a database of measurements performed in Q/V bands. Figure 5 shows he Complementary Cumulative Distribution Function (CCDF) of atmospheric attenuation at 37 GHz for six sites in Europe. The additional attenuation can exceed 30 dB. At such high frequency bands, it is fundamental the use of

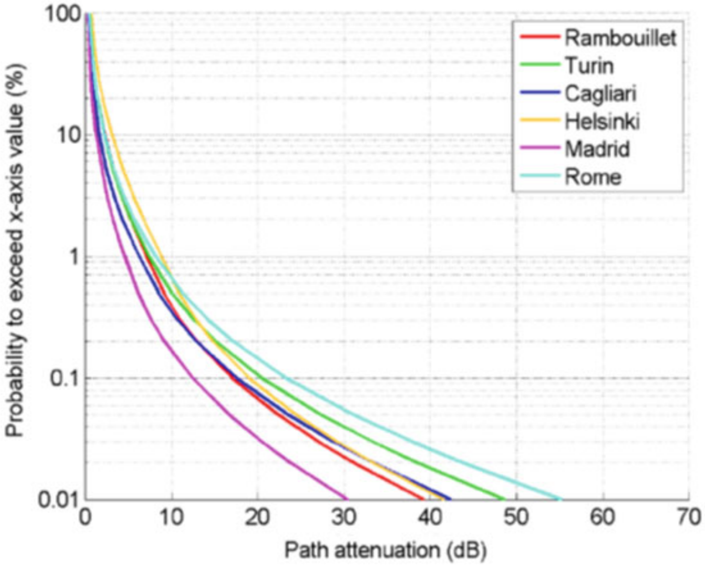


Fig. 5 CCDF of atmospheric attenuation at 37 GHz in six locations [32]

propagation impairment mitigation techniques (PIMT), such as Adaptive Coding and Modulation, Power Control and Smart Diversity. The optimization of the design of such techniques require a better understanding of the channel at such frequencies. The Alphasat satellite launched in 2013, carries “The Aldo Paraboni” payload to perform two experiments in Q/V band: the propagation experiment with the main objective to characterize the second order statistics of the channel; the communication experiment, whose main objective is the study and optimization of PIMTs and smart diversity strategies [30, 31].

About high data rates ISLs, which will be a key element of the future space networks, both optical or EHF bands are considered. Optical ISLs are well suited as they offer small size, weight and power advantage with respect to traditional RF links. On the other hand, optical ISLs are susceptible to pointing errors, which could be rather challenging and complex to implement over a small satellite in case of inter-plane ISLs. Several planned LEO constellations suggest the use of laser ISLs that may achieve data rates of 100 Gbps. RF EHF ISLs are characterized by wider beams that reduce the sensitivity to the pointing errors and enables neighbor discovery procedures. Q/V and also W-band are the most interesting options for RF ISLs. Recently a Ka-band ISL has been tested with the following characteristics. For a distance of 4700 km, a de-pointing of 2° and an antenna beamwidth of 5°, a data-rate higher than 1.5 Mbps in nominal mode has been achieved.

2.3 LEO Satellite Application Scenarios in 5G and beyond

Within the 5G framework, satellites play a key role in:

- fostering the roll out of 5G service in un-served areas that cannot be covered by terrestrial 5G network (isolated/remote areas, on board aircrafts or vessels) and underserved areas (e.g. sub-urban/rural areas). The performance of the terrestrial networks can be upgraded in a cost effective manner;
- reinforcing the 5G service reliability by providing service continuity for passengers on board moving platforms (e.g. passenger vehicles-aircraft, ships, high speed trains, bus) or ensuring service availability anywhere especially for critical communications, future railway/maritime/aeronautical communications;
- enabling 5G network scalability by providing efficient multicast/broadcast resources for data delivery towards the network edges or even user terminal.

Undoubtedly, LEO satellites will play a key role in the so-called Internet of Remote Things (IoRT) scenarios, where devices are distributed over wide and remote geographical areas where terrestrial networks are unavailable or out of reach, such as on remote land (think at the case of areas such as, e.g., forests) as well as offshore (e.g., in the oceans) [28], such as in the following application scenarios:

- Agriculture -The production of food must be made more efficient, resource saving and ecologically compatible. The possibility to collect data related to the health of soil, moisture analysis, water contamination level, water quantity level and other geodata as well as data on the equipment used on remote fields or on the animals freely moving in the area could be essential to make the agriculture processes more productive, efficient and environmentally friendly. Moreover, by processing such a data it could be possible to make predictions and better planning for the future.

- Long-term continuous environmental monitoring, such as the monitoring of air and water pollution, and wildlife position and activity;
- Detection of destructive phenomena such as landslides, avalanches, forest fires, volcanoes eruptions, floods, and Earthquakes;
- Critical Infrastructure Monitoring- Any accident, malfunctioning or man- made threat to a critical infrastructure such as railways, gas & oil pipelines and electrical grids poses risk to population safety, and finding a remedy is time-consuming and economically expensive. Continuous monitoring of the long pipelines of gas & Oil, the transmission lines of electricity, remote energy generators (winds), the track of railways require coverage in areas that might be non covered by the terrestrial infrastructure and the use of satellite to collect the data will be essential.

In all the mentioned scenarios, the sensors network must be characterized by: large number of nodes, very low cost, ease of deployment, low maintenance, and very long battery duration (possibly using solar energy). Nodes could be highly mobile (i.e., monitoring of wild animals). The collection of data from sensors can be done in two ways: (1) direct access and (2) indirect access. In the indirect access mode, each sensor may communicate with the satellite through a terrestrial sink node. Such sink node is equipped with a satellite terminal, expensive and power hungry while the other sensors nodes can be less expensive and power hungry. This solution allows to decrease the system costs and the complexity of the installation (in terms of antenna pointing and power generation facilities). On the other hand, the more challenging direct access to satellite allows much more flexibility.

However, the satellite component will be fundamental to reach the convergence, envisioned in 6G, between communication, navigation and sensing. The same infrastructure, both terrestrial or non terrestrial, should provide not only communications but also navigation and sensing services. In this framework, it must be mentioned the growing interest in using LEO megaconstellations designed to provide broadband communications, for PNT services.

2.3.1 LEO Constellations for PNT Services

The possibility to use LEO constellations for navigation has gained more and more interest. LEO provides several advantages with respect to MEO satellites. First of all, signals might be around 30 dB more powerful, and this allows to consider them also for indoor applications. Moreover, as LEO satellites move faster than MEO, localization based on Doppler measurements become attractive [32]. There are two main options to use LEO constellations for PNT services:

- dedicated LEO constellations which are optimized for PNT services [33];
- use of Signals of Opportunity (SoOP) transmitted by already deployed LEO constellations, such as OneWeb, Startlink, which are not originally meant for navigation purposes [27, 34, 35].

The first approach has the advantage of simple receiver architectures and navigation algorithms. However, they require the development of dedicate infrastructures and also spectrum allocation which might require a cost that private companies are not willing to pay. Moreover, the development of new space infrastructures would worse the problem of the space sustainability. The use of SoOP is very interesting for the operator that does not have to invest for a new expensive space infrastructure. Moreover, the possibility to use already deployed infrastructures for other services goes in the direction of a more sustainable use of the space but it poses some challenges [34]:

- satellites that are not meant for navigation do not usually transmit satellite ephemerides. Using the information that can be found in the two-line-element files which are tracked and publicly published on a daily basis by the North American Aerospace Defence Command (NORAD) (https://celestrak.org/NORAD/elements/, n.d.), introduces an error of kilometers due to several sources of perturbations.
- LEO satellites for communications services are not equipped with atomic clocks so they are not tightly synchronized.
- LEO satellites are owned and operated by private entities which use proprietary protocols and hence novel specialized receivers must be developed that are capable for extracting navigation observables.

Some recent papers have started to address the above-mentioned problems [36, 37], which represent a hot research area.

2.4 3GPP Standardization Activities on NTN

Figure 6 shows the 3GPP standardization activities from release 15 to release 17 on the integration of the satellite access into the terrestrial network [38, 39].

Initial activities started with Release 15. In Rel. 15, two SIs under RAN1 and SA2 where developed: (i) *Study on NR to support NTN* which have defined deploying scenarios and related system parameters, identified and assessed the potential key impact areas on NR and identified the required adaptations to the 3GPP channel models for NTNs, proposing also some preliminar solutions; (ii) *Study on using Satellite Access in 5G*, where a set of use cases for the integration of the

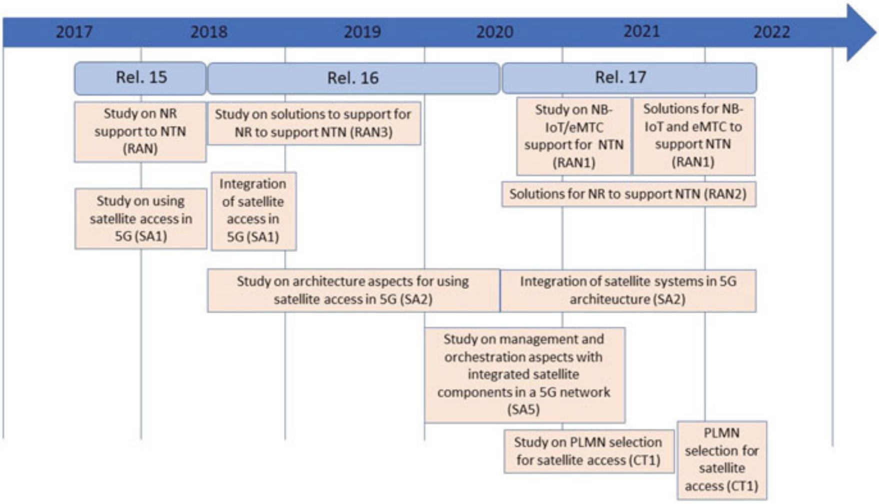


Fig. 6 Timeline 3GPP activities from Rel. 15 to Rel. 17

satellite component in NR have been identified, as well as potential services and categorization of use cases based on service continuity, ubiquity and scalability. Rel. 16 have developed: (i) a Work Item (WI) for NTN on *Integration of Satellite Access in 5G* supervised by SA1, which was then completed under Rel. 17; (ii) three additional SIs were initiated such as *Study on architecture aspects for using satellite access in 5G* under SA2, *Study on management and orchestration aspects with integrated satellite components in a 5G network* within SA5 (management) and *Study on solutions for NR to support NTN* within RAN3. In the last one, starting from the results on release 15, a set of required adaptations enabling NR technologies and operations in the NTN context were addressed covering several issues in RAN1 (PHY), RAN2 (layer 2 and 3) and RAN3 (interfaces). Moreover, the performance assessment of NR in scenarios including GEO and LEO satellites has been provided at both system and link level, together with a preliminary set of potential solutions for NR adaptations at layer 2 and 3. Finally, some architecture aspects were modified with respect to TR 38.811, which is superseded by TR 38.821. It is worth outlining that Rel. 15 and Rel. 16 have mainly focused on the GEO and LEO components of the NTN. Also Rel. 17 was mainly focused on that, but it started to consider the compatiblity to support also HAPs and air-to-ground scenarios. Moreover, Rel. 17 has started to consider also Internet of Things applications and the NB-IoT standard and have initiated the SI on *Study on NB-IoT / enhanced Machine Type Communication (eMTC) support for NTN* and the WI on *Solutions for NB-IoT and eMTC to support NTN.*

The specific characteristics of satellite links, and in particular of LEO satellite links, such as much larger propagation delays and also the variability of the propagation delay along with the much high Doppler shifts, raise the need to modify some of the procedure of 5G-NR standard and NB-IoT. On example of modification that is needed is the *Time Advance Adjustment*.

In 5G-NR, the gNodeB provides the UE with a common Time Advance (TA) that signals the RTT between the satellite and the gNB. The UE adds the RTT between the UE and the satellite to the TA to get the full TA which is used as an offset between the received downlink timing and uplink transmission timing at the UE. Therefore, if downlink slots n starts at time t_0, then the uplink slot n starts at time t_0-full TA. This is important for both random access and data transmissions in connected mode. When the communication occurs with LEO constellations satellites, two issues arise:

- TA needs to be dynamically updated since the propagation delay is highly variable (the differential delay is 4 ms in case of LEO at 600 km and 7 ms in case of LEO at 1500 km) due to the changing distance from UE over Satellite to BS;
- The delay, of the order of 28 ms in case of LEO at 600 km and 51 ms in case of 1500 km, is much longer over a satellite link than one TTI (equivalent to one frame), which is equal to or less than 1 ms.

The estimation of such TA can be done both at the UE side and at the network (satellite) side. However, in both cases the knowledge of the UE position would be needed.

Also the high Doppler shifts and its variation rates call for pre-post compensation techniques. Some solutions have been identified, which rely on the assumption that the UE is equipped with a GNSS receiver and knows its own position. The knowledge of the position would also help in improving the frequent handovers management and the paging when LEO constellations are considered. However, the assumption that the UE is equipped with a GNSS receiver could not be compatible with some application scenarios. For instance, in IoRT scenarios, terminal could have stringent power and dimension constraints and equipping them with a GNSS receiver could be not feasible. As a matter of fact, this assumption has been removed by Rel. 18. This opens a new research area, which is the possibility to geolocalize a UE at the satellite side, using the same satellites that are used for communications. Only few works have so far focused on this novel research area.

3GPP Rel-18 is continuing to work on a further list of enhancements for both NR-NTN and IoT-NTN. Plans are also underway to further define the enablers for NR based satellite access in bands above 10 GHz to serve fixed and moving platforms (e.g., aircraft, vessels, UAVs) as well as building- mounted devices (e.g., businesses and premises). The goal of these efforts is to further optimize satellite access performance, address new bands with their specific regulatory requirements, and support new capabilities and services as the evolution of 5G continues.

2.5 The Issue of Space Sustainability

There are currently more than 4800 active satellites around the Earth. Table 1 shows the number of satellites of some constellations. Generation 1 of SpaceX's Starlink will consist of 11,926 satellites, and generation 2 will have 30,000 more. We are witnesses of a new era of space activity that promises new opportunities for economic development, global education, rural healthcare, location-based services, and advancements in environmental science. However, the increasingly congested operating environment raises concerns over space sustainability and the safety risks. Together with active satellites, there are currently an estimated 330 million pieces of space debris, including 36,500 objects bigger than 10 cm, such as old satellites, spent rocket bodies and even tools dropped by astronauts orbiting around Earth. This crowded situation poses several challenges. First of all, the margin of error in space operations is narrower as separation between satellites is reduced. Therefore, there is a higher probability of collisions which will lead to an increase in the debris thus making the space less and less safe. Moreover, such crowded situation raises interference issues. On one hand, astrophysics are worried for the interference caused to astronomical observations [40]. On the other hand, the spectrum management will become very complex to mitigate the radio frequency interference to other communication systems. Urgent actions are needed to make the space more sustainable and safe.

The issue of a sustainable use of space and its protection has become very urgent. Several recent papers have provided an overview of the problem and outlined several technological issues related to the space sustainability but also regulatory issues [41, 42]. Most of the initiatives and research activities are focused on mitigating the damages caused by a wild use of satellite orbits and an uncontrolled generation of space garbage [43, 44]. The mitigation alone of the already made damages – without taking other actions - will not guarantee a long term space sustainability and safe use of space. In [45] some enablers for safe and sustainable future satellite communication systems have been identified, such as space traffic management, debris detection techniques, spectrum sharing and cyber-security aspects. However, a new design approach should be developed aiming to reduce the amount of objects and "mass" that are launched in the space [46]: future missions should be designed to use as much as possible the already developed space infrastructure (BW-compatibility) and should be open and flexible enough to be used for different missions in the future (FW-compatibility).

2.6 Concluding Remarks

Several megaconstellations of LEO satellites have been recently developed or are under development to provide internet access or IoT services. They provide connectivity also to rural and low income areas, where geographic accessibility and cost of laying fibre optic cables are the primary barriers to ubiquitous Internet access. One big barrier to face the digital divide and cover rural areas in low-income zones is that LEO constellation user terminals are very costly (Starlink currently charges around 500 dollars). The monthly rate for the subscription could be several dozens of dollars. This could limit market penetration. Providers are trying to make moves to target this issue and terminal cost of end-user terminals is estimated to fall to approximately 130–300 dollars over the next decade. However, the terminal costs could be really reduced and hence, they could really play a role in closing the digital divide, once an effective integration between terrestrial and satellite networks will be a reality. The first steps for integrating of satellite nodes with terrestrial nodes has been done in the 5G 3GPP standardization. 6G foresees a more seamless integration of terrestrial and non terrestrial nodes, where LEO satellites will be an important component, but part of a more complex network including also other types of non terrestrial nodes. To be able to take full advantage of such a complex network, very high level of flexibility both at node level (or payload level) and network level will be needed. It should be possible to establish opportunistic links between very heterogeneous type of nodes (LEO or GEO satellites, UAVs, HAPs) on a dynamic basis. Communication links between such nodes and ground terminals, but also between non terrestrial nodes will be crucial and characterized by variable delays, high Doppler shifts. Novel waveforms and higher and higher frequency bands are being investigated. To handle such complexity, AI and ML tools will play a key role. Already in the current 3GPP standards, it has become evident that the knowledge of absolute or relative position of the communicating nodes could greatly help the integration between terrestrial and satellites nodes. Such a knowledge would be important also for proper resource and interference management in coexistence and spectrum sharing scenarios with terrestrial and non terrestrial nodes. However, it is not straightforward to assume that user terminals are equipped with GNSS receivers, at least not in all possible scenarios. This issue opens the way for an interesting research area: the geolocalization via one single satellite, used for communication, of the user terminal.

Finally, it is worth outlining that 6G will benefit of the satellite components only if the robust and effective solutions to the issue of space sustainability will be put in place in the near future.

3 On the Role of Non-terrestrial Networks in 6G

Debashisha Mishra • Evgenii Vinogradov • Enrico Natalizio

3.1 Non-Terrestrial Radio Access Network and 6G

The integration of heterogeneous networks, such as ‘5G with satellites’ or ‘5G with UAVs’, is already a part of the standardization process for 5G by 3rd Generation Partnership Project (3GPP). Due to the enormous potential of UAVs, it is anticipated that they will soon become an essential technology enabler of the airspace. Global deployment of these UAVs is anticipated as the technology develops and the necessary rules are put in place. Their inherent mobility in three dimensional space and portability makes them useful in a wide variety of contexts, including package delivery, pollution control, farming, and search-and-rescue missions. 6G wireless systems envision to bring transformational changes to global ubiquitous coverage that establishes the connectivity of the future via universal and seamless accessibility for a variety of different Radio Access Technologies (also known as multi-RATs) [47].

There are continuing discussions to build and construct network architecture that combines cross-layer, high and low altitude platforms and space satellites into conventional cellular networks in order to inject more capacity and improve coverage for underserved regions in a cost-effective way [48]. Motivated by this necessity, this section details workable strategies for integrating space-air-ground networks and provides insight into studies that have embraced the multi-dimensional and inter-operational network, key to the 6G concept.

3.1.1 Integrated Space-Aerial-Ground Communication Network

This section describes the three different parts of the network and places special emphasis on the “aerial” segment as an enabling layer between the ground and the space, as seen in Fig. 7. The links that connect the ground (terrestrial) network with aerial ones make it possible to communicate ground-to-aerial (G2A) and aerial- to-ground (A2G), which complements the wireless broadband access provided by the ground infrastructure [50–52]. This kind of cross-layer inter-working is the most advanced approach in terms of the implementation of already existing solutions that have been standardized. 3GPP Rel-15 anticipates UAVs having the ability to handle high transmission rates (10 Gbps), tight latency (one millisecond round trip delay), user traffic offloading, and upgrades to radio access technologies [53]. In addition, 3GPP Rel-17 outlines the requirements and key performance indicators (KPIs) for a variety of use cases including dependable and beyond visual line of sight (BVLOS) operations using UAVs [54].

Satellites and other means of long-distance communication fall under the “space” category. Satellite communication backbone (SATCOM) is centered on earth expanding to different orbital depths of the universe. Typical SATCOM backbone network consists of low earth orbiting (LEO), middle earth orbiting (MEO) and geosynchronous earth orbiting (GEO) satellites delivering data to the “ground” network segment. SATCOM allows satellites in various orbits to provide varying ser-

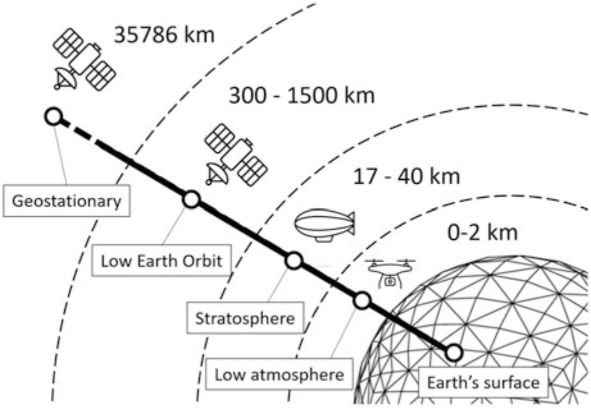


Fig. 7 Platforms of non-terrestrial networks: GEO, LEO, HAP, and UAV [49]

vice quality. SATCOM connections are reliable and secure, with the added benefit of being able to reach every corner of the globe. However, significant barriers to entry in this market include high costs associated with infrastructure, long propagation distances, and slow connections. Space-to-aerial (S2A) and aerial-to-space (A2S) transmission lines are an effective cross-layer inter-working method for transferring information from satellites to "aerial" and then to "ground" segment. LEO and GEO satellites, which operate at 2000 and 35,000 km altitudes, respectively, are the two most often utilized systems for S2A/A2S communications, providing cost-effective and backhaul-aware transmissions to HAPs [55].

The aerial segment is a key layer in implementing the multi-dimensional and inter-operational network of 6G vision because to its capacity to efficiently bind the other two levels, which are separated by thousands of kilometers. In this market category, commercial multirotors, fixed-wing UAVs, balloons, and airships serve as high or low altitude platforms (HAPs or LAPs). HAPs work with decreased transmission latency, cheaper cost, simple mobility in emergency scenarios, and broad coverage with high elevation angles, in contrast to SATCOM. Existing works demonstrate the cooperation between SATCOM and HAP for robust beamforming or boosting communication confidentiality. LAPs serve as a connecting connection to ground devices in cases when HAPs are not desired owing to their expensive cost in common civilian applications (e.g., temporary hotspots, sporting events). The LAPs could then bridge the link to HAPs and then to the space segment, enabling end-to-end communication from the ground device to the space network.

3.1.2 Unmanned Aerial Vehicles

Unmanned aerial vehicles (UAVs) are expected to play a defining role in the future of low-altitude airspace due to their widespread use in fields as diverse as civil and public safety, Industrial Internet of Things (IIoT), cyber-physical systems, atmospheric monitoring, environmental monitoring, and more [56]. Their ability to move in three dimensions (3D), be deployed in a variety of ways, and run autonomous missions with dynamic altitude control make them ideal candidates for providing on-demand, agile communication services to ground-based user devices (UEs). Due to its self-organizing maneuverability, it can establish a distance-efficient line-of- sight (LOS) link for ground UEs and among UAVs [57, 58]. Gradual innovations on aerial robotics and its widespread demand on several interesting civilian use cases, the cost to manufacture commercial drone has become much more affordable. In this viewpoint, it serves as an on-demand, temporary, cost-effective platform to supplement to the network load of terrestrial BS infrastructure, especially in contexts with high spatio-temporal dynamics of communication demand, e.g.: areas where there is a sudden influx of people, such as at a flash crowd, a hotspot, disaster [59], or a temporary event. UAV-assisted cellular radio access network (RAN) or flying RAN refers to airborne UAVs operating as flying base station or relay infrastructures within the scope of 5G and 6G to deliver flexible, on-demand communication services to UEs on the ground [60].

3.1.2.1 Prospects of UAVs in 6G

The following points elaborate on certain elements of planned 6G communication for airborne RAN.

- **Aerial Internet:** UAVs deployed as aerial base station (UAV-BS) can extend Internet services to remote/rural areas and under-served regions of interest with poor or no connectivity.
- **Emergency and Temporary Network:** UAVs offer high capacity links for wireless coverage required for temporary events e.g., political rallies, sports event, exhibitions, hotspots, etc. Additionally, for cell-edge or resource constrained user, it provides extra capacity and fair communication services by optimally placing itself in a better serving position.
- **Aerial Backhaul:** UAVs can serve the backhaul demands arising from ground infrastructure and also extend to locations without any wired backhaul solutions. It decreases the cost of traditional fiber-like deployments.
- **Sensing:** UAVs can be used to harvest data from ground sensor devices. While such data collection is dependent on efficient placement before the data collection begins, an online optimization on UAV placement that copes with the dynamic environment by learning or discovering the uncertain environment during flight is much more promising. An optimal trajectory path that maximizes the metric associated with data collection accelerate the sensing performance.
- **Caching, Computing, Control (3C):** In latency-critical applications, UAVs with edge computing functionality can exploit closeness to the user and provide rapid deployment solution for content distribution applications. The practicality of deploying edge servers on satellites is limited due to cost, latency and hardware calibration. However, UAVs perform as the best candidate in such requirements,

thus bringing content and computational resources close to the origin of requests (users).

3.1.2.2 Air-to-Ground (A2G) Channel

In wireless communication networks, the propagation channel is the medium between the transmitter and the receiver. As the medium properties largely define the physical limitations of wireless networks' performance (e.g., range, achievable throughput, latency) and directly impact the technology design choices, channel characterization and modeling become a crucial first step toward achieving the ambitious 6G performance goals.

The nature of A2G channels implies a high probability of line-of-sight (LOS) propagation (this is especially important for higher frequencies such as mmWaves and THz). This results in higher link reliability and lower transmission power required to ensure the targeted link budget [61]. Even for non-line-of-sight (NLOS) links at lower frequencies, power variations are less severe than in terrestrial communication networks due to the fact that only the ground-based side of the link is surrounded by the objects that affect the propagation [62]. Figure 8 illustrates A2G propagation channel and introduces the main geometrical parameters as well as drawing the important distinction between LOS and NLOS channels.

Propagation Basics The transmitter radiates electromagnetic waves in several directions. Waves interact with the surrounding environment through various propagation phenomena before they reach the receiver. As illustrated in Fig. 9, different phenomena such as specular reflections, diffraction, scattering, penetration or any combination of these can be involved in propagation [63]. Moreover, the blockage must be considered (see Fig. 8). Therefore, multiple realizations of the transmitted signal, often termed as Multipath Components (MPC) arrive at the Rx with different amplitudes, delays and directions. The resulting signal is the linear coherent superposition of all copies of the transmitted signal, which can be constructive or destructive depending upon their respective random phases.

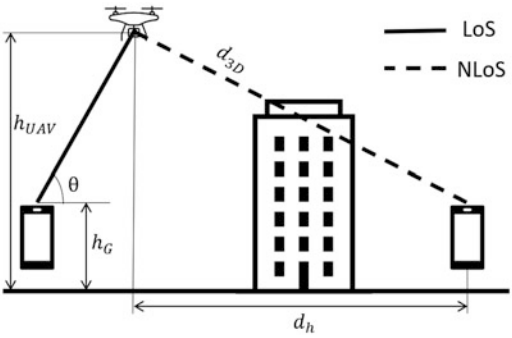


Fig. 8 Geometry of Air-to-Ground propagation

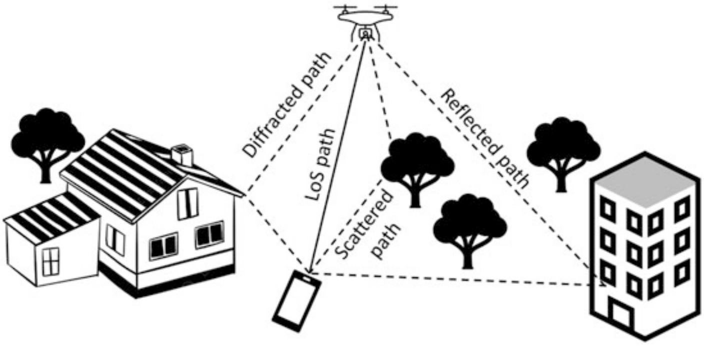


Fig. 9 Air-to-ground propagation phenomena

Typically, radio channels can be represented as a superposition of several separate fading mechanisms:

$$\mathrm{H} = \Lambda + \mathrm{X}_{\mathrm{sh}} + \mathrm{X}_{\mathrm{SS}}, \tag{1}$$

where, Λ is the distance-dependent Pathloss (PL), X_{sh} is the Shadow fading (also known as shadowing) consisting of large-scale power variations caused by the environment, and X_{SS} is the Small-Scale fading (alternatively, fast fading). Next, let us describe models of the components presented in (1) separately.

There exist several channel models not drawing an explicit distinction between LOS and NLOS channels. However, the most common channel modeling approach consists of the four following steps:

- Define the link state (LOS/NLOS);
- Generate pathloss accordingly;
- Generate Shadow fading;
- Generate SS fading.

Line-of-Sight Modeling In the case when the distinction between LOS and NLOS links is made, the LOS probability P_{LOS} modeling becomes critical. In this chapter, we describe a popular approach suggested by ITU[1] [64]. In [64], the LOS probability is given by

$$P_{LOS} = \prod_{n=0}^{m} \left[1 - \exp\left(- \frac{\left[h_{UAV} - \frac{(n+1/2)(h_{UAV} - h_G)}{m+1} \right]^2}{2\Omega^2} \right) \right], \tag{2}$$

---

[1] Note that another popular model suggested by 3GPP in [53] is not applicable since it considers an aerial user connected to an elevated terrestrial BS.

where we m = floor($d_h \sqrt{\varsigma\xi} - 1$), d_h have is the horizontal distance between the UAV and the ground node, h_{UAV} and h_G are the terminal heights, ς is the ratio of land area covered by buildings compared to the total land area, ς is the mean number of buildings per km^2 and Ω is the scale parameter of building heights distribution (assumed to follow Rayleigh distribution). In some cases, it is more convenient to express the LOS probability as a function of incident or elevation angle (e.g., in [62]). This representation can be found in [65].

It is worth noting that, the NLOS probability is computed from the LOS probability by following equation.

$$\mathrm{P_{NLOS}} = 1\text{–}\mathrm{P_{LOS}} \tag{3}$$

Pathloss Modeling Pathloss is the distance-dependent attenuation experienced by radio signal while it propagates through the media. The simplest PL model assumes an LOS link between the transmitter and receiver, and propagation in free space. In this case, the received signal power is given as [66]

$$P_R = P_T G_T G_R \left(\frac{\lambda}{4\pi d}\right)^{\eta}, \tag{4}$$

where P_T is the transmitted power, G_T and G_R are the transmit and receive antenna gains, respectively, λ is the carrier wavelength, and d is the distance between the Tx and Rx[2]. Note that the Pathloss exponent (PLE) η (the power of the distance dependence) in this equation is 2 for free-space propagation. So that the path loss can be expressed for a generalized case as

$$\Lambda = \left(\frac{4\pi d}{\lambda}\right)^{\eta}. \tag{5}$$

Unfortunately signals in real-life A2G wireless communications do not always experience free space propagation. In the majority of literature, the well-known log-distance PL model with free-space propagation reference is used for PL (in dB) modeling:

$$\Lambda\left(d\right) = \Lambda_0 + 10\eta \, log\left(d/d_0\right), \tag{6}$$

where Λ_0 is the PL at reference distance d_0 (Λ_0 can be specified or calculated as free space pathloss $\Lambda_0 = 20\log\left[\frac{4\pi d_0}{\lambda}\right]$)

Finally, pathloss can be predicted based on a combination of LOS and NLOS components [67–69]:

---

[2] For simplicity of notation, $d = d_{3d}$ in Fig. 8.

Table 2 Parameters of Pathloss and Shadow fading models

Scenario	Frequency, GHz	η	Λ_0, dB	σ, dB	References
		2.54–3.037	21.9–34.9	2.79–5.3	[70]
		2.2–2.6			[71]
		2.01			[72]
Suburban, urban, open field	0.968	4.1		5.24	[73]
		2–2.25			[74]
		1.6	102.3		[75]
	5.06	1.9	113.9		[75]
	0.968	1.7	98.2–99.4	2.6–3.1	[76]
	5.06	1.5–2	110.4–116.7	2.9–3.2	[76]
Urban (LOS)	28	2.1		3.6	
Urban (NLOS)	28	3.4		9.7	
Urban (LOS)	38	1.9–2		1.8–4.4	[77]
urban (NLOS)	38	2.2–2.8		4.1–10.8	
Urban (LOS)	73	2		4.2–5.2	
Urban (NLOS)	73	3.3–3.5		7.6–7.9	

$$\Lambda = \mathrm{P}_{LOS} \cdot \Lambda_{LOS} + (1 - \mathrm{P}_{LOS}) \cdot \Lambda_{N\,LOS}, \tag{7}$$

where $\Lambda_{LOS,NLOS}$ are the path loss for the LOS and NLOS cases, respectively, P_{LOS} denotes the probability of having an LOS link between the UAV and the ground node.

The PLE values for Λ_{LOS} and $\Lambda_{N\,LOS}$ can be found in Table 2 [62]. Note that these results are valid for the indicated frequency ranges and environments. Additionally, atmospheric absorption and rain attenuation can also lead to significant power loss for mmWaves and THz frequency bands.

Shadowing and Small-Scale Fading Modeling Apart from pathloss, large obstacles (such as buildings, vegetation, and vehicles) can cause position-specific random variations of received power. Typically, these variations are changing relatively slowly (in the order of a few tens or hundreds of wavelengths). This channel component is called Shadow fading. At any distance d, Shadow fading X_{Sh} measured in dB is usually modeled as a normal random variable with a variance σ, which takes into account random variations of the received power around the pathloss curve. In Table 2, the standard deviation for the Normal distribution N(0, σ^2) describing Shadow fading can be found.

Small-scale fading describes the random fluctuations of the received power over short distances, typically a few wavelengths, due to constructive or destructive interference of MPCs impinging at the receiver. Different distributions are proposed to characterize the random fading behavior of the signal envelope, suitable for different wireless systems and propagation environments. The Rayleigh and Rice distributions, both based on a complex Gaussian distribution, are the most commonly used models. Considering a large number of MPCs with amplitudes and random phases, the signal envelope of small-scale fading thus follows a Rayleigh distribution [66]. For A2A and A2G channels, where the impact of LOS propagation is high, the Ricean distribution [66] provides a better fit. Of course, other distributions such as the Nakagami [78], chi-squared (χ^2) and non-central χ^2 [68, 79], and Weibull distributions might also be employed. The family of χ^2 distributions is attracting our attention since many of the distributions listed above are particular cases of it.

Small-scale fading models apply to narrow-band channels or taps in tapped delay line wideband models. Due to the stochastic nature of these signal variations, fading is usually modeled using statistical approaches and its models are obtained through measurements or through geometric analysis and simulations. The most popular type of small-scale models is Geometry-Based Stochastic Channel Models (GBSCM) [63].

3.1.2.3 Research Challenges

While there are several research issues associated with UAV-assisted cellular communication, we aim to highlight some notable challenges that must be considered before successful roll-out of system encompassing UAVs as part of aerial network infrastructure [80].

Optimal Positioning and User Mobility It is clear that proximity to the user is helpful for enhancing the quality of service (QoS) provided to the user due to the greater SINR and LOS likelihood [81]. UAVs take use of their great relocating flexibility to position themselves in appropriate areas in order to provide the user with a higher QoS. However, making optimal placement selections in airborne networks is not simple. First, both the user and the UAVs are mobile during the operation, resulting in a time-varying channel. In other words, a little user or UAV or both movement invalidates a superior serving position. Second, users with varying QoS requirements create trade-offs in which the best placement option cannot be reached with fine temporal granularity. In an ideal circumstance, for instance, a better serving position would be determined in each transmission time interval (TTI) when user scheduling occurs, however this is extremely difficult from a realistic viewpoint. Third, the precise deployment of unmanned aerial vehicles (UAVs) must be based on accurate radio maps, which may not be available in an unpredictable environment. This will result in a considerable amount of time spent constructing the radio map based on the various user locations.

Cooperative Control Due to the rising complexity of missions and the limiting capabilities of a single UAV, a multi-UAV system (UAV fleet or UAV swarm) is becoming the obvious choice for accelerating mission completion. Multi-UAV deployment with cooperative route planning and control is essential for cooperatively optimizing user services and preventing UAV collisions [82]. This could take one of two forms: (a) centralized control, in which a centralized entity is responsible for pre-computing the time-dependent flight paths for UAVs and assigning them to each UAV on mission; or (b) decentralized control, in which each UAV runs its own movement control and reacts autonomously to the behavior of other UAVs in the fleet. A decentralized category offers greater benefits in autonomous UAV behavior as a result of a collaborative planning and optimization process that reaches an agreement among them.

Wireless Backhaul In terrestrial infrastructure, the backhaul is frequently a high-speed, multi-gigabit, dedicated fiber link distributed from the BS site to the core network. When UAVs are working as the aerial base stations, this link is wireless and suffer from numerous constraints in terms of gathering user data and deliver them to the core network. Recently, Millimeter wave has been proposed as an enabler to answer the speed need in the backhaul, although such an implementation increase the complexity owing to high atmospheric absorption loss, high frequency range of operation, and need of directional antennas for transmission [83].

Energy Constraints and Recharging In every case where a UAV ecosystem is used, the amount of energy available is one of the most important factors to think about in order to provide continuous, high-quality service [80]. UAVs are energy-constrained machines that runs largely on battery or solar power. There has been a flurry of studies lately looking into the feasibility of gasoline fuel that might last for days at a time with no interruptions. For the most part, this will rely on the mass and dimensions of the UAVs involved in the communication operation. UAVs have a total energy consumption that takes into account not just the mission's demands but also the UAV's own weight and any extra payload it carries during landing and takeoff. The restricted onboard energy still substantially inhibits the practical implementation of longer mission, despite the availability of energy harvesting technology and effective storage units. Iterative recharging and mission integration is considered by the authors in [84] and [85] for efficient UAV operation over extended periods of time.

Interference Management When more than one aerial BS is active, it causes interference not just between itself but also with the ground-based stations [86]. Additionally, non-serving UAV base stations may be negatively impacted by uplink interference originating from terrestrial user transmissions. As a result, there has to be appropriate interference mitigation mechanisms in place to reduce the possibility of interference between UAVs and terrestrial BS.

3.1.3 High Altitude Platforms (HAPs)

While UAVs are a great tool for network densification [87], macro BSs remain a crucial component in wireless access architectures to provide coverage and support capacity. Although network coverage and capacity can be improved through the addition of UAV-mounted BSs, their Size, Weight, and Power (SWAP) constraints limit the lifetime and coverage area of UAV BSs. Also, the mobility of UAV BSs introduces challenges related to frequent BS activation/deactivation.

3.1.3.1 Prospects of HAPs in 6G

Compared to UAVs, HAP-based systems have more computational power, a larger footprint [88], and better LOS communication links [64]. A HAP-mounted SMBS (HAP-SMBS) can be a powerful platform to enhance connectivity. However, HAP-SMBS are rather a complementary solution for network management and control than alternatives to terrestrial BSs [89]. Moreover, in metropolitan areas, the use of multiple coordinated HAP-SMBS systems, equipped with multi-antenna arrays, can enable further flexibility of the extremely precise beams through a distributed MIMO set-up. Further, HAP-SMBS systems can serve as computational platforms. They can function as intelligent infrastructure to enable communication, computation, and caching. We envision that future HAP-SMBS architectures will support data acquisition, computing, caching, and processing in diverse application domains, as exemplified in Fig. 7, and detailed below. These potential use cases have been recently presented in [89, 90].

- **Massive IoT connectivity**: The ever-increasing proliferation of IoT technologies presents substantial challenges regarding connectivity, reliability, and latency requirements of a massive number of connected devices. The wide footprint of HAPs is ideal for providing greater coverage to a high number of IoT devices. Moreover, IoT devices might be located in areas where there is no or limited terrestrial network coverage such as forests, mountains, oceans, etc.
- **Services for Emergency and Spontaneous high-demand events**: To ensure service during unexpected and temporary events such as flash crowds, wireless networks might require additional support. As we indicated above, UAV-mounted BSs can offer a solution. However, compared to UAV-mounted aerial BSs, HAP-SMBS can provide greater capacity for ground users due to their large platforms, massive-MIMO capabilities, and higher transmission power. However, UAVs offer a more agile solution.
- **Backhaul:** Installing fiber for BS backhauling may not be an efficient solution for many environments due to its high cost. A cost-effective backhauling solution is the use of wireless links. Due to relaxed SWAP limitations of HAPs, backhauling can leverage wider channel bandwidths of mmWave bands and MIMO digital beamforming with high gain. Moreover, THz frequencies or Free Space Optical links can be utilized.
- **Aerial Data Centers:** HAP-SMBS systems can operate as aerial data centers to support agile computational offloading. One of the important design issues in data centers is reducing response delays. Analyzing data in the sky will reduce response delays and decrease the burden on communication links. As the atmospheric temperature at these altitudes is quite low (on average in the range of [-15 °C,- 50 °C], the energy consumption of the cooling subsystem can be

reduced. However, the size of the HAP-SMBS data center will be limited by the payload capacity.

- **Handover support for the NTN subsystems:** The high speeds of LEO satellites necessitate frequent handovers at terrestrial gateways. On the other hand, mobile UAV-mounted base stations need to handover between different wireless backhaul links. Fortunately, HAPs-SMBS systems can cover many satellites and UAV-BSs simultaneously due to their wide upper footprint.
- **Global connectivity:** the satellite stations combined with HAPs nodes can play a significant role in global connectivity when terrestrial BSs are overloaded. Moreover, users with high throughput located outside terrestrial BS coverage areas (e.g., suburban and remote areas) can be served via this infrastructure [91].

3.1.3.2 Channel Modeling for HAPs

Channel modeling for HAPs links is similar to the one described above for UAVs: (i) Define whether the link is LOS/NLOS; (ii) generate pathloss, shadowing, and small-scale fading values accordingly. Consequently, in the following, we only highlight some important differences.

ITU identified a bandwidth of 0.324–2.969 GHz as necessary for the HAPs-to-ground and ground-to-HAPs links. Additionally, at the World Radiocommunication Conference in 2019 (WRC-19), it was agreed to append the 31–31.3 GHz, 38–39.5 GHz bands for HAPs usage, in addition to the already dedicated 47.2–47.5 GHz and 47.948.2 GHz bands for worldwide usage. Moreover, the usage of higher frequencies is currently under discussion.

As higher frequencies and flight altitudes are considered for HAPs, propagation becomes more affected by several phenomena the influence of which is less pronounced in the case of communication with UAVs. The list of these phenomena: (i) Rain Attenuation, (ii) Gaseous Absorption, (iii) Scintillation. Additionally, depolarization can occur [92]. For very large horizontal distances (*i.e.,* tens of kilometers), the Earth's curvature should be taken into account.

Rain Attenuation The troposphere consists of a mixture of particles having a wide range of sizes and characteristics, from the molecules in atmospheric gases to raindrops and hail. While an electromagnetic wave is passing through such a medium comprising many small particles, its total loss is composed of two additive contributions (when expressed in decibels) arising from absorption and scattering processes [93].

Note that the scattering process is strongly frequency-dependent since signals having a higher frequency (e.g., mmWave and THz) are scattered more as the wavelength becomes comparable with the particle size. The main scattering particles of interest to HAP systems are hydrometeors, including raindrops, fog and clouds. In these cases, the scattering component of attenuation is only significant to systems operating above around 10 GHz. The absorption component also rises with frequency, however, not so rapidly.

The total rain attenuation can be calculated as

$$\Lambda_{r\ ain} = d_{r\ a} R^b, \tag{8}$$

where R is the rainfall rate (in millimetres per hour), d_R is the distance traveled through the rain (remember that HAPs altitude is high), a and b are empirical parameters dependent on frequency and average rain temperature [94]. For example, at 20 °C, horizontal polarisation, and a central frequency of 40 GHz, $a = 0.35$ and $b = 0.939$ (for more details see [94]).

Gaseous Absorption Gaseous molecules found in the atmosphere may absorb energy from radio waves passing through them, thereby causing attenuation. In normal atmospheric conditions, only oxygen and water molecules contribute significantly to absorption, although other atmospheric gases may be significant in very dry air at above 70 GHz. The main resonance peaks of oxygen is at 60 (10–20 dB/km), whereas those of water vapour are 183.3 and 323.8 GHz (80 and 100 dB/km, respectively). For more details refer to [95].

Scintillation Scintillation is the result of radio wave propagation through a medium with small-scale random variations for the refraction index (in the troposphere, the variations are caused by air turbulence). When the wind blows, the horizontal layers of equal refractive index in the troposphere tend to become mixed due to turbulence, leading to rapid refractive index variations over small distances and over short time intervals. This effect generates constructive and destructive interference at the receiver (similar to multipath). Note that scintillation results in time-varying received signal strength even in cloudless sky conditions (so-called dry tropospheric scintillation). This effect is more pronounced in warm, humid climates and is greatest during summer days.

Scintillation does not cause absorption: the mean level of the signal is essentially unchanged. The effect is strongly frequency-dependent, its magnitude is measured by its standard deviation σ_S (in decibels), measured usually over 1 min intervals. The distribution of fluctuations (in decibels) is approximately Gaussian [96], with $\sigma_S < 1$ dB. However, the problem of scintillation does not lie in its amplitude but in the rate at which it occurs. The oscillations are faster at high frequencies, low elevation angles and low antenna gains. To reduce the effect of scintillation, MIMO techniques can be applied since it becomes possible to average the scintillation effect across the slightly different paths taken to each antenna element.

3.1.3.3 Relevant Tools for HAPs

Stochastic Geometry To study the coverage and capacity of a HAP or a constellation of HAPs, we recommend tools from stochastic geometry for modeling the spatial locations of UEs. These tools have been widely adopted for investigating various aspects of terrestrial networks as well as UAV-enabled systems. In general, these tools allow exploring the average behavior of the network to anticipate easy-

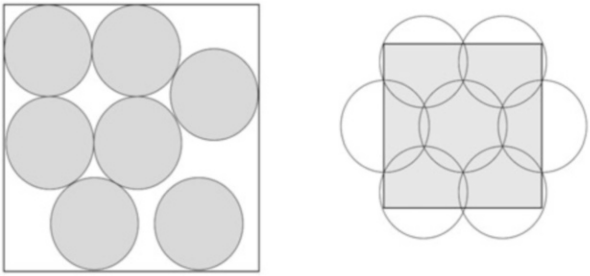


Fig. 10 Optimization problems: packing circles inside a square (left) and covering a square with circles (right)

to-use performance bounds of the network, which are important for understanding the large-scale impact of various system parameters. For example, the performance of large-scale HAPs-enabled networks can be assessed by modeling the location of HAPs via sophisticated point processes, such as the Determinantal Point process [97] and Ginibre Point Process [98], as these mathematical models allow for the inclusion of the (deliberate) repulsion that exists between the stations.

An accurate account of the inter-cell interference between HAPs cells and terrestrial cells can be included in the analysis. In systems based on HAPs constellations, the effect of inter-cell interference can be severe due to highly dominated LOS air-to-ground/ground-to-air channels. Signal radiated by the side lobes of the HAPs' antennas can still cause severe interference for ground users many kilometers away. This means that more advanced resource allocation along with sophisticated antenna techniques should be adopted at the stations. On the other hand, each platform may be a collection of several macro BSs. This means that the typical assumptions regarding the independency of large-scale path-loss attenuation (including the LOS probability) and shadowing need to be revisited. This is a novel challenge in the performance evaluation of HAPs-based systems.

Finally, we expect that stochastic geometry will play a key role in understanding the performance of a HAPs system for robotic and connected autonomous vehicle applications. In this case, Poisson Line Process can become a valuable tool.

Optimization Theory: Circle Packing and Coverage Problems Many complex optimization problems will have to be solved while HAPs-based wireless systems are designed. For example, HAPs locations have to be optimized to achieve the maximum network performance (e.g., in terms of coverage) or the antenna beam- forming codebook has to be optimized in order to cover as many users as possible. While optimizing the HAPs deployment location, traditional non-convex and convex optimization methods can be used. For example, the class of Packing and Coverage problems. In a packing problem, the main objective is to fill a single container as densely as possible (See Fig. 10, left). A dual problem is called a coverage problem in which we minimize the number of overlapping objects necessary to completely cover the container (See Fig. 10, right). Note that the circles can be of equal or unequal sizes.

The packing density corresponds to the proportion of the given surface covered by circles. In 2D space, the maximum circle packing is $\frac{\pi\sqrt{3}}{6}$ which results in a hexagonal packing structure. These two problems can be generalized to the Sphere packing/coverage problems. Usually, spheres of equal sizes are considered. These problems can become useful if HAPs are used to ensure coverage not only on the ground but also for (i) conventional aviation and (ii) Urban or Advanced Aerial Mobility (e.g., aerial taxis; cargo, freight, and delivery drones).

References

1. Di Renzo M et al (2019) Smart radio environments empowered by reconfigurable AI meta-surfaces: an idea whose time has come. EURASIP J Wirel Commun Netw 129:1–20
2. Massa A, Benoni A, Da Rù P, Goudos SK, Li B, Oliveri G, Polo A, Rocca P, Salucci M (2021) Designing smart electromagnetic environments for next-generation wireless communications. Telecom 2(2):213–221
3. Yang F, Erricolo D, Massa A (2022) Guest editorial smart electromagnetic environment. IEEE Trans Antennas Propag 70(10):8687–8690
4. Jia X et al (2015) Spectrum and energy efficiencies for multiuser pairs massive MIMO systems with full-duplex amplify-and-forward relay. IEEE Access 3:1907–1918
5. Oliveri G, Rocca P, Salucci M, Massa A (2021) Holographic smart EM skins for advanced beam power shaping in next generation wireless environments. IEEE J Multiscale Multiphys Comput Tech. 6:171–182
6. Oliveri G, Zardi F, Rocca P, Salucci M, Massa A (2022) Building a smart EM environment - AI-enhanced aperiodic micro-scale design of passive EM skins. IEEE Trans Antennas Propag 70(10):8757–8770
7. Chiaraviglio L, Di Paolo C, Blefari-Melazzi N (2022) 5G network planning under service and EMF constraints: formulation and solutions. IEEE Trans Mob Comput 21(9):3053–3070
8. Oliveri G, Rocca P, Salucci M, Erricolo D, Massa A (2022) Multi-scale single-bit RP-EMS synthesis for advanced propagation manipulation through system-by-design. IEEE Trans Antennas Propag 70(10):8809–8824
9. Dai L et al (2020) Reconfigurable intelligent surface-based wireless communications: antenna design, prototyping, and experimental results. IEEE Access 8:45913–45923
10. He Y, Cai Y, Mao H, Yu G (2022) RIS-assisted communication radar coexistence: joint beamforming design and analysis. IEEE J Sel Areas Commun 40(7):2131–2145
11. Luo H, Liu R, Li M, Liu Y, Liu Q (2022) Joint beamforming design for RIS-assisted integrated sensing and communication systems. IEEE Trans Veh Technol 71(12):13393–13397
12. Salem AA, Ismail MH, Ibrahim AS (2023) Active reconfigurable intelligent surface-assisted MISO integrated sensing and communication systems for secure operation. IEEE Trans Veh Technol., in press. https://doi.org/10.1109/TVT.2022.3227319
13. Amin Momeni Hasan Abadi SM, Booske JH, Behdad N (2016) Exploiting mechanical flexure as a means of tuning the responses of large-scale periodic structures. IEEE Trans Antennas Propag 64(3):933–943
14. Duan J, Liu R, Xu H, Zhang R, Zhang B, Kim B-S (2020) EMT conversion of composite broadband absorbent metamaterials for stealth application over X-bands. IEEE Access 8:153787–153798
15. Cui TJ, Smith D, Liu R (eds) (2010) Metamaterials: theory, design, and applications. Springer, New York
16. Engheta N, Ziolkowski R (eds) (2006) Metamaterials: physics and engineering explorations. Wiley, New York

17. Yang F, Rahmat-Samii Y (2019) Surface electromagnetics with applications in antenna, microwave, and optical engineering. Cambridge University Press, Cambridge, UK
18. Liang JC, Cheng Q, Gao Y, Xiao C, Gao S, Zhang L, Jin S, Cui TJ (2022) An angle-insensitive 3-bit reconfigurable intelligent surface. IEEE Trans Antennas Propag 70(10):8798–8808
19. Yang H, Yang F, Xu S, Li M, Cao X, Gao J (2017) A 1-bit multipolarization Reflectarray element for reconfigurable large-aperture antennas. IEEE Antennas Wirel Propag Lett 16:581–584
20. Pendry JB, Schurig D, Smith DR (2006) Controlling electromagnetic fields. Science 312:1780–1782
21. Rocca P, Da Rù P, Anselmi N, Salucci M, Oliveri G, Erricolo D, Massa A (2022) On the design of modular reflecting EM skins for enhanced urban wireless coverage. IEEE Trans Antennas Propag 70(10):8771–8784
22. Benoni A, Salucci M, Oliveri G, Rocca P, Li B, Massa A (2022) Planning of EM skins for improved quality-of-service in urban areas. IEEE Trans Antennas Propag 70(10):8849–8862
23. Oliveri G, Zardi F, Rocca P, Salucci M, Massa A (2023) Constrained-design of passive static EM skins. IEEE Trans Antennas Propag., in press. https://doi.org/10.1109/TAP.2022.3225593
24. Oliveri G, Salucci M, Massa A (2022) Towards efficient reflectarray digital twins - an EM-driven machine learning perspective. IEEE Trans Antennas Propag 70(7):5078–5093
25. Salucci M, Tenuti L, Oliveri G, Massa A (2018) Efficient prediction of the EM response of reflectarray antenna elements by an advanced statistical learning method. IEEE Trans Antennas Propag 66(8):3995–4007
26. Massa A, Salucci M (2022) On the design of complex EM devices and systems through the system-by-design paradigm - a framework for dealing with the computational complexity. IEEE Trans Antennas Propag 70(2):1328–1343
27. Morales-Ferre R, Lohan ES, Falco G, Falletti E (2020) Gdop-based analysis of suitability of LEO constellations for future satellite-based positioning. In: 2020 IEEE international conference on wireless for space and extreme environments (WiSEE), pp 147–152. https://doi.org/10.1109/WiSEE44079.2020.9262624
28. De Sanctis M, Cianca E, Araniti G, Bisio I, Prasad R (2016) Satellite communications supporting internet of remote things. IEEE Internet Things J 3(1):113–123. https://doi.org/10.1109/JIOT.2015.2487046
29. Propagation data and prediction methods required for the design of earth-space telecommunication systems. In: ITU-R P.618, 2017
30. Rossi T, De Sanctis M, Domenico SD, Ruggieri M, Cianca E (2021) Low-complexity blind equalization in q/v band satellite links: an experimental assessment. IEEE Trans Aerosp Electron Syst 57(6):4465–4471. https://doi.org/10.1109/TAES.2021.3111781
31. Aurizzi MM, Rossi T, Raso E, Funari L, Cianca E (2021) An sdn-based traffic handover control procedure and sgd management logic for ehf satellite networks. Comput Netw 196:108260. https://doi.org/10.1016/j.comnet.2021.108260
32. Benzerrouk H, Nguyen Q, Xiaoxing F, Amrhar A, Rasaee H, Landry RJ (2019) LEO satellites based doppler positioning using distributed nonlinear estimation. IFAC-PapersOnLine 52(12):496–501, 21st IFAC Symposium on Automatic Control in Aerospace ACA 2019. https://doi.org/10.1016/j.ifacol.2019.11.292
33. Ma F, Zhang X, Li X, Cheng J, Guo F, Hu J, Pan L (2020) Hybrid constellation design using a genetic algorithm for a LEO-based navigation augmentation system. GPS Solutions 24. https://doi.org/10.1007/s10291-020-00977-0
34. Z. M. Kassas, J. Morales, J. Khalife, New-age satellite-based navigation – stan: simultaneous tracking and navigation with LEO satellite signals, 2019
35. Reid TG, Neish AM, Walter T, Enge PK (2018) Broadband LEO constellations for navigation. Navigation 65(2):205–220. https://doi.org/10.1002/navi.234
36. Khairallah N, Kassas Z (2021) Ephemeris closed-loop tracking of LEO satellites with pseudorange and doppler measurements. In: Proceedings of the 34th international technical meeting of the satellite division of the Institute of Navigation (ION GNSS+ 2021)

37. Neinavaie M, Khalife J, Kassas ZM (2022) Acquisition, doppler tracking, and positioning with starlink LEO satellites: first results. IEEE Trans Aerosp Electron Syst 58(3):2606–2610. https://doi.org/10.1109/TAES.2021.3127488
38. Study on new radio (nr) to support non-terrestrial networks (release 15), in: 3GPP TR 38.811, September 2020
39. Solutions fornrto support non-terrestrial networks (ntn) (release 16), in: 3GPP TR 38.821, May 2021
40. Witze A (2022) Unsustainable: how satellite swarms pose a rising threat to astronomy. Nature 606:236–237
41. Lawrence A, Rawls ML, Jah M, Boley A, Vruno FD, Garrington S, Kramer M, Lawler S, Lowenthal J, McDowell J, McCaughrean M (2022) The case for space environmentalism. Nat Astron 6(4):428–435
42. Palmroth M, Tapio J, Soucek A, Perrels A, Jah M, Lönnqvist M, Nikulainen M, Piaulokaite V, Seppälä T, Virtanen J (2021) Toward sustainable use of space: economic, technological, and legal perspectives. Space Policy 57:101428
43. Murtaza A, Pirzada SJH, Xu T, Jianwei L (2020) Orbital debris threat for space sustainability and way forward (review article). IEEE Access 8:61000–61019
44. Saito K, Hatta S, Hanada T (2019) Digital currency design for sustainable active debris removal in space. IEEE Trans Comput Social Syst 6(1):127–134
45. Hoyhtya M, Boumard S, Yastrebova A, Jarvensivu P, Kiviranta M, Anttonen A (2022) Sustainable satellite communications in the 6g era: a european view for multilayer systems and space safety. IEEE Access 10:99973–100005
46. Ruggieri M, Rossi T (2020) New fascinating challenges for space systems: softwarization, ai-based robotization and sustainability: which role for cubesats? Adv Astronaut Sci 173:609–615
47. Mishra D, Zema NR, Natalizio E (2021) A high-end IoT devices framework to foster beyond-connectivity capabilities in 5G/B5G architecture. IEEE Commun Mag 59(1):55–61
48. Mishra D, Vegni AM, Loscri V, Natalizio E (2021) Drone networking in the 6G era: a technology overview. IEEE Commun Stand Mag 5(4):88–95
49. Geraci G, Adrian Garcia-Rodriguez M, Azari M, Lozano A, Mezzavilla M, Chatzinotas S, Chen Y, Rangan S, Di Renzo M (2022) What will the future of uav cellular communications be? A flight from 5g to 6g. IEEE Commun Surv Tutorials 24(3):1304–1335. https://doi.org/10.1109/COMST.2022.3171135
50. Rinaldi F et al (2020) Non-terrestrial networks in 5G & beyond: a survey. IEEE Access 8:165178–165200
51. 3GPP TR 22.822 (2018) Study on using satellite access in 5G. http://www.3gpp.org/ftp//Specs/archive/22_series/22.822/22822-g00.zip
52. 3GPP TR 38.811 (2018) Study on new radio to support non-terrestrial networks. http://www.3gpp.org/ftp//Specs/archive/38_series/38.811/38811-f00.zip
53. 3rd Generation Partnership Project (3GPP) (2017) 36.777: study on enhanced LTE support for aerial vehicles, December 2017
54. 3GPP, TS 23.255 (2021) Application layer support for Uncrewed Aerial System (UAS), Functional architecture and information flows. https://portal.etsi.org/webapp/workprogram/Report_WorkItem.asp?WKI_ID=67770
55. Ye H, Chen M, Saad W (2020) Joint access and backhaul resource management in satellite-drone networks: a competitive market approach. IEEE Trans Wirel Commun 19(6):3908–3923. https://doi.org/10.1109/TWC.2020.2979127
56. Mishra D, Natalizio E (2020) A survey on cellular-connected UAVs: design challenges, enabling 5G/B5G innovations, and experimental advancements. Comput Netw 182:107451
57. Mishra D, Trotta A, Traversi E, Di Felice M, Natalizio E (2022) Cooperative cellular UAV-to-everything (C-U2X) communication based on 5G sidelink for UAV swarms. Comput Commun 192:173–184

58. Mishra D, Trotta A, Di Felice M, Natalizio E (2021) Performance analysis of multi-hop communication based on 5G sidelink for cooperative UAV swarms. In: 2021 IEEE international mediterranean conference on communications and networking (MeditCom). IEEE, pp 395–400
59. Bushnaq OM, Mishra D, Natalizio E, Akyildiz IF (2022) Unmanned aerial vehicles (UAVs) for disaster management. In: Nanotechnology-based smart remote sensing networks for disaster prevention. Elsevier, pp 159–188
60. Mozaffari M, Saad W, Bennis M, Nam Y-H, Debbah M'r (2019) A tutorial on UAVs for wireless networks: applications, challenges, and open problems. IEEE Commun Surv Tutorials 21(3):2334–2360
61. Bai T, Vaze R, Heath RW (2014) Analysis of blockage effects on urban cellular networks. IEEE Trans Wirel Commun 13(9):5070–5083. https://doi.org/10.1109/TWC.2014.2331971
62. Vinogradov E, Sallouha H, De Bast S, Azari MM, Pollin S (2018. ISSN 1550-4654.) Tutorial on UAVs: a blue sky view on wireless communication. J Mob Multimed 14(4):395–468. https://doi.org/10.13052/jmm1550-4646.1443
63. Clerckx B, Oestges C (2013) MIMO wireless networks, 2nd edn. Elsevier Academic Press
64. ITU-R (2012) Propagation data and prediction methods required for the design of terrestrial broadband radio access systems operating in a frequency range from 3 to 60 GHz, February 2012. https://www.itu.int/rec/R-REC-P.1410-5-201202-I/en
65. Al-Hourani A, Kandeepan S, Jamalipour A (2014) Modeling air-to-ground path loss for low altitude platforms in urban environments. In: 2014 IEEE global communications conference, pp 2898–2904. https://doi.org/10.1109/GLOCOM.2014.7037248
66. Molisch AF (2011) Wireless communications. Wiley, Wiley - IEEE
67. Al-Hourani A, Kandeepan S, Lardner S (2014. ISSN 2162-2337.) Optimal LAP altitude for maximum coverage. IEEE Wireless Commun Lett 3(6):569–572. https://doi.org/10.1109/LWC.2014.2342736
68. Azari MM, Rosas F, Chen K-W, Pollin S (2016) Joint sum-rate and power gain analysis of an aerial base station. In: IEEE Globecom workshops (GC Wkshps), pp 1–6
69. Mozaffari M, Saad W, Bennis M, Debbah M (2016) Efficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage. IEEE Commun Lett 20(8):1647–1650
70. Khawaja W, Guvenc I, Matolak D (2016) UWB channel sounding and modeling for UAV air-to-ground propagation channels. In: 2016 IEEE global communications conference (GLOBECOM), pp 1–7. https://doi.org/10.1109/GLOCOM.2016.7842372
71. Yanmaz E, Kuschnig R, Bettstetter C (2011) Channel measurements over 802.11a-based UAV-to-ground links. In: 2011 IEEE GLOBECOM workshops (GC Wkshps), pp 1280–1284. https://doi.org/10.1109/GLOCOMW.2011.6162389
72. Yanmaz E, Kuschnig R, Bettstetter C (2013) Achieving air-ground communications in 802.11 networks with three-dimensional aerial mobility. In: 2013 proceedings IEEE INFOCOM, pp 120–124. https://doi.org/10.1109/INFCOM.2013.6566747
73. Newhall WG, Mostafa R, Dietrich C, Anderson CR, Dietze K, Joshi G, Reed JH (2003) Wideband air-to-ground radio channel measurements using an antenna array at 2 GHz for low-altitude operations. In: IEEE military communications conference, 2003. MILCOM 2003, vol 2, pp 1422–1427. https://doi.org/10.1109/MILCOM.2003.1290436
74. Tu HD, Shimamoto S (2009) A proposal of wide-band air-to-ground communication at airports employing 5-GHz band. In: 2009 IEEE wireless communications and networking conference, pp 1–6. https://doi.org/10.1109/WCNC.2009.4917538
75. Matolak DW, Sun R (2015) Air-ground channel characterization for unmanned aircraft systems: the near-urban environment. In: MILCOM 2015–2015 IEEE military communications conference, pp 1656–1660. https://doi.org/10.1109/MILCOM.2015.7357682
76. Matolak DW, Sun R (2017. ISSN 0018-9545.) Air-ground channel characterization for unmanned aircraft systems—part III: the suburban and near-urban environments. IEEE Trans Veh Technol 66(8):6607–6618. https://doi.org/10.1109/TVT.2017.2659651

77. Maccartney GR, Rappaport TS, Samimi MK, Sun S (2015. ISSN 2169-3536.) Millimeter-wave omnidirectional path loss data for small cell 5G channel modeling. IEEE Access 3:1573–1580. https://doi.org/10.1109/ACCESS.2015.2465848
78. Azari MM, Rosas F, Chiumento A, Pollin S (2017) Coexistence of terrestrial and aerial users in cellular networks. In: 2017 IEEE Globecom workshops (GC Wkshps), pp 1–6. https://doi.org/10.1109/GLOCOMW.2017.8269068
79. Azari MM, Rosas F, Chen K, Pollin S (2016) Optimal UAV positioning for terrestrial-aerial communication in presence of fading. In: 2016 IEEE global communications conference (GLOBECOM), pp 1–7. https://doi.org/10.1109/GLOCOM.2016.7842099
80. Zeng Y, Zhang R, Lim TJ (2016) Wireless communications with un- manned aerial vehicles: opportunities and challenges. IEEE Commun Mag 54(5):36–42
81. Chakraborty A, Chai E, Sundaresan K, Khojastepour A, Rangarajan S (2018) SkyRAN: a self-organizing LTE RAN in the sky. In: Proceedings of the 14th international conference on emerging networking EXperiments and technologies, pp 280–292
82. Kimura T, Ogura M (2020) Distributed collaborative 3D-deployment of UAV base stations for on-demand coverage. In: IEEE INFOCOM 2020-IEEE conference on computer communications. IEEE, pp 1748–1757
83. Gapeyenko M, Petrov V, Moltchanov D, Andreev S, Himayat N, Koucheryavy Y (2018) Flexible and reliable UAV-assisted backhaul operation in 5G mmWave cellular networks. IEEE J Sel Areas Commun 36(11):2486–2496
84. Trotta A, Di Felice M, Chowdhury KR, Bononi L (2017) Fly and recharge: achieving persistent coverage using small unmanned aerial vehicles (SUAVs). In: 2017 IEEE international conference on communications (ICC). IEEE, pp 1–7
85. Trotta A, Andreagiovanni FD, Di Felice M, Natalizio E, Chowdhury KR (2018) When UAVs ride a bus: towards energy-efficient city-scale video surveillance. In: IEEE infocom 2018-IEEE conference on computer communications. IEEE, pp 1043–1051
86. Mozaffari M, Kasgari ATZ, Saad W, Bennis M, Me'rouane Debbah. (2018) Beyond 5G with UAVs: foundations of a 3D wireless cellular network. IEEE Trans Wirel Commun 18(1):357–372
87. Lagum F, Bor-Yaliniz I, Yanikomeroglu H (2018) Strategic densification with UAV-BSs in cellular networks. IEEE Wireless Commun Lett 7(3):384–387. https://doi.org/10.1109/LWC.2017.2779483
88. Oestges C, Vanhoenacker-Janvier D (2001) Coverage modelling of high-altitude platforms communicaton systems. Electron Lett 37(2):1
89. Kurt GK, Khoshkholgh MG, Alfattani S, Ibrahim A, Darwish TSJ, Alam MS, Yanikomeroglu H, Yongacoglu A (2021) A vision and framework for the high altitude platform station (haps) networks of the future. IEEE Commun Surv Tutorials 23(2):729–779. https://doi.org/10.1109/COMST.2021.3066905
90. Alam MS, Kurt GK, Yanikomeroglu H, Zhu P, Du˜ng a'o N (2021) High altitude platform station based super macro base station constellations. IEEE Commun Mag 59(1):103–109. https://doi.org/10.1109/MCOM.001.2000542
91. Alsharoa A, Alouini M-S (2020) Improvement of the global connectivity using integrated satellite-airborne-terrestrial networks with resource optimization. IEEE Trans Wirel Commun 19(8):5088–5100. https://doi.org/10.1109/TWC.2020.2988917
92. Oestges C (2008) Impact of tropospheric depolarization on the performance of dual-polarized high-altitude platform systems. In: VTC spring 2008 - IEEE vehicular technology conference, pp 2897–2901. https://doi.org/10.1109/VETECS.2008.631
93. Saunders SR, Arago'n-Zavala. A (2007) Antennas and propagation for wireless communication systems. Wiley
94. ITU-R (1992) Specific attenuation model for rain for use in prediction methods
95. ITU-R (1997) Attenuation by atmospheric gases
96. Peeters G, Marzano FS, d'Auria G, Riva C, Vanhoenacker-Janvier D (1997) Evaluation of statistical models for clear-air scintillation prediction using olympus satellite measurements. Int J Satell Commun 15(2):73–88

97. Li Y, Baccelli F, Dhillon HS, Andrews JG (2014) Fitting determinantal point processes to macro base station deployments. In: 2014 IEEE global communications conference, pp 3641–3646. https://doi.org/10.1109/GLOCOM.2014.7037373
98. Deng N, Zhou W, Haenggi M (2015) The Ginibre point process as a model for wireless networks with repulsion. IEEE Trans Wirel Commun 14(1):107–121. https://doi.org/10.1109/TWC.2014.2332335

Artificial Intelligence and Machine Learning in 6G

Shashwat Mishra, S. Muhammad Karam, Chung Shue Chen, and Luca Rose

1 Introduction

The demand for connectivity in our society is ever on the rise. Internet traffic has been growing at an annual rate of about 26%, from 122 exabytes per month in 2017 to 396 exabytes per month by 2022 [1]. As 5G deployments are rolling out publicly worldwide, the industry and academia are already gearing up for the next leap forward. With the scarcity of available communication spectrum in the wireless domain, the most efficient use of resources is paramount to achieving connectivity, latency, and reliability targets. The future networks present us with a unique set of challenges characterized by harsh propagation conditions in the newly opened-up spectrum coupled with increasing demands on all fronts including user connection density, peak data rates, ubiquitous network availability, and imperceptible latency. At the same time, it is imperative to ensure backward compatibility and harmonious co-existence through techniques such as spectrum sharing.

The proliferation of device-centric communication, as opposed to human-centric communication and the advances in the field of artificial intelligence (AI), presents us with a novel opportunity to develop AI-driven networks [2, 3]. Native AI-driven network architecture refers to the design and implementation of artificial intelligence (AI) systems that are integrated into network infrastructure and operation. Although AI is often thought of as a system in itself, it is a set of technologies implemented in a system to enable it to reason, learn, and solve complex problems. These systems use machine learning (ML) algorithms and techniques to analyze data, make decisions, and perform tasks without human intervention. ML is a subset of artificial intelligence that enables a machine or system to learn autonomously

---

S. Mishra (✉) · S. Muhammad Karam · C. S. Chen · L. Rose
Nokia Bell Labs, Paris-Saclay Center, Massy, France
e-mail: shashwat.mishra@nokia.com; muhammad.shehzad@nokia.com

V. Loscri et al. (eds.), *The Road towards 6G: Opportunities, Challenges, and Applications*, https://doi.org/10.1007/978-3-031-42567-7_5

and improve from experience. Instead of explicit programming, ML depends on algorithms to analyze large amounts of data, derive crucial insights, and then make informed decisions. While AI encompasses the idea of a machine that can mimic human intelligence, ML does not. Rather, ML aims to teach a machine or system how to perform a specific task and provide accurate results by identifying underlying patterns [4].

There are several key features that distinguish native AI-driven network architecture from other types of AI systems.

- Self-learning capabilities: Native AI-driven network architecture has the ability to learn and adapt to new data and situations on its own. This is achieved through the use of machine learning algorithms that can analyze data and improve the performance of the system over time.
- Real-time decision-making: Native AI-driven network architecture is designed to make decisions and perform tasks in real-time, without the need for human intervention. This allows the system to respond quickly to changing conditions and perform tasks more efficiently than humans.
- Scalability: Native AI-driven network architecture is designed to handle large volumes of data and operate at scale. This is important for network infrastructure, where there may be millions of devices and users generating data that needs to be analyzed and acted upon.
- Integration with existing systems: Native AI-driven network architecture is designed to be integrated with existing network infrastructure and systems. This allows the system to seamlessly operate alongside other technologies and techniques, improving efficiency and decision-making.
- Robustness: Native AI-driven network architecture is designed to be robust and reliable. This is important for network infrastructure, where downtime or failure can have significant consequences.

Motivated by the promising advances in the signal-processing capabilities of general hardware, increased access to cloud computing resources and developments made in the area of ML algorithms, we analyze the impact of AI and ML technologies on the upcoming generation of communication networks. The rest of this chapter is organized as follows:

- To begin with, in Sect. 1 we further explore the key drivers of the 6G growth in a holistic manner followed by quantifying the actual requirements proposed for the future networks.
- In Sect. 2, we summarize the principle advantages of ML techniques in future networks and make a case for utilizing learning-based paradigms to improve the connectivity, reliability and robustness of 6G networks.
- We consolidate the current state-of-the-art literature on the use of AI for enhancing the physical layer of 6G networks, ranging from waveform design to channel coding, in Sect. 3.
- In the subsequent Sect. 4, we extend our analysis to the MAC layer and highlight the advantages of ML-based radio resource allocation.

- We show how ML-based processing may be used to reinforce existing communication networks, where we take the example of orthogonal frequency division multiplexing and investigate the various improvements that may be made with the use of ML.
- Section 6 presents the possible directions of future communication systems, their use cases, deployment scenarios, target metrics and how they benefit from AI strategies.
- We provide a brief description of some popular ML techniques and evaluate their applicability to certain communication scenarios in Sect. 7.
- To complete the discussion on the inclusion of ML in 6G, in Sect. 8 we elaborate on how ML solutions may be designed and deployed in practical networks through the application of MLOps workflow.
- Finally, in Sect. 9, we introduce some datasets and code repositories that are pertinent to ML-enhanced communication schemes and may be used as a first stepping stone for interested practitioners to develop their own schemes.

1.1 Drivers for 6G Development

The demand for extreme connectivity through 6G can be attributed to the exponential growth in many parts of the world. The growth in the economy is usually driven by an expansion in production capabilities which in turn necessitates the development of the supply-chain infrastructure. With the advancement in technology, a lot of automation is helping drive productivity to levels that have never been witnessed before. As such, we may classify the advent of 6G due to four major causes:

- Economic Growth: The contribution of data to revenues from Internet advertising would amount to USD 54 billion [5]. In Europe, the funding gap to meet the 2025 Digital Agenda and Gigabit Strategy objectives are projected at USD 453 billion [6]. A substantial amount of sales now happen through online platforms, with the US estimating an e-commerce activity of USD 3576.5 billion in the manufacturing sector alone for the year 2020 [7]. Even factory floors are heavily automated and equipped with sensor networks for monitoring activity and production levels. The workhorse for Industry 4.0 is a densely connected cyber-physical environment, the end users of which are catered by 5G networks at present. As the economy grows, we would require 6G and beyond networks to be already in place and capable of handling this explosive growth.
- Environmental Sustainability: All advancements that the human collective makes must be carefully balanced against our impact on the environment. For 6G, this issue becomes more pressing due to the limited availability of resources. An alternative direction for green communications is to utilize energy harvesting techniques which adopt renewable energy or ambient energy to reduce the usage of fossil resources. AI techniques can be adopted to address the uncertainty and dynamics in the energy harvesting process as well as to save energy through

strategic standby for base stations [8]. Furthermore, we envision using satellite imaging for safeguarding the illegitimate exploitation of natural resources such as excessive mining and deforestation and taking rapid action against the aggressors utilising unmanned aerial vehicles (UAVs). Additionally, we may leverage future networks for an active survey of the climate and weather phenomenon [9–11].

- Social Advancement: Future networks can empower individuals and enable data-driven innovation in the private and public sectors to improve scientific and education outcomes. Access to the internet is quickly getting recognised as a fundamental necessity at par with water, food and housing. Existing networks have proven instrumental in managing public health emergencies such as the COVID-19 pandemic and enhancing energy resilience and the clean energy transition [12]. The availability of network access would prove to be a great equalizer, creating opportunities for employment across borders, harsh landscapes and war-affected areas. The design of future networks must take into account fair-access opportunities for people across the globe and be collaboratively managed to avoid isolated bottlenecks due to differences in political ideologies.
- Technology Innovation: The real costs of both computer processing power and data storage have halved every 15 months over the second half of the last century [5]. This has generated fertile ground for the development of feasible infrastructure for economical ubiquitous connectivity. Simultaneously improving the performance of transmission while enabling the capabilities of services such as sensing in the 6G network are key directions in 6G research. The link transmission capability of 6G is expected to be improved by 100 times that of 5G to achieve a terabit per second (Tbps) target and support the demands of versatile applications such as virtual reality (VR)/augmented reality (AR), automated vehicular systems, remote surgery and so on. In addition to the advanced radio interface technologies that aim to improve spectral efficiencies, such as waveform and modulation, increasing the frequency bandwidths and considering a variety of carrier frequencies are needed. However, considerations should be taken for the best system efficiency rather than simply the spectral efficiency, which must include the minimization of total power consumption and supporting heterogeneous quality of service for devices.

1.2 Key Performance Indicators and Related Use Cases

6G is still in the early stages of development and there is no universal agreement on its end goals. Nonetheless, we may expect an improvement in some core metrics based on past trends and the requirements for prevailing use cases.

1.2.1 Data Rates

One key performance indicator for 6G communication would be the data transfer rate. With 5G technology, data transfer rates are already incredibly fast, reaching speeds of up to 10 Gbps. However, 6G is expected to push these speeds even further, with some estimates suggesting that it could reach peak data rates of up to 1 Tbps [13, 14]. Such high data rates are necessary to support throughput-intensive use cases like high-fidelity video streaming, AR gaming, online digital twins and virtual digital persona such as the metaverse. We also envision the user-experienced data rates going up to 1 Gbps from the present 100 Mbps to enable enhanced mobile broadband coverage with better quality of service for the users.

1.2.2 Connection Density

6G technology will need to provide near-universal coverage in order to support this growing demand. This will likely involve the deployment of a dense network of small cells, which are low-power wireless access points that can be deployed quickly and inexpensively. Alternatively, the number of served devices in a geographical area is expected to increase tenfold from 10^6 devices/km^2 in 5G to 10^7 devices/km^2. Additionally, a network densification growth factor of approximately 30% is expected every 3 years [15]. This increase is necessary to enable machine-type communication and the omnipresent Internet of Things (IoT). This goal of network densification is the key to achieving automated factory floors, smart cities with automatic metering, and wireless sensor networks for wide-area safety monitoring to name a few. We envision a 1000-fold increase in network capacity compared to the current networks.

1.2.3 Energy Efficiency

As the number of connected devices grows, the power demands of wireless networks are only going to increase. 6G technology will need to be designed with energy efficiency in mind in order to adhere to the environmental commitment that we discussed in Sect. 1.1. This could involve the use of new technologies such as energy-efficient modulation schemes and advanced antenna designs. Even though the power consumption per unit of traffic (Watt/bit) is greatly decreased, the power consumption of 5G increases greatly compared to that of 4G with the maximum power consumption of a 64T64R active antenna unit is 1000–1400 W, and that of a base-band unit is about 2000 W [16]. In 6G networks, we expect to achieve 2 times the current energy efficiency. This metric is also important for low-power devices such as the RedCap class of devices defined in the 3GPP release 17 specification [17]. Here energy efficiency equates to longer battery life with industrial sensors requiring up to a few years of service through a single charge.

1.2.4 Latency and Reliability

There is a trade-off between throughput, reliability and latency for classical communication. Low latency may be achieved by a shorter packet but it will cause a degradation in channel coding gain and result in a decrease in reliability. Conversely, to improve reliability at a constant spectral efficiency, a bigger number of re-transmissions can be used and latency requirement limits the number of re-transmissions in URLLC transmission.

Moreover, if more time domain resources are consumed due to an increase of parity check bits in the low code rates, it also increases latency and reduces the system efficiency. For 5G, the 3GPP requirement for the ultra-reliable low latency communication (URLLC) use case for one transmission of a packet is 10^{-5} for 32 bytes with a user plane latency of 1 ms [18]. This reliability requirement poses a challenge in URLLC design because it is much higher than the typical block error rate of the Long-Term Evolution (LTE) system of 10^{-2}. However, 6G pushes the boundary of reliability even further with an air interface latency of 0.1 ms and a reliability target of 10^{-7}. These targets are crucial to supporting mission-critical use cases like automated vehicles, fleets of unmanned aerial vehicles and remote actuator movement for medical procedures [19, 20].

1.2.5 Mobility and Positioning Accuracy

The targeted mobility in 5G today is up to 500 km/h [21]. This already covers most of the connected objects including high-speed trains and flying objects such as drones. It is likely that this target would stay unchanged for the short term. For the longer-term 6G may be designed to support flying objects travelling in excess of 500 km/h (e.g. UAVs, aeroplanes).

Several use cases, especially in industrial control require 100 cm level positioning accuracies which are in agreement with the IEEE 802.11az (next-generation positioning) targets. The aim here is to support automated warehouses and manufacturing facilities. There are provisions in the 3GPP R18 RAN1 work group specifically designed to study the role of AI/ML in positioning. In current discussions on enhanced positioning in 3GPP Rel-17, the expected positioning accuracy is about 10–30 cm for several use cases. The move to higher frequencies and wider bandwidths is anticipated to increase the positioning accuracy. Finer resolutions are achievable in higher frequency bands. Furthermore, cm-level accuracy is possible today through sensing mechanisms (e.g. LiDAR). The target accuracy is therefore envisioned to improve to below 1 cm for the indoor environment and about 50 cm for outdoor environments. Such accuracy is essential for user tracking for beamforming as well as vehicular coordination.

In the subsequent sections, we take a look at the avenues in 6G communication where ML paradigms can significantly enhance network performance and discuss the underlying use cases (Fig. 1).

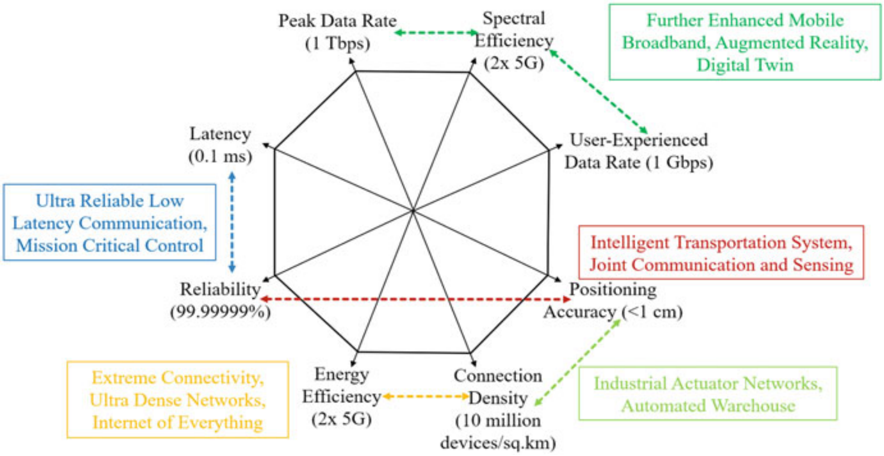


Fig. 1 Major KPIs and relevant use cases for 6G next generation networks

2 The Role of AI and ML in 6G

In recent years, we have seen astonishing progress in the field of AI and ML, with these technologies being used in a wide range of applications, from image and voice recognition to autonomous vehicles and beyond. As we move into the era of 6G, it is certain that these technologies will continue to play a central role in driving innovation and progress. Industries have geared up to show enthusiastic support for the amalgamation of ML techniques into their existing architectures and solutions. This has also pushed several regulatory bodies around the world to look into establishing guidelines governing the scope and effect of AI in communication systems.

As we move into the 6G era a wide range of new technologies will be developed, including advanced sensors, smart devices, and other technologies that will help to drive the next wave of innovation. AI and ML algorithms will be used to analyze data from these technologies, identify patterns, and make predictions and decisions. This will help to drive the development of new and exciting products and services that will benefit people around the world. However, these new developments come with a unique set of challenges that can be resolved efficiently through the application of well-designed ML techniques. AI shows immense promise to enhance existing networks and provide solutions to problems in future networks, often caused due to the complex interaction of a fast-changing environment, user traffic and resource availability.

The key contributions of AI to future networks are broadly summarized into the following classes:

2.1 Decreased Computational Complexity

One key advantage of using AI is that approximate solutions may be obtained to traditionally complex problems, such as sphere decoding for MIMO systems [22] and decoding error-correcting codes [23]. These solutions usually outperform classical computation-intensive solutions in terms of the floating point operations required, since the ML model is trained to achieve a specific level of accuracy under a certain performance time. The process of training the model is often the most labour-intensive part of the entire solution [24]. Once a fully-trained model is obtained, the inference may be performed in a fraction of a second.

This approach is synonymous with heuristics instead of calculating the exact solution, with the difference that we aim at developing machine-aided heuristics for a specific use case and some specific performance goal. The use of ML makes this process faster and can even be extended to novel use cases that do not have a closed-form solution so far. As a rule of thumb, the ML model must be used as a rough draft to develop upon the heuristic as the models are usually agnostic to certain domain knowledge that a human expert may have. As such, having a model is not enough. The designer must also be able to explain the underlying pattern to some extent, which falls under the purview of explainable AI. Due to the rapid development of hardware capabilities, even simple devices can now run ML models locally for inference and most medium to high-end consumer mobile handsets now have dedicated ML hardware. Additionally, many problems in wireless communication fit perfectly in the framework of well-studied ML problems like sequence detection [25] and reinforcement learning based on a partially observed environment. This helps us solve key problems such as interference management, power allocation and user-resource pairing.

2.2 Improved Robustness and Versatility

The use of ML is especially fruitful in the area of multi-objective optimization, even more so when the objectives are conflicting. For example in the case of low-power sensor networks, reliability must also be taken into account. However, ensuring reliability through added parity bits or channel coding will cause the power consumption to increase. A similar problem occurs when trying to ensure user fairness in a resource-constrained network. In such cases, ML techniques can help us find the balance between various parameters to achieve the desired system performance and are often easier to execute than traditional optimization techniques. Furthermore, ML models are re-trainable on the fly. A single model trained on multiple scenarios can very well maintain good performance in the changing morphologies [26]. Additionally, ML models can adapt to scenarios for which they have never been trained, which makes an ML-supported system more robust. For instance, the problem of automatically adapting radio maps to changing deployment in Wi-Fi-based indoor localization can be readily addressed using transfer learning, without necessarily conducting the site survey [27].

2.3 Leveraging Data Traces

Most processes keep a log of various system parameters for monitoring and alarm triggering. ML techniques have shown great promise in their pattern-detection abilities [28]. Using system logs, ML models can generate resource provisioning schedules for base stations that can help us conserve up to 90% more energy than the current standard deployments [29].

Similarly, we can utilise user activity statistics to perform efficient scheduling as well as predictive resource reservation based on predicted traffic load. This is especially beneficial for supporting high-density use cases, where a limited set of network resources can be proactively re-used to serve a larger user base. Conventional radio resource management does not take into account specific seasonal patterns as the algorithms are usually data agnostic, however, with ML schedulers we get a twofold advantage. Firstly, we can actually discover patterns based on correlated user activity, location or operating conditions and secondly, we can use these prompts to create solutions that are more suited for a specific deployment. Although specificity and generalizability seem to be diametrically opposite performance goals, through careful fine-tuning, we can make use of both these regimes depending on the context of service.

2.4 Knowledge Transfer and Adaptability

Even though AI seems to be competing against traditional algorithms, it is actually a complementary tool that can be used to enhance the performance of existing schemes. More often than not, the initial training data for popular ML schemes are taken from conventional solutions obtained using numerical methods like successive convex approximation. In effect, ML techniques may be seen as advanced numerical solvers and as such may be employed side-by-side with traditional techniques to either extend their results through domain knowledge-guided interpolation or to find a suitable candidate solution through a data-driven approach.

Transfer learning is a promising technique for quick prototyping. The goal here is to start from an already developed solution for one scenario and adapt it as an initial point for finding the solution for another scenario rather than starting from scratch. This methodology works best when trying to find solutions for dual problems. As an example, power consumption and device density may be considered dual problems in the downlink for many scenarios. We have a larger literature on power consumption maximization or energy efficiency maximization that may be used to address the metric of scheduling throughput which refers to the number of

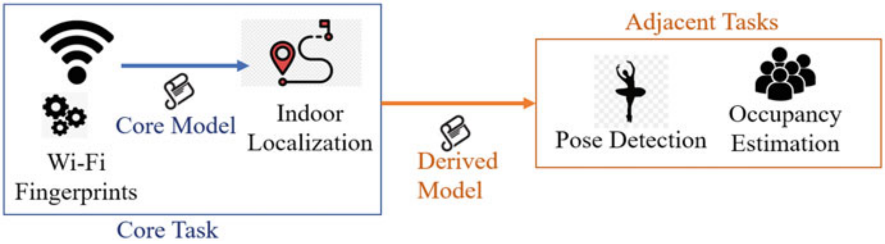


Fig. 2 Robust models can address multiple related problems by transferring the learned characteristic from one domain to another

connected devices. Another example exists in the domain of positioning, where computer-vision data sets can be used to augment positioning results obtained through Wi-Fi positioning. This type of cross-domain combining helps in reducing the size of location fingerprint databases which are often a bottleneck in terms of initial construction and storage for dense deployments. Additionally, we may solve multiple tasks using a single model such as pose detection and occupancy estimation using a positioning model [30] (Fig. 2).

3 AI Enhanced PHY Layer

The physical layer is the soul of all communication systems. At the physical layer, machine learning algorithms can be used to optimize the design of antennas, transmitters, and receivers. By using data-driven optimization techniques, the hardware can be designed to provide improved coverage, better beamforming, increased power efficiency, and enhanced interference management. In the signal processing stage, machine learning algorithms can be used to develop advanced techniques for channel estimation, equalization, and multi-user detection. These algorithms can learn from the data collected from the network and adapt to the changing channel conditions to provide improved performance and reliability. In the following sections, we shall demonstrate the efficacy of learning-based methodology in some key aspects of the physical layer.

3.1 Waveform Design

One of the key challenges in the development of 6G is the ability to efficiently transmit and receive data in the presence of dynamic environmental challenges, such as interference, noise, shadowing and fading. Traditionally, waveform design has been based on mathematical optimization techniques, which focused on finding the optimal waveform that maximizes the performance metrics, such as spectral efficiency, energy efficiency, and robustness. However, these methods have several limitations, such as the assumption of perfect channel state information, the high computational complexity, and the inability to adapt to the changing environment.

By using data-driven techniques, such as deep learning and reinforcement learning, machine learning algorithms can learn the characteristics of the communication environment and adapt to changing conditions. This enables the waveform design to be more robust and efficient, improving the overall performance of the communication systems. The pioneering works in [31, 32] demonstrate the efficacy of end-to-end learning systems applied to legacy OFDM waveform, achieving a PAPR reduction of 2 dB without any change in the hardware. Additionally, we can design waveforms suited to specific purposes in a given scenario such as anti-jamming [33]. Here the ML model leverages the spatiotemporal characteristics of the communicating agents to propose candidate waveforms. Furthermore, we may use ML techniques to support next-generation multiple-access techniques where waveform design focuses on joint user detection and channel estimation [34].

6G is expected to further support advanced multiplexing techniques, such as non-orthogonal multiple access (NOMA) and rate-splitting multiple access (RSMA) for increasing the spectral efficiency of networks while utilising limited spectrum, These multiplexing techniques often lead to superimposed constellations that are difficult to decode using the Euclidean distance metric, leading to worse bit-error performance. This problem can be eliminated by deep learning techniques where the bit-to-symbol mapping is based upon the interaction between the multiple user data streams. This use case benefits from the strong classifying ability of neural networks, where great accuracy can be achieved in differentiating near-identical data points in higher metric spaces with low computational complexity.

3.2 Antenna Design

By leveraging the flexibility and adaptability of machine learning algorithms, the waveform design can be more robust and efficient, leading to better communication performance in 6G networks. Machine learning has become a significant tool in the field of antenna design, particularly in the development of 6G technology. The traditional approach to antenna design involves the use of analytical and numerical techniques to optimize antenna parameters, such as size and shape, for a specific application. However, this approach can be time-consuming and labour-intensive, and it may not be able to capture the complex interactions between the antenna and its environment. Machine learning, on the other hand, can automate the optimization process and enable the design of antennas that can adapt to changing conditions. This is achieved through the use of algorithms that can learn from data and make predictions based on that knowledge.

One application of machine learning-based antenna design is the use of deep learning algorithms to optimize the antenna radiation pattern. By training a deep learning model on a large data set of antenna designs and their performance characteristics, the model can learn to identify patterns and predict the best antenna shape for a given application. This approach can greatly reduce the time and effort required for antenna design, and it can also improve the performance of the antenna by finding more optimal solutions than traditional methods [35]. Another crucial advantage of using ML for antenna design is the ability to optimize for non-conventional use cases such as joint communication and sensing as well as re-configurable intelligent surfaces [36].

3.3 Channel Prediction

The prediction of channels is instrumental for the efficient operation of 6G networks. It allows the network to determine the optimal transmission parameters, such as the frequency and power, for each communication link, which in turn ensures the quality of service for the user. To predict channels, machine learning algorithms are trained on a large dataset of channel measurements collected from various locations and environments. These algorithms can then be used to estimate the characteristics of a channel based on the available information, such as the distance and orientation between the transmitter and receiver, and the type of environment. In contrast to this, another approach is to use unsupervised learning, where the algorithm is trained on unlabeled data and learns to identify patterns and relationships in the data without explicit guidance. This can be useful in scenarios where the true channel characteristics are not known, or when the environment is changing rapidly. ML-based predictors have lesser operational complexity than Kalman Filter (KF)-based predictors [37]. Additionally, ML models can take crucial practical factors like channel ageing into consideration without the need for an explicit mathematical model and perform well on complex non-linear signal models [38, 39]. ML techniques enable channel prediction not only for future transmit time intervals but also over different spatial locations by leveraging geographical traffic information.

In addition to predicting the characteristics of channels, machine learning algorithms can also be used to optimize the transmission parameters of 6G networks. For example, the algorithm can learn the relationship between the transmission parameters and the channel characteristics and determine the optimal amount of feedback information. Experimental verification on real-world datasets has shown near-perfect reconstruction based on quantized channel state information reporting while greatly reducing overheads [40, 41].

3.4 Source and Channel Coding

Traditionally, source and channel coding are designed using statistical models of the source and the channel, respectively. These models are typically based on assumptions and simplifications that may not accurately reflect the true characteristics of the source and channel. Shannon's separation theorem states that the two-step separate source and channel coding approach is theoretically optimal in the asymptotic regime of infinitely long source and channel blocks [42]. However, in practical applications, joint source and channel coding are known to outperform the modular approach. This becomes especially relevant for short-blocklength and low-SNR regimes which are important for emerging technologies like IoT, mMTC and ITS. The exact characterization of the optimal joint source and channel codes in non-asymptotic regimes remains an open problem, even for fully known source and channel distributions. A deep joint source and channel coding architecture can be trained to map the underlying signal samples directly to channel inputs [43]. There is an intuitive correspondence between the autoencoder structure and a typical communication system. The work in [44] presents an end-to-end model that replaces the modular structure of conventional communication systems consisting of separate blocks for data compression, channel coding, modulation and channel estimation while performing well in short blocklength regimes.

4 AI Supported MAC Layer

Machine learning-aided medium access control (MAC) is an emerging methodology that has the potential to revolutionize the way wireless networks operate in the 6G era. With the increasing demand for higher data rates and lower latency, MAC protocols need to be intelligent and adaptable to the changing network conditions. Machine learning algorithms can provide the necessary flexibility and adaptability to MAC protocols, allowing them to better serve the needs of 6G networks. The MAC layer plays a key role in providing functionalities such as user fairness, device authentication, connection grant and QoS guarantees.

4.1 Distributed Grant Free Access

MAC layer design is increasingly becoming distributed and grant-free. This is in part owing to the emerging wireless communication systems on unlicensed spectra, such as Wi-Fi, ZigBee, and SigFox, due to their lack of centralized management and well-governed infrastructure. Additionally, the growth of user traffic supersedes the available radio resources. This is especially true for the mMTC use case where the cost of scheduling can often be higher than the size of the information packet itself, promoting the superposition of user requests and messages on the same time-frequency resources.

Most conventional distributed access mechanisms are based on random access schemes, which require efficient collision resolution. The most popular random access scheme is carrier-sense multiple access with collision avoidance (CSMA/CA). Upon a collision, the user picks up a random back-off time period to defer its transmission. However, the intrinsic random deferment feature makes them upper-bounded by a relatively low MAC efficiency and suffers from severe fairness issues in many realistic scenarios [45], which is unsuitable for the low latency requirements of future networks. Multi-agent reinforcement learning-based solutions have shown great promise, outperforming CSMA/CA and even its theoretical performance bound in various scenarios including saturated traffic, unsaturated traffic and delay-sensitive traffic at marginally increased computational costs [46].

In 5G, grant-free (GF) or 2-step random access schemes are introduced for reducing signalling overhead in dense machine networks. While GF random access can be more efficient in terms of spectral efficiency due to a high spatial multiplexing gain with massive MIMO, its performance is limited by the number of preambles [47]. Therefore, we may use ML-designed preambles that utilize the underlying user-activity patterns.

4.2 Radio Resource Allocation

Resource allocation in 6G networks would be increasingly challenging as the mobile environment would become increasingly complex due to the interaction of various simultaneously existing competing technologies. As such, typical tasks such as multi-cell scheduling would become untenable without proper design considerations. Deep learning (DL) is a powerful tool where a multi-layer neural network can be trained to model a resource management algorithm using network data. This design paradigm has shown substantial gains, reducing the execution time for multi-cell power allocation by a factor of 10 as compared to traditional heuristic methods [48].

The increasing heterogeneity of future networks introduces new resource management tasks such as the distribution of resources for sensing and communication, caching in edge networks and energy harvesting. Most existing ML algorithms require accurate reward feedback and large-scale sample collection. In many cases, no prior samples are available in advance due to sparse activity or rarity of events, and performance feedback is corrupted by noise and interference. In modern networks, the underlying traffic distribution requires fast adapting RRM policies due to the complex nature of mixed user traffic. ML techniques enable various network components such as base stations, edge servers, gateway nodes, and user devices to make autonomous and local decisions while allowing for various levels of cooperation that help in performing joint optimization of several resource allocation tasks that makes networks more robust.

4.3 Network Optimization

The MAC layer has been increasingly supporting the paradigm of modularity through techniques such as network functions, which decouple the functionalities of a network from the physical devices on which they run such as routers and switches. The focus instead is to develop software-controlled generic hardware that can be used flexibly which leads to significant reductions in expenses and facilitate the deployment of new services with increased agility and faster time-to-value.

Extending the same ideology, network slicing has been developed as a driver for 5G heterogeneous service support [49]. Network slicing allows for the creation of multiple virtual networks on a single physical network infrastructure. The key reason for the development of this technology is supporting a wide range of verticals with a diverse set of performance and service requirements, which is only going to further expand going towards 6G. Machine learning algorithms can be used to predict the demand for network resources and allocate slices to heterogeneous services accordingly.

Furthermore, machine learning can be used to optimize the performance of individual slices. For instance, machine learning algorithms can be used to identify the most efficient routing paths and allocate network resources accordingly. This is made possible through predictive traffic analysis, aimed at finding characteristic trigger events and seasonality [50]. Such data-driven schemes improve the overall performance of the network slices and enable them to support applications with contrasting resource requirements.

4.4 Security

The principal concerns for data-driven systems are user privacy, transparency in the use of the collected data and model security. With the increasing push for network functionalities to be software-defined, the risk for malicious attacks has become multi-fold. In the cloud-based learning architecture, a malicious attack during the transmission of DNN parameters trained in the cloud could seriously reduce RRM efficiency and incur excessive training costs. There is also the risk of privacy disclosure in the uploaded data samples by RRM entities since the data may be location-dependent. AI can help to identify and mitigate these threats, by analyzing network data to identify anomalies and potential attacks, and by learning from previous security incidents to continuously improve security measures. Adversarial models may be employed to harden the network against threats and blind spots that may be too complex to be hand-crafted. For efficiently supporting diverse distributed systems, learning the data locally and generating the learning models and inference with limited coordination and information exchange is instrumental [51, 52].

5 ML Enhanced OFDM Systems: A Case Study

Orthogonal Frequency Division Multiplexing (OFDM) has been the backbone of modern wireless technologies such as WiFi, 4G, and 5G. The technology for supporting OFDM-based communication is mature and extensive theoretical and experimental studies have been conducted to establish the bounds for communication with this modulation scheme. Yet there are limitations pertaining to the use of OFDM to support beyond 5G communication, especially in terms of spectral efficiency and the need for high linearity power amplifiers. We present this example as it is backwards compatible with existing technologies and would most likely remain the dominant signalling scheme for many 6G use cases with limited hardware capabilities. In the case of an OFDM receiver, the machine learning algorithm can be trained to recognize patterns in the received signals and use these patterns to estimate the CIR and the transmitted symbols. Such convolutional neural network-based de-mappers can jointly process a large number of OFDM symbols to compute better log-likelihood ratios (LLRs) by compensating for channel ageing without requiring perfect channel state information [53]. This system can guarantee a bit-error rate of 10^{-3} even at high speeds of 130 km/h. There are lightweight models that enable efficient online channel estimation and equalization even with non-linear distortion, making such implementation deployable in real systems [54].

ML techniques also allow for the end-to-end design and optimization of the system, allowing us to perform conventionally challenging tasks such as blind equalization. Another major challenge in OFDM systems is the high Peak-to-Average Power Ratio (PAPR). High PAPR results in higher power consumption and can cause non-linear distortion in the transmitters, leading to reduced system performance. To reduce PAPR, various techniques have been proposed, such as clipping and filtering, selective mapping, and tone reservation. However, these techniques have their own limitations, such as reduced data rate, increased complexity, and reduced system performance. To overcome these limitations, AI-based OFDM receive techniques have been proposed where constellation mapping and demapping of symbols on each sub-carrier in an OFDM system are learned through the training of DNN to lower the PAPR while simultaneously minimizing the degradation of the bit-error rate, with negligible operational overhead and up to 2 dB lower PAPR than the current state-of-the-art schemes [55].

6 Next Generation Networking Paradigms

6.1 *Cell-Free Massive MIMO*

Cell-free massive MIMO (CFmMIMO) is a new approach to wireless communication that uses a large number of distributed antennas to serve multiple users, such that there are no coverage holes and hard cell edges due to deployment geometry.

This allows for increased data rates and improved signal quality overtaking the performance of small-cell systems, making it an ideal technology for 6G networks [56]. However, the sheer amount of data generated by a CFmMIMO system can be overwhelming, making it difficult to effectively manage and process. New hardware developments such as radio stripes are being envisaged to support CFmMIMO. Machine learning algorithms can be used to analyze and process the vast amounts of data generated by a CFmMIMO system. This allows for real-time optimization of the system, resulting in improved performance and efficiency. Machine learning can be used to detect and mitigate interference from other devices, resulting in improved signal quality. This is particularly important in dense urban environments where there are many competing wireless deployments. Machine learning can also facilitate power allocation, which is usually computationally intensive, making it 50–100× faster than conventional methods [57]. Other processes that benefit from ML-based processing include beam-forming, pre-coding, pilot allocation and channel state acquisition [58].

6.2 Ultra Massive Machine-Type Communication

6G traffic is expected to be dominantly machine-type, where millions of devices are connected to the network and need to share data in real time. These kinds of deployments produce a unique challenge as the device communication needs to be nearly automated. The traffic is sporadic and the communication is dominated by short packet transmission making traditional grant-based scheduling difficult. One way ML can be used for mMTC in 6G is through the use of clustering algorithms that group devices with similar requirements together while creating minimum interference, ensuring that each device receives the necessary quality of service. ML also enables the use of non-orthogonal resource allocation for such systems through efficient code-book design and beam selection in the case of multi-antenna systems. mMTC traffic has two characteristic types: event-driven and periodic. It is difficult to find theoretical models to generate an accurate representation of such mixed traffic scenarios, where the frequency of device activity may vary from a second to a few months. Here again, we may leverage learning techniques to infer correlated device activation and generate predictive fast grant schedules for resource reservation in case of guaranteed services [59] and perform faster active user-detection than prevalent sparse detection techniques in case of grant-free transmission.

6.3 Sub-terahertz Communication

Sub-Terahertz (THz) communication is a new form of wireless communication that uses electromagnetic waves in the terahertz frequency range of up to 200 GHz. These frequency ranges are promising bands for 6G and beyond networks owing to the unused and unexplored spectrum [60]. These waves have high frequency, high bandwidth, and low penetration, making them ideal for high-speed, short-range outdoor communication such as person area networks for augmented reality and health care. However, the use of the THz spectrum has been limited due to the challenges posed by their unique properties. For instance, THz waves are easily absorbed by atmospheric water molecules, leading to poor transmission range and reliability. In addition, THz communication systems require complex and expensive equipment, which makes them challenging to implement on a large scale. So far, THz waves have been used for medical imaging and material analysis, where ML techniques are often employed to improve imaging and detection quality. ML can play a crucial role in overcoming these challenges and enabling the widespread adoption of THz communication, especially through reducing the effect absorption losses by strategic transmission planning based on learning environmental factors. ML can be used to develop adaptive beamforming algorithms that can steer the pencil beams spanning only a few radians precisely focused in the receiver's direction, improving the transmission range and reliability.

6.4 Intelligent Transportation Systems

Intelligent transportation systems (ITS) refers to the use of advanced technologies to improve the efficiency, safety, scalability, and sustainability of public and private transportation systems [61]. These systems often rely on real-time data and low-latency communication to optimize the flow of traffic, reduce congestion, and prevent accidents. As the world moves towards the deployment of 6G networks, the potential for ITS to transform transportation is even greater, not only limited to the popular autonomous driving but extending to UAV fleets, automated ship freights, and warehouse storage systems. Autonomous vehicles rely on a variety of sensors and communication technologies to navigate and make decisions, and the high speeds and low latency of 6G networks could significantly improve the performance and reliability of these systems. With 6G, autonomous vehicles could potentially communicate with each other and with the surrounding infrastructure in real time, allowing them to coordinate their movements and avoid accidents. This use case is covered under the umbrella term of V2X communication [62].

Another area where 6G could have a major impact on ITS is in the realm of public transportation. Many public transportation systems, such as buses and trains, are equipped with sensors and communication systems that allow them to be tracked in real-time and to optimize their routes leading to faster and more reliable public transportation and streamlined automation. By enabling the integration of renewable energy sources, such as solar panels and wind turbines, into transportation systems, 6G could help to reduce the carbon footprint of transportation. Electric vehicles (EVs) could potentially be charged using renewable energy sources, and the high speeds and low latency of 6G could enable the real-time monitoring and optimization of the charging process. This could ultimately lead to a more sustainable transportation system. UAVs often operate in swarms, with multiple drones working together to complete a task. With 6G, drones can request precise positioning on the cm level and low latency encrypted network slices enabling them to coordinate their actions and make decisions based on the information received from other drones. The availability of non-terrestrial links will extend the coverage region for UAV communication [63].

6.5 Joint Communication and Sensing

Joint communication and sensing is an evolving set of advanced communication paradigms supported by the integration of communication and sensing capabilities in a single device or system. The integration of advanced technologies, such as sensors, communication networks, and intelligent software improves the efficiency and safety of transportation systems. Well-designed sensing signals can reduce the signalling overhead by a factor of 2.67 for road traffic monitoring use case [64]. There are two strong driving forces behind the adoption of joint communication and sensing in future networks. First, the use of THz frequencies allows for effective monitoring and sensing of environmental factors such as moisture. Furthermore, the ubiquitous presence of distributed antenna systems spanning different wireless spectra provides a stable framework for ranging and detection applications. Through the application of ML-powered strategic data fusion, we can leverage the diversity of measurements for the same target across different frequencies. The use of support vector machines and auto-encoders with granularity matching allows the use of one wireless infrastructure such as Wi-Fi for multiple tasks such as pose detection, occupancy sensing and localization on top of communication [30]. A critical challenge in the development of joint communication and sensing systems is the need to balance the trade-offs between communication and sensing capabilities. In some cases, it may be more beneficial to prioritize one over the other, depending on the specific needs and goals of the system. For example, in a transportation system, it may be more important to prioritize communication in order to improve the flow of traffic, while in a healthcare system, it may be more important to prioritize sensing in order to improve the accuracy and reliability of medical devices. With ML techniques we may robustly allocate resources for this division on the fly.

6.6 Semantic Communication

Semantic communication is the exchange of information that is understood by both the sender and the receiver in a way that is meaningful and contextually relevant. An example of this method of communication is directly mapping the pixel values to the samples transmitted over the wireless channel, which bypasses the transformation of the pixel values to a sequence of bits. The use of a multi-modal network to enable the transformation from speech to text can greatly reduce the speech representation redundancy allowing for more energy-efficient transmission. Semantic-driven adaptive encoding often uses a lower number of symbols than fixed encoding to represent each word. In addition to these technical advancements, semantic communication in 6G networks is expected to improve the user experience by providing personalized and relevant information faster through content-aware caching. We may also expect better language and gesture translation accuracy leading to an improved quality of life for hearing or speech-impaired people. Through the combination of representation learning and reasoning models, we can design computationally efficient goal-oriented solutions.

6.7 ML-Based Standardization

Standardization is essential for the smooth and efficient operation of any telecommunications network, as it ensures that different devices and systems are able to communicate with each other in a consistent and reliable manner. However, the rapid pace of technological advancement in the telecommunications industry has made it difficult for traditional standardization processes to keep up, resulting in fragmented and incompatible systems that hinder the deployment of new technologies and services. Machine learning can help to overcome these challenges by providing a more flexible and adaptable approach to standardization. Through the use of machine learning algorithms, it is possible to automatically identify common patterns and characteristics in different 6G technologies and networks, and use this information to develop standardized protocols and protocols that can be easily adopted by all stakeholders. This leads to increased transparency, better explainability and enhanced readability thus creating white boxes for fairly complex process descriptions. This is made possible by advances in natural language processing and intuitive logical reasoning-based machine learning frameworks [65].

The prime benefit of using machine learning for standardization is that it can help to reduce the time and resources required for the standardization process. Traditional standardization processes are often labour-intensive and time-consuming, requiring the involvement of multiple stakeholders and expert committees. By contrast, machine learning algorithms can quickly and accurately identify common patterns and characteristics in different 6G technologies and networks over much wider knowledge bases, enabling the development of standardized protocols in a much shorter time frame. Another advantage of using machine learning for standardization is that it can help to ensure the compatibility and interoperability of different 6G technologies and networks.

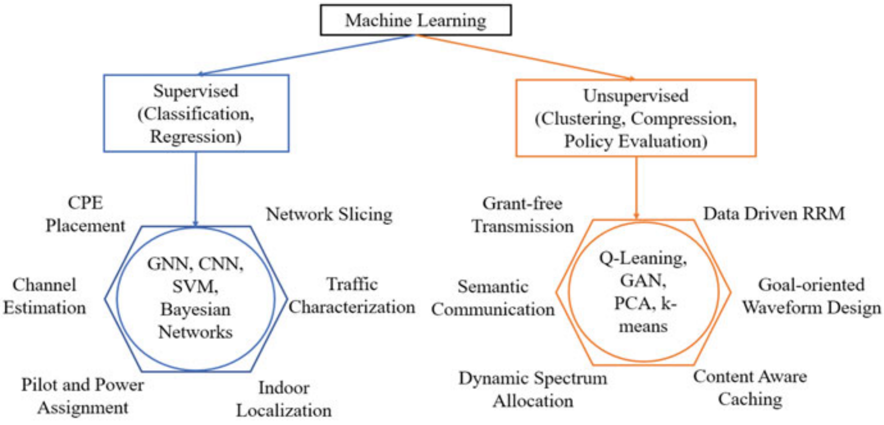


Fig. 3 An overview of machine learning techniques and their possible use cases

7 Machine Learning Frameworks for 6G

We now discuss some machine learning frameworks and discuss their suitability for different wireless deployment scenarios (Fig. 3).

7.1 Federated Leaning

Federated learning enables the training of multiple models on distributed datasets located on different devices rather than a single centralized model. This decentralized approach fits intuitively with several operational infrastructures for 6G such as distributed MIMO and sensor networks. Traditional machine-learning approaches require large amounts of data to be collected and centralized in order to train models which are not scalable. Federated learning offers a solution to this problem by enabling the training of multiple models on decentralized datasets, which can be aggregated to create a single, relevant and accurate model. Another benefit of federated learning for 6G communication is improved privacy and security. In a centralized machine learning system, sensitive data must be shared with a central server in order to be used for training. This can expose that data to potential breaches or misuse. Federated learning, on the other hand, allows data to remain on the devices where it is generated as only model updates are shared with the central server, reducing the risk of sensitive data being compromised. Furthermore, federated learning can improve the performance and efficiency of 6G communication systems. In a centralized machine learning system, the central server must receive and process all the data in order to train the model. This can be resource intensive and may require significant amounts of bandwidth and computing power. With federated learning, the training can be done on the devices themselves, which reduces the need for data transfer and can improve the speed and over-the-air spectrum usage of the learning process.

7.2 Graph Neural Networks

Graph neural networks (GNNs) are a type of neural network that operates on graph-structured data, such as social networks or networks of connected devices in a 6G communication system. GNNs are capable of learning from and making predictions on graphs by leveraging the structural information inherent in the connections between nodes, especially in problems presenting some form of permutational invariance which can be induced in the context of certain network metrics like MIMO SINR calculation. This allows them to effectively model complex relationships and patterns in the data, which is crucial for applications such as predicting traffic patterns in a 6G network or identifying important nodes in a social network. The strength of GNNs is their ability to handle node and edge features, which are additional pieces of information associated with each node and edge in the graph. This allows GNNs to incorporate diverse types of information, such as the location, type, or communication capabilities of a device in a 6G network, or the attributes of a user in a social network. Another important characteristic of GNNs is their ability to propagate information between connected nodes in the graph. This allows GNNs to effectively capture the influence of one node on another, and to make predictions based on the overall structure of the graph. GNNs can be used for a variety of tasks, such as network modelling and optimization, traffic prediction and management, and anomaly detection. For example, GNNs could be used to predict the traffic demands on different parts of the network and to optimize the allocation of resources in real time. They could also be used to identify anomalies or malicious activity in the network, such as attempts to disrupt communication or steal data.

7.3 Q-Learning

Q-learning can be used to optimize the performance of a 6G network by learning from past experiences and making decisions that maximize the long-term rewards for the system. Q-learning works by using a mathematical model called a “Q-table” to store the information that the algorithm has learned. This Q table is essentially a matrix that maps the state of the system (e.g. the current channel conditions in a 6G network) to the optimal action to take in that state (e.g. selecting a particular frequency band for transmission). The algorithm then updates the Q-table based on the rewards it receives from the environment, allowing it to continually improve its decision-making over time. Q-learning can adapt to changing conditions and environments thus improving the robustness of the system. Another advantage of

Q-learning is that it can be used to optimize a wide range of performance metrics, such as data throughput, energy efficiency, and network reliability. In a 6G network, these metrics are likely to be very important, as they will determine the quality of the user experience and the overall performance of the network. By using Q-learning, a 6G network can learn to make decisions that maximize these metrics, leading to better overall performance.

7.4 Adversarial Learning

Adversarial learning is a machine learning technique that involves training a model to recognize and classify data by learning from examples and making decisions based on patterns in the data. Adversarial learning has been used in a variety of applications, including image recognition, natural language processing, and speech recognition. With the emergence of 6G technology, adversarial learning is expected to play a critical role in enabling advanced communication and networking capabilities. Adversarial learning involves training a model using a dataset that consists of input data and corresponding labels. The model is then trained to predict the correct label for a given input data. However, adversarial learning differs from traditional machine learning techniques in that it involves the use of an additional component known as an adversarial network. This network is designed to challenge the model by generating data that is similar to the training data but with subtle differences that are difficult for the model to recognize. The adversarial network is trained to generate data that are classified incorrectly by the model, and the model is then trained to correctly classify this adversarial data.

Adversarial learning can improve the robustness of the model even in cases where sufficient training data is not available such as sparse transmission or event-driven communication. By training the model to correctly classify adversarial data, it becomes more resistant to errors and is less likely to be fooled by subtle differences in the data. This is particularly important in the context of 6G technology, where the demand for high-speed and reliable communication is expected to increase significantly. The goal of adversarial learning is to support the safe adoption of ML/DL solutions to emerging applications in the presence of adversaries. Another advantage of adversarial learning is its ability to improve the generalization ability of the model. By challenging the model with a variety of different data, it becomes better at handling and classifying new data that it has not seen before. This is particularly useful in the context of 6G, where the complexity and diversity of communication scenarios is expected to increase significantly. There are several approaches to implementing adversarial learning, including generative adversarial networks (GANs) and adversarial training. GANs involve training two neural networks, one to generate data and one to classify it. The generator network is trained to generate data that is similar to the training data, while the classifier network is trained to correctly classify the generated data. Adversarial training involves adding a small amount of adversarial data to the training dataset and training the model to correctly classify this data.

7.5 Auto-encoders and Transformers

Autoencoders are a type of neural network that is used for unsupervised learning. They learn to represent data in a lower-dimensional space, called the latent space, and then learn to reconstruct the original data from this lower-dimensional representation. Autoencoders were first introduced in the 1980s, but have gained increased attention in recent years due to their ability to perform dimensionality reduction and feature learning. Autoencoders could be used to learn the underlying structure of the communication data, and then use this information to compress the data for transmission. This could potentially lead to higher data rates and more efficient use of the spectrum, as well as improved robustness to noise and interference. Another potential application of autoencoders in 6G is in the area of network security. Autoencoders could be used to detect anomalies in network traffic, such as attempts to infiltrate the network or unauthorized access to sensitive data. By learning the normal patterns of network traffic, autoencoders can identify deviations from these patterns and alert system administrators to potential security threats.

The transformer is a sequence-to-sequence deep neural network model consisting of a multi-layered encoder and a decoder module where the input and output sequences are converted to vector embeddings. Each encoder/decoder layer has a similar structure and consists of a self-attention layer followed by a multi-layer perception. Self-attention mechanism relates different positions in a single sequence to compute a representation of the sequence, which can also be regarded as a non-local filtering operation. The transformer-based solutions can perform dimensionality reduction networks for pilot design and channel estimation [66]. In contrast to local attention, self-attention in the transformer can extract long-range correlations so that the global features of the channel matrix can be extracted for enhanced estimation accuracy.

8 Streamlining the Deployment of ML Models

The end-to-end system design workflow for 6G using ML can be described as follows:

1. Identify the business problem: The first step would be to identify the business problem that needs to be solved using 6G and ML. This could be anything from optimizing network performance to improving customer experience to reducing operational costs.
2. Gather data: Once the business problem is identified, the next step would be to gather relevant data that can be used to train ML algorithms. This data could be collected from various sources such as sensors, devices, and applications.
3. Pre-processing and cleaning: The collected data would then need to be pre-processed and cleaned to ensure that it is in the right format and free of any errors or inconsistencies.

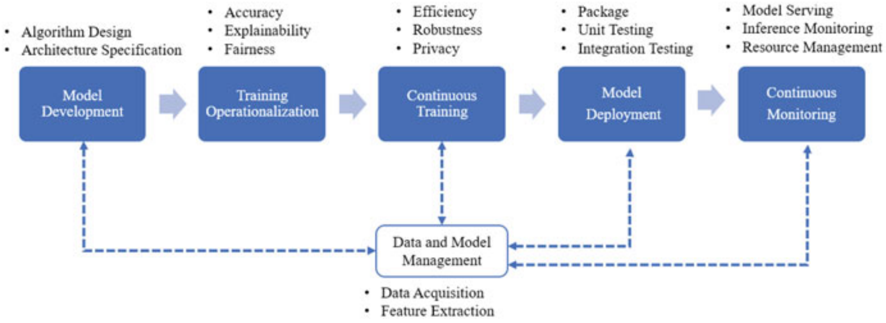


Fig. 4 The typical MLOps workflow

4. Feature engineering: The next step would be to perform feature engineering to extract relevant features from the data that can be used to train the ML algorithms.
5. Model training: After the data has been cleaned and features have been extracted, the next step would be to train ML models using the data. This could involve using various algorithms such as decision trees, random forests, or neural networks.
6. Model evaluation: Once the models have been trained, they would need to be evaluated to ensure that they are performing well and can accurately solve the business problem.
7. Deployment: If the models are performing well, they can then be deployed in the 6G network to solve the business problem.
8. Monitoring and maintenance: The final step would be to continuously monitor the performance of the models and make any necessary adjustments to ensure that they are still performing well over time.

This workflow falls under the purview of Machine Learning Operations (MLOps). It consists of a set of practices and tools that are designed to facilitate the collaboration between data scientists and IT professionals in the development and deployment of machine learning models, enforcing the use of best practices and a readily verifiable process pipeline. The goal of MLOps is to make the process of building, testing, and deploying machine learning models more efficient, reliable, and scalable. MLOps enables data scientists and IT professionals to work together more effectively. Data scientists are often responsible for developing the algorithms and models that power machine learning applications, while IT professionals are responsible for managing the infrastructure and deployment of those applications. By bringing these two groups together, MLOps helps to ensure that machine learning models are developed and deployed in a way that is consistent with the needs of the business (Fig. 4).

MLOps improves the reliability and scalability of machine learning models. When machine learning models are deployed in a production environment, they need to be able to handle large volumes of data and be able to adapt to changing business needs. MLOps ensures that models are tested and optimized for these requirements and re-trained as and when required. To effectively implement MLOps, organizations need to establish clear roles and responsibilities for data scientists and IT professionals along with following industry standards for network automation [67] closely. One of the key tools used in MLOps is version control software, which helps to track changes to machine learning models and ensure that the most current and reliable versions are being deployed. Other tools that are commonly used in MLOps include continuous integration and delivery platforms, which automate the process of building and deploying machine learning models, and machine learning platforms, which provide a consistent environment for developing and testing machine learning models.

9 Some Useful Tools and Resources

High-quality datasets are essential for wireless AI research, so it is urgent to maintain a sustainable community jointly constructing and sharing wireless AI datasets. To this effect, CAICT (China academy of information and communication technology) published the ray tracing-based wireless AI research datasets with contributions from universities and companies [68]. The dataset consists of 10k scenarios from real maps of more than 40 major cities around the world with 5 BSs and 30 UEs in each scenario. The dataset consists of:

- Configurations: frequencies, bandwidths, sampling rates, antenna patterns.
- Environmental information: channel gains, ray traces, and geographical deployment maps.

This dataset enables testing a variety of tasks such as localization, environment reconstruction, beamforming, CSI feedback, and channel estimation. Sionna is a GPU-accelerated open-source library for link-level simulations based on TensorFlow [69]. It is immensely helpful for the quick prototyping of complex neural network-aided communication system architectures. Sionna has verified implementations for the state-of-the-art algorithms that can be used for benchmarking, allowing researchers to focus on their research while saving time implementing components outside their area of expertise. The following is the GitHub repository of Prof. Emil Bjornson, professor of wireless communication at KTH Sweden [70]. The repository contains implementations for many of his works focusing on deep neural networks which provide a good example of standard practices for environment modelling and ML implementation in use cases such as IRS, CF mMIMO and grant-free random access. PyTorch Lightning is the deep learning framework for professional AI researchers and machine learning engineers who need maximal flexibility without sacrificing performance at scale [71]. It is a simple plug-and-play framework for model development that supports a variety of functions including hyperparameter tuning, bottleneck profiling and distributed execution management. Deepsig Inc. has created a small corpus of standard datasets called RadioML 2018.01A which can be used for original and reproducible research, experimentation and measurements [72]. This dataset includes both synthetic simulated channel effects and over-the-air recordings of 24 digital and analog modulation types.

References

1. Cisco (2019) Cisco visual networking index: forecast and trends, 2017–2022. Version c11
2. Shehzad MK, Rose L, Butt MM, Kovács IZ, Assaad M, Guizani M Artificial intelligence for 6G networks: Technology advancement and standardization. IEEE Veh Technol Mag 17(3):16–25 (2022)
3. Tataria H, Shafi M, Molisch AF, Dohler M, Sjöland H, Tufvesson F (2021) 6g wireless systems: Vision, requirements, challenges, insights, and opportunities. Proc IEEE 109(7):1166–1199
4. Shalev-Shwartz S, Ben-David S (2014) Understanding machine learning: from theory to algorithms. Cambridge University Press
5. OECD (2022) Measuring the value of data and data flows. https://doi.org/10.1787/2182ce9f-en
6. OECD (2022) Harnessing the power of ai and emerging technologies. https://doi.org/10.1787/923230a6-en
7. Bureau TUC (2022) E-stats 2020: Measuring the electronic economy. https://www.census.gov/library/publications/2022/econ/2020-e-stats.html
8. Mao B, Tang F, Kawamoto Y, Kato N (2022) AI models for green communications towards 6G. IEEE Commun Surv Tutorials 24(1):210–247
9. Dhingra S, Madda RB, Gandomi AH, Patan R, Daneshmand M (2019) Internet of things mobile–air pollution monitoring system (IoT-mobair). IEEE Internet Things J 6(3):5577–5584
10. Qi C, Huang S, Wang X (2020) Monitoring water quality parameters of Taihu lake based on remote sensing images and LSTM-RNN. IEEE Access 8:188068–188081
11. Matracia M, Saeed N, Kishk MA, Alouini MS (2022) Post-disaster communications: Enabling technologies, architectures, and open challenges. IEEE Open J Commun Soc 3:1177–1205
12. OECD (2022) Responding to societal challenges with data
13. David K, Berndt H (2018) 6G vision and requirements: Is there any need for beyond 5G? IEEE Veh Technol Mag 13(3):72–80
14. Yang P, Xiao Y, Xiao M, Li S (2019) 6G wireless communications: Vision and potential techniques. IEEE Netw 33(4):70–75
15. Mourad A, Yang R, Lehne PH, de la Oliva A (2020) Towards 6G: Evolution of key performance indicators and technology trends. In: 2020 2nd 6G wireless summit (6G SUMMIT), pp 1–5
16. Huawei (2019) 5G power whitepaper. https://www.huawei.com/en/news/2022/3/5g-advanced-mwc22-white-paper
17. 3GPP (2021) Study on support of reduced capability NR devices. Technical Report (TR) 38.875, 3rd Generation Partnership Project (3GPP), Version 17.0.0
18. 3GPP (2019) Study on physical layer enhancements for NR ultra-reliable and low latency case (URLLC), Technical Report (TR) 38.824, 3rd Generation Partnership Project (3GPP). Version 16.0.0
19. Popovski P, Nielsen JJ, Stefanovic C, Carvalho Ed, Strom E, Trillingsgaard KF, Bana AS, Kim DM, Kotaba R, Park J, Sorensen RB (2018) Wireless access for ultra-reliable low-latency communication: Principles and building blocks. IEEE Netw 32(2):16–23
20. Ge X (2019) Ultra-reliable low-latency communications in autonomous vehicular networks. IEEE Trans Veh Technol 68(5):5005–5016
21. Liu Y, Wang CX, Huang J (2019) Recent developments and future challenges in channel measurements and models for 5G and beyond high-speed train communication systems. IEEE Commun Mag 57(9):50–56
22. Mohammadkarimi M, Mehrabi M, Ardakani M, Jing Y (2019) Deep learning-based sphere decoding. IEEE Trans Wirel Commun 18(9):4368–4378

23. Tonnellier T, Hashemipour M, Doan N, Gross WJ, Balatsoukas-Stimming A (2021) Towards practical near-maximum-likelihood decoding of error-correcting codes: An overview. In: ICASSP 2021 - 2021 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp 8283–8287
24. Adolf R, Rama S, Reagen B, Wei GY, Brooks D (2016) Fathom: Reference workloads for modern deep learning methods. In: 2016 IEEE international symposium on workload characterization (IISWC). IEEE, pp 1–10
25. Shao Y, Liew SC, Wang T (2020) Alphaseq: Sequence discovery with deep reinforcement learning. IEEE Trans Neural Netw Learn Syst 31(9):3319–3333
26. Li P, Cui H, Khan A, Raza U, Piechocki R, Doufexi A, Farnham T (2021) Deep transfer learning for wifi localization. In: 2021 IEEE radar conference (RadarConf21), pp 1–5
27. Yang J, Zhao X, Li Z (2021) Updating radio maps without pain: An enhanced transfer learning approach. IEEE Internet Things J 8(13), 10693–10705
28. Ripley BD (2007) Pattern recognition and neural networks. Cambridge University Press
29. El-Amine A, Iturralde M, Haj Hassan HA, Nuaymi L (2019) A distributed q-learning approach for adaptive sleep modes in 5g networks. In: 2019 IEEE wireless communications and networking conference (WCNC), pp 1–6
30. Yu J, Wang P, Koike-Akino T, Wang Y, Orlik PV, Buehrer RM (2022) Multi-band wi-fi sensing with matched feature granularity. IEEE Internet Things J 9(23):23810–23825
31. Aoudia FA, Hoydis J (2021) End-to-end waveform learning through joint optimization of pulse and constellation shaping. In: 2021 IEEE Globecom Workshops (GC Wkshps). IEEE, pp 1–6
32. Aoudia FA, Hoydis J (2022) Waveform learning for next-generation wireless communication systems. IEEE Trans Commun 70(6):3804–3817
33. Liu X, Xu Y, Jia L, Wu Q, Anpalagan A (2018) Anti-jamming communications using spectrum waterfall: A deep reinforcement learning approach. IEEE Commun Lett 22(5):998–1001
34. Zhang J, Pan P, Yang LL, Maunder RG (2020) Channel estimation and user activity identification in massive grant-free multiple-access. IEEE Open J Veh Technol 1:296–316
35. Maeurer C, Futter P, Gampala G (2020) Antenna design exploration and optimization using machine learning. In: 2020 14th European conference on antennas and propagation (EuCAP), pp 1–5
36. Huang C, Alexandropoulos GC, Yuen C, Debbah M (2019) Indoor signal focusing with deep learning designed reconfigurable intelligent surfaces. In: 2019 IEEE 20th international workshop on signal processing advances in wireless communications (SPAWC), pp 1–5
37. Kim H, Kim S, Lee H, Jang C, Choi Y, Choi J (2021) Massive MIMO channel prediction: Kalman filtering vs. machine learning. IEEE Trans Commun 69(1):518–528
38. Yuan J, Ngo HQ, Matthaiou M (2020) Machine learning-based channel prediction in massive MIMO with channel aging. IEEE Trans Wirel Commun 19(5):2960–2973
39. Hu Q, Gao F, Zhang H, Jin S, Li GY (2021) Deep learning for channel estimation: Interpretation, performance, and comparison. IEEE Trans Wirel Commun 20(4):2398–2412
40. Shehzad MK, Rose L, Assaad M (2021) Dealing with CSI compression to reduce losses and overhead: An artificial intelligence approach. In: 2021 IEEE international conference on communications workshops (ICC Workshops), pp 1–6
41. Shehzad MK, Rose L, Wesemann S, Assaad M (2022) ML-based massive MIMO channel prediction: Does it work on real-world data? IEEE Wirel Commun Lett 11(4):811–815
42. Cover TM, Thomas JA (2006) Elements of information theory, 2nd edn., Wiley series in telecommunications and signal processing. Wiley-Interscience
43. O'Shea T, Hoydis J (2017) An introduction to deep learning for the physical layer. IEEE Trans Cogn Commun Netw 3(4):563–575
44. O'Shea TJ, Karra K, Clancy TC (2016) Learning to communicate: Channel auto-encoders, domain specific regularizers, and attention. In: 2016 IEEE international symposium on signal processing and information technology (ISSPIT), pp 223–228
45. Dai L, Sun X (2013) A unified analysis of ieee 802.11 DCF networks: Stability, throughput, and delay. IEEE Trans Mob Comput 12(8):1558–1572

46. Guo Z, Chen Z, Liu P, Luo J, Yang X, Sun X (2022) Multi-agent reinforcement learning-based distributed channel access for next generation wireless networks. IEEE J Sel Areas Commun 40(5):1587–1599
47. Choi J, Ding J, Le NP, Ding Z (2022) Grant-free random access in machine-type communication: Approaches and challenges. IEEE Wirel Commun 29(1):151–158
48. Ahmed KI, Tabassum H, Hossain E (2019) Deep learning for radio resource allocation in multi-cell networks. IEEE Netw 33(6):188–195
49. Foukas X, Patounas G, Elmokashfi A, Marina MK (2017) Network slicing in 5G: Survey and challenges. IEEE Commun Mag 55(5):94–100
50. Wang H, Wu Y, Min G, Miao W (2022) A graph neural network-based digital twin for network slicing management. IEEE Trans Ind Inf 18(2):1367–1376
51. Shi Y, Davaslioglu K, Sagduyu YE (2020) Generative adversarial network in the air: Deep adversarial learning for wireless signal spoofing. IEEE Trans Cogn Commun Netw 7(1):294–303
52. Li D, Xu J, Yang Z, Lu Y, Zhang Q, Zhang X (2021) Train once, locate anytime for anyone: Adversarial learning based wireless localization. In: IEEE INFOCOM 2021-IEEE conference on computer communications. IEEE, pp 1–10
53. Goutay M, Aoudia FA, Hoydis J, Gorce JM (2021) Machine learning for MU-MIMO receive processing in OFDM systems. IEEE J Sel Areas Commun 39(8):2318–2332
54. Liu J, Mei K, Zhang X, Ma D, Wei J (2019) Online extreme learning machine-based channel estimation and equalization for OFDM systems. IEEE Commun Lett 23(7):1276–1279
55. Kim M, Lee W, Cho DH (2018) A novel PAPR reduction scheme for OFDM system based on deep learning. IEEE Commun Lett 22(3):510–513
56. Ngo HQ, Ashikhmin A, Yang H, Larsson EG, Marzetta TL (2017) Cell-free massive MIMO versus small cells. IEEE Trans Wirel Commun 16(3):1834–1850
57. Salaün L, Yang H (2021) Deep learning based power control for cell-free massive MIMO with MRT. In: 2021 IEEE global communications conference (GLOBECOM), pp 01–07
58. Song N, Yang T (2021) Machine learning enhanced CSI acquisition and training strategy for FDD massive MIMO. In: 2021 IEEE wireless communications and networking conference (WCNC), pp 1–6
59. Mishra S, Salaun L, Gorce JMS, Chen CS (2022) Maximizing downlink user connection density in NOMA-aided NB-IoT networks through a graph matching approach. In: VTC2022-Fall-IEEE 96th vehicular technology conference
60. Rappaport TS, Xing Y, Kanhere O, Ju S, Madanayake A, Mandal S, Alkhateeb A, Trichopoulos GC (2019) Wireless communications and applications above 100 ghz: Opportunities and challenges for 6G and beyond. IEEE Access 7:78729–78757
61. Cheema MA, Shehzad MK, Qureshi HK, Hassan SA, Jung H (2021) A drone-aided blockchain-based smart vehicular network. IEEE Trans Intell Transp Syst 22(7):4160–4170
62. Garcia MHC, Molina-Galan A, Boban M, Gozalvez J, Coll-Perales B, Şahin T, Kousaridas A (2021) A tutorial on 5G NR V2X communications. IEEE Commun Surv Tutorials 23(3):1972–2026
63. Mishra D, Vegni AM, Loscrí V, Natalizio E (2021) Drone networking in the 6G era: A technology overview. IEEE Commun Stand Mag 5(4):88–95
64. Wild T, Braun V, Viswanathan H (2021) Joint design of communication and sensing for beyond 5G and 6G systems. IEEE Access 9:30845–30857
65. Davies A, Veličković P, Buesing L, Blackwell S, Zheng D, Tomašev N, Tanburn R, Battaglia P, Blundell C, Juhász A, et al (2021) Advancing mathematics by guiding human intuition with AI. Nature 600(7887):70–74
66. Wang Y, Gao Z, Zheng D, Chen S, Gunduz D, Poor HV (2022) Transformer-empowered 6G intelligent networks: From massive MIMO processing to semantic communication. IEEE Wirel Commun 1:1–9
67. 3GPP (2019) Study of enablers for network automation for 5G, Technical Report (TR) 23.791, 3rd Generation Partnership Project (3GPP), June 2019. Version 16.20
68. CAICT (2022) Mobile communication open dataset. https://www.mobileai-dataset.com/

69. Hoydis J, Cammerer S, Aoudia FA, Vem A, Binder N, Marcus G, Keller A (2022) Sionna: An open-source library for next-generation physical layer research. https://github.com/nvlabs/sionna
70. Björnson E, Hoydis J, Sanguinetti L (2017) Massive MIMO networks: spectral, energy, and hardware efficiency. Found Trends Signal Process 11(3–4):154–655. http://dx.doi.org/10.1561/2000000093
71. Falcon W (2019) Pytorch lightning. Version=1.4
72. O'Shea TJ, Roy T, Clancy TC (2018) Over-the-air deep learning based radio signal classification. IEEE J Sel Top Signal Process 12(1):168–179. https://doi.org/10.1109/JSTSP.2018.2797022

Cyber Security in 6G Networks

Valeria Loscri

1 Transitioning of Security Aspects from 5G to 6G

In this section we will try to capture the more characteristic transitioning points from 5G to 6G technology, with tracing the important implications from users and applications point of view. Based on this analysis, the evolution of the security threats that this technological changes imply will be analyzed. A simplified representation of this transitioning is provided as reference in Fig. 1. It can be seen that the requirements of the new applications envisaged in 6G systems are very strict with very high data rate, very low latency and localization precision in the order of cm, by considering Non-Terrestrial Dimension (NTD), that make the development of solutions very challenging.

1.1 A Glimpse to the Differences Between 5G and 6G Technologies

The transition from 5G networks to 6G networks is characterized by key driving factors, that need to be carefully considered to better understand the unique features of 6G systems [1, 2]. An overall evolution of the different aspects of different generations is shown in Fig. 2. It is clear how the massive use of AI and ML approaches will permit from one side to achieve the high performance required by the 6G networks, but create some important security issues, by enabling evolved types of attacks or new kinds attacks that were not possible before.

V. Loscri (✉)
Inria research centre Lille - Nord Europ, Villeneuve d'Ascq, France
e-mail: valeria.loscri@inria.fr

V. Loscri et al. (eds.), *The Road towards 6G: Opportunities, Challenges, and Applications*, https://doi.org/10.1007/978-3-031-42567-7_6

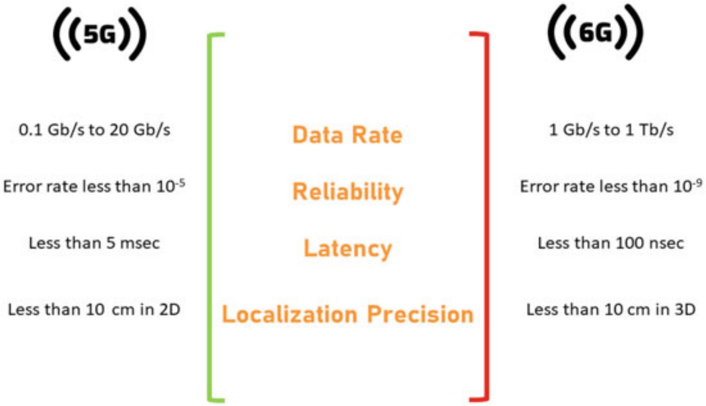


Fig. 1 Requirements transitioning from 5G to 6G

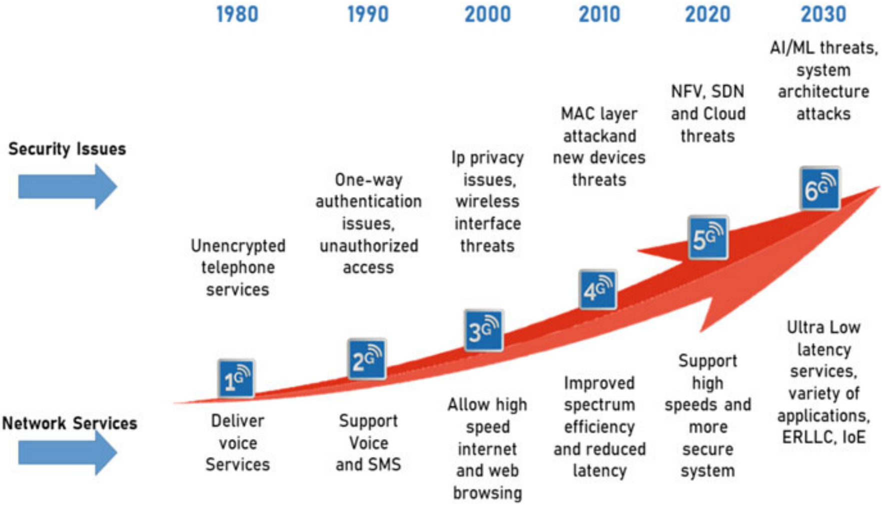


Fig. 2 The evolution of specific features of the different networks generations

A first important aspect is related to the need to meeting higher user experience in terms of higher data rates, connection density, reduced latency, in an increased connection density. The connected objects will be with massive density and with seamless coverage across space, air, ground and sea. These characteristics need to be integrated from an architectural point of view and the new types of 6G platforms considered have to meet cost-effective requirements [3]. The specific cost considered cannot be only confined to the energy consumption during the usage of the platform, but has to represent a more global vision of the cost, encompassing the creation of the equipment, the energy/cost impact during their usage and the cost associated to the dismiss the various components.

By considering the most advanced services that the 6G technology will be able to realize, it is easy to figure out that the operation and maintenance of the new architectures will be also more complex. It is for these reasons that these aspects are crucial and need to be considered and deeply analyzed from the beginning.

In the following, we discuss the key differences between 5G and 6G, to better analyze the security and privacy aspects.

The fifth generation of cellular network technology has been the successor of the 4G Long Term Evolution (LTE), but differently from the previous technologies, it was not a simple incremental technological advancement but a true bounce. One key aspect of 5G, in contrast to previous generations, is the exploitation of higher frequency bands (i.e., mmWaves), enabling higher capacity and data rates.

This enhanced technology makes in sort that evolved uses cases can be considered in the context of Smart City applications, Internet of Things, Health-care, etc.

The sixth generation 6G of cellular network technology is expected to be available by 2030 and will rely on higher frequency bands up to 300–3000 GHz. One key question is, what does the usage of higher frequency imply and which are the main differences between 5G and 6G? Roughly speaking, the utilization of higher frequency bands is the key difference between the two technologies, with tremendous impact on the scarcity spectrum challenges that 5G is experiencing, with unprecedented speeds, theoretically 100 times faster than 5G. These speeds have the potential to enable instantaneous connections through different digital devices, spanning from tablets, smartphones, wearable and very tiny devices, with an effect of producing a completely connected cyber ecosystem.

1.2 Security Changes from 5G and 6G

Anyway, all this big evolution does not come without important side effects in terms of potential security threats. Firstly, 6G systems have a larger attack surface than 5G networks, with applications and services in cyber critical infrastructures.

Overall, 6G network is based on massive data processing, much more than in 5G networks. This feature triggers increased security issues, that need to be faced by introducing decentralized security systems, in order to handle the data locally and in a dynamic way. This assertion is more evident when high mobility is considered. Indeed, devices may change their interconnected networks, and services that sense functionality, can be required to other networks, making security and privacy more complicated to be handled [4].

It is expected that future generation networks will strongly rely on softwarization, as abstraction of the underlying hardware and physical parts to realize more flexible and innovative solutions, where the main functionalities are mainly based on software [5, 6]. Softwarization will be combined with more dynamic resources allocation, through advanced Artificial Intelligence (AI) and Machine Learning (ML) approaches [7]. Softwarized architectures, which will characterize the advent of 6G will be capable to perform automatic and seamless selection of the most suitable model to meet the target performance of the underlying application.

Two important characteristics of 6G turn on in potential threats:

- the increased relying on open source software
- the massive use of Artificial Intelligence (AI) and Machine Learning (ML).

Both these aspects will be considered and analyzed in the following sections, from an offender point of view and a detection and prevention point of view.

2 Cyber-Security Vulnerabilities, New Risks and Threats in 6G

One of the recurrent concepts in 6G networks is the "openness", that will be translated in softer separation lines between inside and outside network. This aspect has an enormous impact on security, since actual solutions are not suitable anymore to protecting the network. In this new security landscape vision, concept of Zero Trust (ZT), where pitfalls can lurk inside the network and the system is accessible to outside untrustworthy users, should be considered as a must. Based on that, the underlying 6G architecture has to be build around this ZT paradigm and concepts as collaborative and non-cooperative Zero Trust Architectures are recently arising [8, 9].

2.1 Open Source Software: The New Evil for 6G Security?

The 6G network is envisioned to be more open and characterized with an unprecedented level of heterogeneity than previous generations. he heterogeneity stems from different domains, including e.g., the set of devices, their features and resources, the adopted communications interfaces and technologies, etc. The sixth generation of networks is expected to seamlessly integrate terrestrial, aerial, satellite and underwater networks through different radio access technologies. A higher level of intelligence is demanded to support the high heterogeneity, above all at the edge layer. A paradigm shift from the Internet of Things to the Internet of Intelligence is occurring in the transition from the 5G to the 6G networks. This results in an increased challenge for the conventional security architectures, where network boundaries are often adopted to develop security perimeter [10].

In a recent white paper on Security and Trust for the 6G era, NOKIA Bell Labs considers the crucial role played by automated software creation to address the increasing threat of open source software. Indeed, it is objectively recognized that one of the most important breaches in terms of security in communication systems is represented by vulnerable software and this is worst for the open source software paradigm shift adopted in 6G. In this context, the support of AI-ML approaches for automated software, is envisaged as a reliable way to provide insights on the code in the different development and integration phases [11]. Coupled with the open source software as a viable solution to support the flexibility required by the 6G architectures, comes the concept of *liquid software*.

The concept of *liquid software* is related to the possibility to dynamically move functionality in a network, among its different components and actors, between hosts at the edge and nodes inside a network. In this revisited version, code becomes *mobile* and it is not embedded in a specific machine. A *mobile code* can be moved from a machine to another and is often referred as *liquid software* to signify its inherent capability to flow from one part of the other in the network, based on the specific requirements. In [12], the authors consider the concept of liquid-software to enable a software-defined network, integrating AI applications. Even though this is a very interesting perspective, as outlined in a white paper "6G White Paper on Edge Intelligence" (2020), this paradigm shift is still far from being realized in actual networks and "Without liquid software as part of future 6G networks, we are stuck with an approach in which we must decide where to locate the intelligence in the network topology at the due design time, because the computations cannot be easily relocated without design-time parations".

2.2 The Double Nature of AI/ML in 6G Security

As a matter of fact, 6G networks will be characterized with embedding intelligence to meet the increased dynamics of these systems. In this sense, intelligent security functions will be programmed to monitor gateways traffic by the means of machine learning algorithms that have increased capacity to detect, contain, prevent and mitigate new or evolved threats and cyber attacks [13–15]. AI/ML mechanisms will not only be a key security technology enabler, but can be also used to create more sophisticated threats against the same AI/ML mechanisms within the 6G architecture [16]. The unprecedented high data rate and reduced latency that can be achieved with the new technologies considered as enabler for the 6G networks, the exploitation of AI/ML techniques is music to the offenders' ears. Indeed, existing attacks can be much improved through the use of advanced machine learning approaches and new types of attacks can be realized, that were simply not possible without the exploitation of intelligence in the networks. This double nature of AI/ML has been largely discussed in [16], where a taxonomy classifying the types of attacks in relation with the types of ML and vice-versa has been considered as summarized in Fig. 3. Different practical applications have been demonstrated in [17, 18].

In [19], a dynamic selection of the potential attack in a next generation network based on intelligent approaches is demonstrated as shown in Fig. 4.

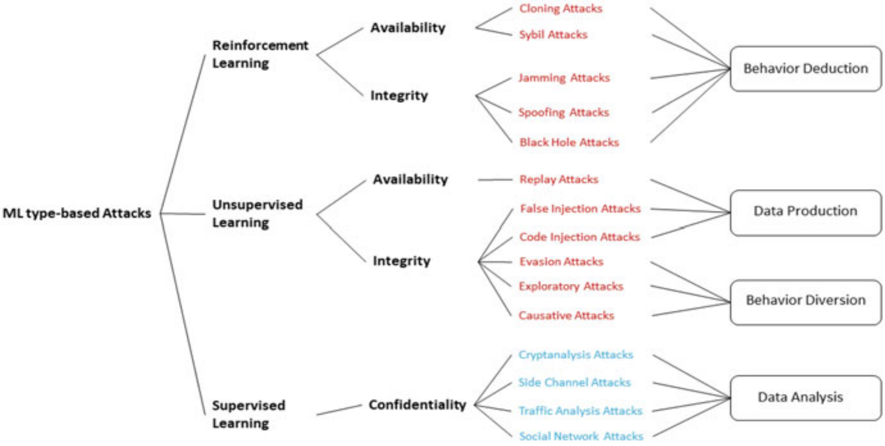


Fig. 3 ML exploitation for Smart Attack generation [16]. Blue text is for Passive attacks, while Red text is for Active attacks

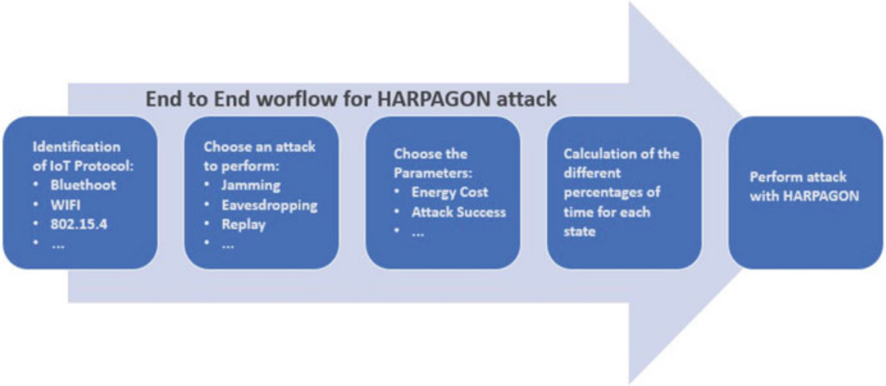


Fig. 4 HARPAGON system flow [19]

3 Physical Layer Security—PLS

Physical Layer Security (PLS) plays a first role in the ultra-dense heterogeneous 6G networks, where a huge amount of (sensitive) data conveys [20]. PLS has been often neglected in the deployment of 5G networks, while its entropic richness offers unique and still unexplored potentiality to enable secrecy capacity and make the propagation links more secure, with low cost solutions. Due to the crucial role that physical layer will play in 6G networks, by supporting lower latencies, higher capacity with lower associated energy consumption, security at this layer cannot be disregarded.

The guarantee of an uniform security level in 6G networks, where the level of unprecedented heterogeneity not only in terms of devices and communication technologies, but also in terms of environment (land, water, air, etc.), cannot be considered without considering PLS.

In the following section we will provide a detailed description of the different PLS approaches categories, following the categorization provided in [20]. In particular, five categories are considered as follows: *Secrecy Capacity*, *Physical Authentication*, *Reconfigurable Intellogent Surfaces (IRS) and Cell-free massive MIMO*, *Spectrum Spreading* and *Cooperation*.

3.1 PLS Categories

In this section we will provide a detailed description of the different PLS categories previously identified and we will discuss the different techniques to improve these solutions in 6G networks.

3.1.1 Secrecy Capacity

Secrecy capacity is the maximum communication rate at which two legitimate nodes can communicate in a secure way, and an eavesdropper is not able to infer key information. Generally, it can be achieved with a better Signal-to-Noise Ratio (SNR) for the two legitimate nodes in respect of the channel's eavesdropper. The recent advent of Reconfigurable Intelligent Surfaces (RISs) or Intelligent Reflecting Surfaces (IRS) have permitted to improve the spectral efficiency of future generation networks [21–23]. It is then straightforward to imagine IRS integrated in 6G networks, to improving their secrecy capacity. An example of application of RIS to improve the security in 6G wireless communication systems has been provided in [24]. In [25], a novel Transmit Antenna Selection (TAS) has been considered to improve the secrecy rate of a mobile 6G system, and a novel Secrecy Outage probability (SOP) has been combined with a Convolutional Neural Network (CNN) to achieve an high accuracy SOP prediction. In [26] authors consider Non-orthogonal multiple access (NOMA) combined with Multiple Input Multiple Output (MIMO) uplink network, in order to maximize the secrecy sum rate. The NOMA technology is characterized by a high spectral efficiency, which is realized by superposing multiple users' signals in the power domain. Consequently, this feature makes it an interesting candidate for 6G systems.

3.1.2 Physical Authentication

Physical authentication relies on the reciprocity property of the wireless link of two legitimate devices. A (cipher) key can be shared by these two nodes, based on the channel analysis. A detailed literature of recent works on authentication has been provided in [27]. The authors in [28] have experimentally demonstrated the key extraction, through USRP devices in MIMO-OFDM systems. In particular, they rely on Channel State Information (CSI) from MIMO-OFDM subcarriers to generate the secret key. In [29] the diversity of multi-antenna has been applied to generate an authentication key. The authors in [30], provide an interesting discussion of the authentication techniques at physical layer. An interesting perspective provided by the authors is that physical layer authentication can be improved through cross-layer approaches, based on a composite security key (CSK), integrating in a seamless way physical and upper layers schemes of authentication. The physical layer authentication is advocated as a key technology to guarantee security in high heterogeneous and complex networks. In [31], the uncertainty of the channel at the transmitter (CSIT) is considered and the different sources if uncertainty are discussed. In particular, CSIT due to an estimation error, due to limited capacity or outdated CSI are considered.

3.1.3 IRS and Cell-Free Massive MIMO

In 5G networks, the most succeeding technology was undoubtedly massive MIMO, supporting a very large number of antennas. In 6G networks the technologies that are gaining momentum are Reconfigurable Intelligent Surfaces (RIS) and Cell free massive MIMO. RIS technology allows the dynamic and adaptable change of phase, frequency and amplitude of incident signals. It is normally coupled to very high frequency technologies, such as mmWave and THz, since for these technologies the control of the signal and its directionality is most intriguing and comes with an enormous potential. Indeed, at mmWave and THz frequencies, the propagation is highly impacted by the high penetration and obstacles loss. One of the main application of RIS is beamforming, to improve the connectivity and coverage of a communication system. Beamforming is a technique consisting in steering of the transmitted signal towards a specific direction. Recently it has been proposed as an effective PLS approach, since the directionality of the signal characterizing this method may provide an increased security towards potential intruder and make the offender’ life more complicated to follow the directional beams. In this scenario, in fact, an offender is required to track the directions beams, a task that is more challenging to achieve compared to legacy networks. A systematic review of beamforming as key technique for PLS in advanced wireless networks is provided in [32]. It is highlighted as the IRS has a very high potential, by playing on the signals with constructively beamforming and in correspondence of the eavesdropper with destructively beamforming [33]. Anyway, this approach is quite complex and it requires the use of sophisticated techniques. Concerning Cell-free massive MIMO, it has been introduced with the scope to keep to keep the same advantages of massive MIMO technology, but with the intent to eliminate its main drawback. Indeed, (conventional) massive MIMO can be characterized with large variations of channel, due to the different distances of the users. The principle of Cell-free Massive MIMO is to deploy more antenna, widely distributed, to diminish this “distance” effect. In [34], the authors demonstrate that Cell-free massive MIMO is as much robust as conventional massive MIMO to passive eavesdropper. In [35] show that Cell-free massive MIMO is more sensitive to pilot contamination attacks, that can have an important detrimental impact on the network.

PLS is at an early stage, both for RIS and Cell-free massive MIMO, and a deep analysis of the physical layer features of these technologies with a critical perspective in terms of physical security should be done.

3.1.4 Spectrum Spreading

Spectrum spreading is widely used in wireless communication, due to recognized advantages, like reduced crosstalk interference and anti-jamming. Initially adopted in military contexts, it is currently used in different wireless standards. The two more recognized spread spectrum techniques are: (1) Frequency Hopping Spread Spectrum (FHSS) and (2) Direct Sequence Spread Spectrum (DSSS). In this domain, there are different techniques to dynamically change the frequency of the legitimate transmission between a transmitter and a receiver. For example, it can be based on a pre-shared sequence [36] or some channel-dependent technique can be implemented [37]. These approaches have been shown more secure, but just for basic jamming attacks. When the jammer becomes smarter, this type of approach has been shown not sufficient. Consequently, more advanced frequency hopping techniques have been developed, relying on Machine Learning approaches as demonstrated in [38]. However, ML technology can be used also to create advanced attacks and smart jammers, able to follow the dynamic selection of the frequency between a transmitter and a receiver [39, 40].

3.1.5 Cooperation

Cooperation among trust nodes is a concept that started to arise interest in 5G networks. In particular, nodes can collaborate to each other, by acting, for example, as friendly jammer. This approach consists in creating an artificial noise with the aim to deteriorating the signal of the jammer node [41, 42]. Even though this technique is quite intriguing, there are some important issues, such as the need to know the exact position of the jammer, the difficult to make the identification of the jammer. Generally, the generation of noisy signal from the cooperative node has important side effects in terms of energy consumption. Moreover, the deterioration of the link quality is often impacting the legitimate transmission as well. Unfortunately, jammers are refining their approach and make use of ML techniques to enhance the effectiveness of their attacks [39, 40]. The concept of cooperation is also applied to complacent jammers, that are able to perform more sophisticated attacks with an increased difficult to localize the jammers deployed in a distributed manner.

4 Discussion

The 6G network comes with very intriguing and breakthrough technologies, developing unprecedented applications and services. With Internet of Things (IoT) paradigm, we started to live an unprecedented pervasiveness of technology and interconnected objects in our daily life, but this is nothing compared with the ubiquitous of 6G networks. Anyway, this open the door to new challenges in terms of cyber security, with cyber attackers constantly at work to improve their performance, with new and more powerful attacks. The design of effective countermeasures cannot be realized without a thorough knowledge of the new and complex technologies enabling 6G systems realization.

References

1. Saad W, Bennis M, Chen M (2020) A vision of 6G wireless systems: applications, trends, technologies, and open research problems. IEEE Netw 34(3):134–142. https://doi.org/10.1109/MNET.001.1900287
2. Nawaz SJ, Sharma SK, Wyne S, Patwary MN, Asaduzzaman M (2019) Quantum machine learning for 6G communication networks: state-of-the-art and vision for the future. IEEE Access 7:46317–46350
3. Zhang Z, et al (2019) 6G wireless networks: vision, requirements, architecture, and key technologies. IEEE Veh Technol Mag 14(3):28–41
4. Letaief KB, Chen W, Shi Y, Zhang J, Zhang YA (2019) The roadmap to 6G: AI empowered wireless networks. IEEE Commun Mag 57(8):84–90
5. Yang P, Xiao Y, Xiao M, Li S (2019) 6G wireless communications: vision and potential techniques. IEEE Netw 33(4):70–75
6. Rappaport TS, et al (2019) Wireless communications and applications above 100 GHz: opportunities and challenges for 6G and beyond. IEEE Access 7:78729–78757
7. Zhang L, Liang Y, Niyato D (2019) 6G visions: mobile ultra-broadband, super internet-of-things, and artificial intelligence. China Commun 16(8):1–14
8. Sedjelmaci H, Ansari N (2023) Zero trust architecture empowered attack detection framework to secure 6G edge computing. IEEE Netw. https://doi.org/10.1109/MNET.131.2200513
9. Ramezanpour K, Jagannath J (2022) Intelligent zero trust architecture for 5G/6G networks: Principles, challenges, and the role of machine learning in the context of O-RAN. Comput Netw 217:109358, ISSN 1389-1286. https://doi.org/10.1016/j.comnet.2022.109358
10. Nguyen VL, Lin PC, Cheng BC, Hwang RH, Lin YD (2021) Security and privacy for 6G: a survey on prospective technologies and challenges. IEEE Commun Surv Tutorials 23(4):2384–2428. https://doi.org/10.1109/COMST.2021.3108618
11. Hashima S, et al (2022) On softwarization of intelligence in 6G networks for ultra-fast optimal policy selection: challenges and opportunities. IEEE Netw. https://doi.org/10.1109/MNET.103.2100587
12. Yang T, Qin M, Cheng N, Xu W, Zhao L (2022) Liquid software-based edge intelligence for future 6G networks. IEEE Netw 36(1):69–75. https://doi.org/10.1109/MNET.011.2000654
13. Ylianttila M, Kantola R, Gurtov A, Mucchi L, Oppermann I (eds) (2020) 6G White paper: research challenges for trust, security and privacy. 6G Flagship, University of Oulu. http://urn.fi/urn:isbn:9789526226804
14. Zhang P, Wang C, Jiang C, Benslimane A (2020) Security-aware virtual network embedding algorithm based on reinforcement learning. IEEE Trans Netw Sci Eng 8(2):1095–1105

15. Benaddi H, IIbrahimi K, Benslimane A, Jouhari M, Qadir J (2022) Robust enhancement of intrusion detection systems using deep reinforcement learning and stochastic game. IEEE Trans Veh Technol (IEEE TVT) 71(10):11089–11102. https://doi.org/10.1109/TVT.2022.3186834
16. Bout E, Loscri V, Gallais A (2022) How machine learning changes the nature of cyberattacks on IoT networks: a survey. IEEE Commun Surv Tutorials 24(1):248–279. https://doi.org/10.1109/COMST.2021.3127267
17. Bout E, Bout V, Brighente A, Conti M, Loscri V (2023) Evaluation of channel hopping strategies against smart jamming attacks. In: IEEE ICC 2023 - IEEE international conference on communications, IEEE, May 2023, Rome, Italy. hal-03950904
18. Bout E, Loscri V, Gallais A (2022) Evolution of IoT security: the era of smart attacks. IEEE Internet Things Mag (2022). hal-03610715. https://doi.org/10.1109/iotm.001.2100183
19. Bout E, Loscri V, Gallais A (2022) HARPAGON: An energy management framework for attacks in IoT networks. IEEE Internet Things J. hal-03658197. https://doi.org/10.1109/jiot.2022.3172849
20. Mucchi L, Jayousi S, Caputo S, Panayirci E, Shahabuddin S, Bechtold J, Morales I, Stoica RA, Abreu G, Haas H (2021) Physical-layer security in 6G networks. IEEE Open J Commun Soc 2:1901–1914. https://doi.org/10.1109/OJCOMS.2021.3103735
21. Cui TJ, Qi MQ, Wan X, Zhao J, Cheng Q (2014) Coding metamaterials, digital metamaterials and programmable metamaterials. Light Sci Appl 3:e218
22. Wu Q, Zhang R (2020) Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless networks. IEEE Commun Mag 58:106–112
23. Loscrí V, Rizza C, Benslimane A, Vegni AM, Innocenti E, et al (2023) BEST-RIM: a mmWave beam steering approach based on computer vision-enhanced reconfigurable intelligent metasurfaces. IEEE Trans Veh Technol 72(6):7613–7626. https://doi.org/10.1109/TVT.2023.3243358
24. Youn J, Son W, Jung BC (2021) Physical-layer security improvement with reconfigurable intelligent surfaces for 6G wireless communication systems. Sensors (Basel). 21(4):1439. https://doi.org/10.3390/s21041439. PMID: 33669520; PMCID: PMC7922280
25. Xu L, Zhou X, Tao Y, Yu X, Yu M, Khan F (2022) AF Relaying secrecy performance prediction for 6G mobile communication networks in industry 5.0. IEEE Trans Ind Inf 18(8):5485–5493. https://doi.org/10.1109/TII.2021.3120511
26. Dursun Y, Wang K, Ding Z (2022) Secrecy sum rate maximization for a MIMO-NOMA uplink transmission in 6G networks. Phys Commun 53:101675. ISSN 1874-4907. https://doi.org/10.1016/j.phycom.2022.101675
27. Hamamreh JM, Furqan HM, Arslan H (2019) Classifications and applications of physical layer security techniques for confidentiality: a comprehensive survey. IEEE Commun Surv Tutorials 21(2):1773–1828. https://doi.org/10.1109/COMST.2018.2878035
28. Cheng W, Xu A, Jiang Y, Wen H, Song H, Ouyang K, Zhu X (2017) The realization of key extraction based on USRP and OFDM channel response. In: IEEE Conf. on Commun. and Netw. Security (CNS), Oct 2017, pp 374–375
29. Zeng K (2010) Exploiting multiple-antenna diversity for shared secret key generation in wireless networks. In: IEEE INFOCOM
30. Wang X, Hao P, Hanzo L (2016) Physical-layer authentication for wireless security enhancement: current challenges and future developments. IEEE Commun Mag 54(6):152–158
31. Hyadi A, Rezki Z, Alouini M (2016) An overview of physical layer security in wireless communication systems with CSIT uncertainty. IEEE Access 4:6121–6132
32. Sharma H, Kumar N, Tekchandani R (2022) Physical layer security using beamforming techniques for 5G and beyond networks: A systematic review. Phys Commun 54:101791, ISSN 1874-4907. https://doi.org/10.1016/j.phycom.2022.101791
33. Cui M, Zhang G, Zhang R (2019) Secure wireless communication via intelligent reflecting surface. IEEE Wirel Commun Lett 8(5):1410–1414
34. Kapetanovic D, Zheng G, Rusek F (2015) Physical layer security for massive MIMO: An overview on passive eavesdropping and active attacks. IEEE Commun Mag 53:21–27

35. Timilsina S, Kudathanthirige D, Amarasuriya G (2018) Physical layer security in cell-free massive MIMO. In: IEEE global communications conference (GLOBECOM), Abu Dhabi, UAE, Dec 2018
36. Hao L, Li T, Ling Q (2007) A highly efficient secure communication interface: Collision-free frequency hopping (CFFH). In: Proc. Work- shop Signal Process. Appl. Public Secur. Forensics, pp 1–4
37. Wang Q, Zhang H, Lyu Q, Wang X, Bao J (2018) A novel physical channel characteristics-based channel hopping scheme for jamming- resistant in wireless communication. Int J Netw Secur 20(3):439–446
38. Van Huynh N, Nguyen DN, Hoang DT, Dutkiewicz E (2019) Jam me if you can: Defeating jammer with deep dueling neural network architecture and ambient backscattering augmented communication. IEEE J Sel Areas Commun 37(11):2603–2620
39. Bout E, Bout V, Brighente A, Conti M, Loscri V (2023) Evaluation of channel hopping strategies against smart jamming attacks. In: IEEE ICC 2023 - IEEE international conference on communications, IEEE, May 2023, Rome, Italy. hal-03950904
40. Bout E, Brighente A, Conti M, Loscri V (2022) FOLPETTI: a novel multi-armed bandit smart attack for wireless networks. In Proceedings of the 17th international conference on availability, reliability and security (ARES ’22). Association for Computing Machinery, New York, NY, USA, Article 21, pp 1–10. https://doi.org/10.1145/3538969.3539001
41. Liu Y, Chen HH, Wang L (2017) Physical layer security for next generation wireless networks: Theories, technologies, and challenges. IEEE Commun Surv Tutorials 19(1):347–376
42. Wu F, Zhang R, Yang LL, Wang W (2016) Transmitter precodingaided spatial modulation for secrecy communications. IEEE Trans Veh Technol 65(1):467–471

6G EMF Exposure

Radiofrequency Electromagnetic Field (RF-EMF) Research, Exposure Limits and Compliance Standards Relevant to 6G

Jack T. Rowley

1 What Is New with 6G from an EMF Point of View

The performance targets for 6G are driven by a vision of the future with Tbs data rates, microsecond latency and increased connection density that cannot be delivered by 5G systems. Some suggested indicators for 6G (relative to 5G) are summarized in Table 1.

The full set features of 6G are yet to be formalized with initiatives underway in various groups, including by the International Telecommunications Union (ITU) where it is known as IMT-2030. A June 2022 ITU[1] workshop heard that 6G is expected to use combinations of bands: below 6 GHz for wide area coverage; millimeter wave (mmWave) frequencies where wide bandwidths are available; and terahertz (THz) frequencies above 300 GHz, but there are technical challenges with the latter.

For the purpose of understanding the implications for 6G radiofrequency EMF (RF-EMF) exposure, important factors relate to the physical implementation of the radio technology. These include:

- Frequency – are the proposed frequencies covered by existing exposure limit and compliance assessment standards?
- Power – what will the transmit power levels mean for exposure levels from devices and networks?

---

[1] Summary of ITU-R Workshop on "IMT for 2030 and beyond" (aka "6G"). *IEEE ComSoc Technology Blog.* 20 June 2022. Available here: https://techblog.comsoc.org/2022/06/20/summary-of-itu-r-workshop-on-imt-for-2030-and-beyond-aka-6g/. Accessed 8 January 2023.

J. T. Rowley (✉)
E M Consultants International Pty Ltd, Melbourne, Australia
e-mail: jrowley@mobileimpact.com.au

V. Loscri et al. (eds.), *The Road towards 6G: Opportunities, Challenges, and Applications*, https://doi.org/10.1007/978-3-031-42567-7_7

Table 1 Selected suggested performance indicators for 6G relative to 5G from [1]

Performance indicator	5G	6G
Peak data rate (Gbps)	20	1000
Experienced data rate (Gbps)	0.1	1
Maximum channel bandwidth (GHz)	1	100
Area traffic capacity (Mbps/m^2)	10	1000
Connection density (devices/km^2)	10^6	10^7
End-to-end latency (ms)	1	0.1
Mobility (km/h)	500	1000

- Network and device antennas – are there novel features that need to be considered when assessing compliance?

Each of these factors will be discussed in the context of what is known about 6G and existing guidelines or standards in order to identify areas where work is required. The approach is similar to that of [2], the focus being on the engineering aspects and relying on independent expert health bodies to consider the possibility of new 6G RF-EMF health hazards.

2 Frequency

The ITU Radio Regulations cover the frequency range 8.3 kHz to 3 THz (3000 GHz). 4G/LTE is specified for frequency bands between 410 and 5900 MHz, though not all bands are in use [3]. 5G New Radio (NR) is specified in two frequency ranges: frequency range 1 (FR1) is similar to that used by existing mobile technologies and frequency range 2 (FR2) covers millimeter wave (mmWave) frequencies where greater bandwidths are available to provide novel higher data rate services [4]. The 5G NR bands are summarized in Table 2.

For reasons of coverage and capacity 6G will combine use of frequency ranges both above and below 100 GHz. Many frequency bands up to 100 GHz are already in use for existing mobile services and other radio applications. There is potential for mobile communication in the range 7 to 24 GHz (between 5G FR1 and FR2) subject to future ITU decisions.[2] As was the case with previous mobile generations, it is likely that spectrum would be refarmed over time for 6G. Frequencies above 100 GHz and up to 3 THz, where wider spectrum bandwidths are available, are proposed to support novel high data rate 6G applications operating over short ranges due to higher path loss [1, 5, 6]. It is important to emphasize that 6G does not mean only THz, in the same way that 5G does not mean only mmWaves. 6G will need to use a combination of lower frequencies for coverage and higher frequencies for extreme data rates.

---

[2] Informally referred to as FR3. Industry publications also refer to FR4 (71–114.25 GHz), FR5 (114.25–275 GHz) and FR6 (275–330 GHz), but these are not agreed official terminologies.

Table 2 3GPP definitions of 5G NR frequency ranges

Frequency range label	Frequency ranges
FR1	410 MHz to 7125 MHz
FR2	24,250 MHz to 71,000 MHz

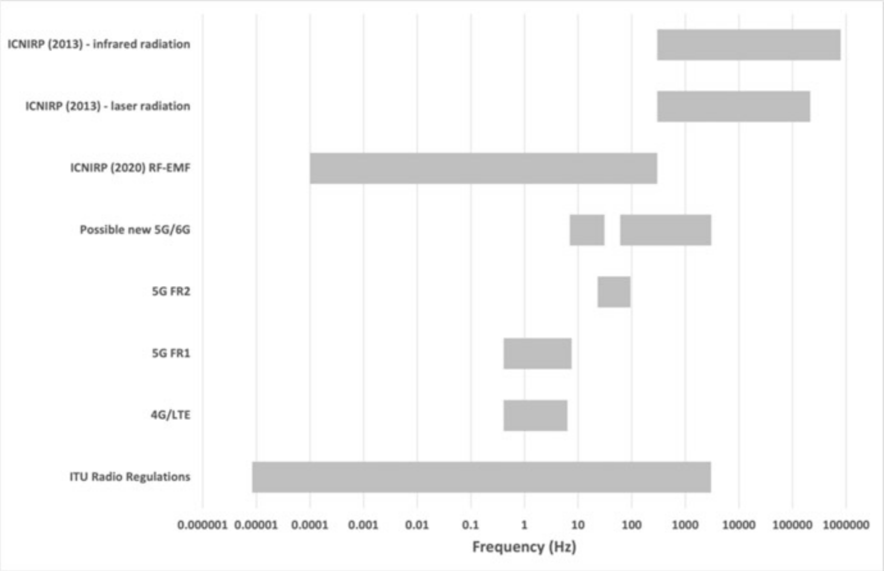


Fig. 1 Frequency ranges of the ITU Radio Regulations; 4G; 5G; possible new 5G/6G bands; ICNIRP (2020); ICNIRP (2013) infrared guidelines (laser and coherent)

In relation to RF-EMF exposure limits, the International Commission on Non-Ionizing Radiation Protection (ICNIRP) guidelines cover the frequency range 100 kHz to 300 GHz [7]. For frequencies above 300 GHz there are also ICNIRP guidelines that apply to infrared (IR) sources, either laser or incoherent [8, 9]. The frequency ranges discussed in this section are summarized in Fig. 1.

It is clear from Fig. 1, that 6G in bands below 300 GHz is covered by the ICNIRP RF-EMF guidelines, however, the applicable ICNIRP limits for potential 6G services operating above 300 GHz need clarification.

2.1 *Scientific Evidence Up to 300 GHz*

The international RF-EMF exposure limits are based on tissue heating as the biophysical mechanism for established health hazards from 100 kHz up to 300 GHz

[7]. This is supported by the consensus of the hundreds of expert reports by independent public health agencies and bodies published since the 1970s.[3]

Prior to 5G, mobile technologies used frequencies below 10 GHz. 5G added mmWave spectrum, which has been used for satellite, point-to-point radio links and radar for decades. More recent mmWave applications include some airport security scanners operating from 24 to 30 GHz. Assessments of mmWave airport security scanners did not identify any health risks [10]. These scanners generate peak exposures during a scan that are in the order of 1000 times below ICNIRP public RF-EMF exposure limits. Such exposure levels are comparable to the highest reported environmental exposures from 5G mmWave base stations.

Much higher exposures are delivered by the truck-mounted Active Denial Technology (ADT) developed by the US military to operate at 95 GHz. The ADT system produces 3-s pulses at more than 1000 times the ICNIRP public limit for continuous exposure, resulting in an increase of skin temperature to about 45 °C, producing an 'intolerable heating sensation' so that the targeted individual will instinctively move away. About 10,000 volunteer subjects were involved in development tests and no harm other than acute local heating of the skin was reported. Mice were exposed to the ADT signal alone and with a chemical carcinogen for skin cancer and no effect on cancer was observed [11].

Medical therapies using mmWaves at various frequencies, typically between 37 and 70 GHz, are used in the countries of Eastern Europe, Russia, and China. However, the therapies are not generally accepted by Western medicine [12]. The average power density incident on the skin is less than 200 W/m^2. This is double the ICNIRP (2020) local exposure limit for workers for this frequency range.

The ICNIRP (2020) public limit for whole-body exposures above 6 GHz is 10 W/m^2. The sun can deliver up to 800 W/m^2 to bare skin [13]. For frequencies from 6 GHz to 1 THz exposure at the ICNIRP public limit, produced a steady-state temperature rise of up to 0.4 °C for the skin [14, 15]. For exposures from distant sources, such as mmWave base stations, the temperature rise would be around 0.001 °C [19].

2.2 *Reviews of 5G mmWave RF-EMF Exposure*

The use of mmWave frequencies for 5G NR is one of the topics linked to misinformation about RF-EMF exposures [16, 17]. Health agencies from Australia [18], France [19] and the Netherlands [20] reviewed research related to 5G mmWave RF-EMF exposure. Their conclusions are summarized in Table 3.

The reviews agreed that compared to lower frequencies, there is less research at frequencies above 6 GHz; that some studies are not informative due to low quality

---

[3] See for example this listing https://www.gsma.com/publicpolicy/emf-and-health/expert-reports. Accessed 2 January 2023.

Table 3 Some authority reviews of 5G mmWave relevant research

Authority (year)	Frequency range	Summary	Follow-up
Health Council (2020) The Netherlands	20–40 GHz	More research needed. Postpone use of 26 GHz.	Rejected by government. Trials underway
ANSES (2022) France	18–100 GHz	More research needed. No conclusions possible.	Commercial launch pending. Trials underway.
ARPANSA (2021) Australia	6–300 GHz	No confirmed evidence of hazards. Future research should improve quality.	Commercial services commenced 2021

Table 4 Calculated peak electric field strength and penetration depth for a 10 W/m^2 incident field for different frequencies

Frequency (THz)	Peak electric field in the skin (V/m)	Penetration depth (μm)
0.1	40	~290
1.0	66	~100
10.0	68	~37

and that there are no confirmed indications of health hazards for 5G mmWave frequencies. They differ is in their view of the strength of the available scientific evidence, possibly due to different definitions of the relevant frequency range.

2.3 *Scientific Evidence Above 300 GHz*

There is significant interest in potential THz frequencies for medical applications; security screening; and for communication applications, as well as the potential for THz health hazards [5, 21–24].

The photon energy of frequencies in the range 0.3 to 3 THz ranges from 0.0012 eV to 0.0124 eV. The upper value is more than 800 times lower than the photon energy of 10 eV taken as the threshold for ionizing radiation. Therefore, THz signals are clearly non-ionizing but the energy difference is much smaller than the factor of more than 2.4 million times for a 1 GHz signal. Visible light, which starts at about 400 THz, has a photon energy of 1.7 eV.

The depth of penetration of RF energy in the skin reduces as the frequency increases as shown in Table 4, data from [24].

The short wavelength of THz signals creates new challenges in the dosimetry of biological experiments, and THz exposures are often highly inhomogeneous. This requires new approaches to the design and characterisation of exposure systems, including development of new THz sources [22, 23]. Additional data on dielectric properties of biological tissues at THz frequencies are also needed [25]. Free-field measurements are possible for THz sources but questions about sensitivity and the possibility of multi-polarization probes need further study [26].

3 Power

The available RF power from portable and mobile devices is generally limited by battery capacities whereas this is not the case for base stations. In modern mobile technologies efficient power control on both the uplink and downlink is important to management of the overall system interference. In the uplink case there is the added benefit of extending time of use for battery operated devices. There is a duality between the downlink and uplink power levels, with exposure from the device higher when exposure from the network is lower [27].

Researchers have investigated the combined RF-EMF exposure from both far-field and near-to-the-body sources as a metric of total exposure. An RF-EMF dose (with units of J/kg) has been described that includes duration in the exposure metric. These dose concepts may be of use to epidemiological studies.

Data was collected on RF-EMF sources operating from 50 MHz to 5.5 GHz to quantify far-field (broadcast, mobile networks, Wi-Fi); near-to-far field (small cell, Wi-Fi access point) and near-field (mobile phones, tablets, body-worn devices) exposures [28]. Whole-body and local SAR was calculated for several different human body models. The approach was illustrated with the example of a 25-year-old female subject exposed to several RF-EMF sources over 24-h. Far-field sources contributed <1% of whole-body dose and 1.6% of whole-brain dose with the rest of the exposure from the near-to-far field and near-field sources.

3.1 Mobile Devices

The time averaged power of mobile devices has been fairly constant over several mobile technology generations, as can be seen in Table 5 (based on [29]). 6G handheld devices are likely to use similar maximum powers.

Table 5 Summary of output powers for mobile devices of different generations

Mobile generation	Maximum time-averaged output power for handheld (mW)	Minimum transmit power relative to maximum (dB)	Typical average transmit power (percentage of maximum)
2G (GSM)	125–250	−33 to −36	20–50
3G (WCDMA)	250	−74	1–2
4G (LTE)	200	−63	1–2
5G (NR)	200	−56	1–2

Measurements on 3G, 4G and 5G devices (operating above and below 6 GHz) show similar output power levels for similar data rate services. Making a voice call places the phone close to the head but many other usages (for example, messaging applications or streaming video) move the device away from the body resulting reductions in RF-EMF exposure of 10 dB or higher [30].

The 3GPP specifications include higher power classes for some assumed applications, such as 5G mmWave customer premises equipment (CPE) for Fixed Wireless Access (FWA) [31]. Such equipment will likely need to be positioned so that line of sight to the network is not obstructed, which will reduce the possibility of a person entering the antenna beam. It is expected that similar output power levels and EIRP will exist for 6G CPE equipment.

3.2 Internet of Things (IoT) Devices

The IEEE provides guidance on RF-EMF exposure assessment for Internet of Things (IoT) devices based on frequency, bandwidth, radiated power, and typical installation configuration [32]. This would need to be extended to the new frequency ranges under discussion for 6G.

Both the RF-EMF levels (30 MHz to 3 GHz) and the duty-cycle of 55 IoT devices were measured in a convenience sample of ten homes [33]. The IoT devices used a range of RF transmission technologies and included smart gas and electricity meters; smart home sensors (for example, temperature and motion); various remote controllers; Wi-Fi networks and enabled equipment and Bluetooth devices. Maximum measured levels at 0.5 m ranged from 0.00000001 to 1.2 W/m^2, with a median level of 0.004 W/m^2. Six-minute duty cycles were between 0.002% and 100%, with a median value of 0.6%. Exposure ratios (relative to ICNIRP public limits) ranged from 3.3×10^{-10} to 0.01, excluding a push-to-talk walkie-talkie. As the IoT devices are unlikely to all transmit simultaneously, measurements of the continuous RF-EMF levels across the ten homes are informative for real-world levels and show a cumulative (worst-case) level of 0.002 W/m^2. This is 1000 times below the most restrictive ICNIRP public limit in this frequency range.

3.3 Mobile Networks

Measurements for more than 20 y of mobile networks shows that the typical environmental RF-EMF level remains a small fraction of the international limits.

Data from a selection of studies of environmental RF-EMF exposure trends are summarized in Table 6 and representative data are plotted in Fig. 2.

Table 6 Selection of studies reporting measured RF-EMF levels over multiple years

Country	Years	Median (W/m^2)	Min (W/m^2)	Max (W/m^2)	Source
Greece	2016–2022	0.00119	0.00220	0.12446	[34]
Romania	2019–2022	0.00357	0.00001	0.36559	
Serbia	2020–2022	0.00605	0.00018	0.05042	
France-Belgium	2021–2022	0.00161	0.00002	0.04372	
Spain (Catalonia)	2013–2022	0.00542	0.00003	1.44250	
Combined-Europe		0.00255	0.00001	1.44250	
Italy[a,b]	2002–2006	0.00079	0.00016	0.03940	[35]
Mauritania	2007–2010	0.00029	0.00000	0.00771	[36]
South Africa	2006–2012	0.00019	0.00000	3.90000	
Ireland[a]	2003–2009	0.00049	0.00000	0.00188	[37]
Greece[a]	2003–2009	0.00323	0.00000	0.06870	
Spain[a]	2002–2008	0.00537	0.00001	0.06010	
UK[a]	2001–2009	0.00001	0.00000	0.01100	
USA (San Francisco)[a]	2003–2009	0.01360	0.00019	0.70000	

[a] Mean rather than median

[b] 25th percentile instead of min and 95th percentile instead of max

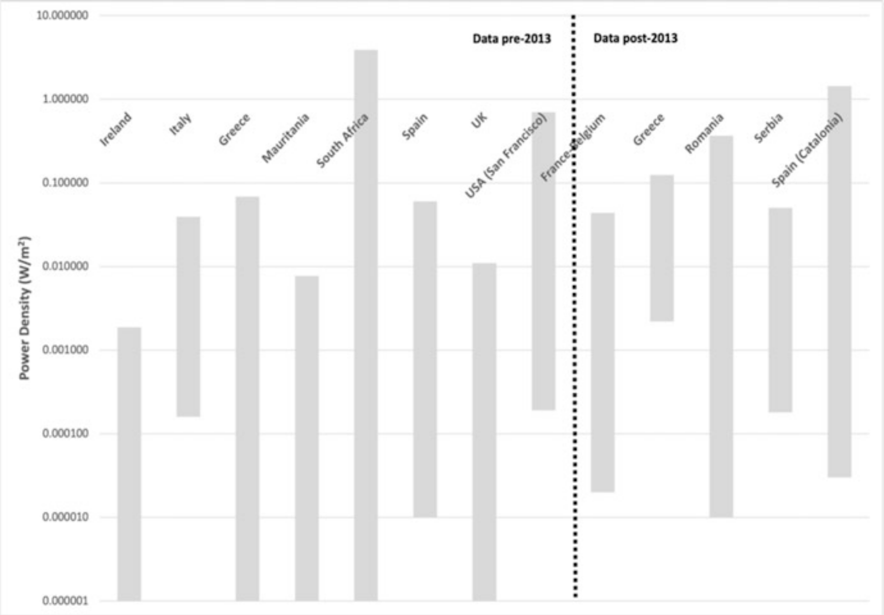


Fig. 2 Range of measured environmental RF-EMF levels for a selection of countries. In 2013, about 20% of people were covered by LTE networks. Smartphones were about 40% of mobile subscriptions

Care needs to be taken in comparing the results between countries due to the different equipment used, locations chosen and whether extrapolated or actual RF-EMF levels are reported. However, Fig. 2 shows no obvious increase in RF-EMF levels for data covering the period 2001 to 2022.

Another approach to assessing RF-EMF exposure is the use of personal exposimeter devices that are carried by trained researchers or volunteers and used to monitor personal exposure during daily activities.

A systematic review with an international sample of personal dosimeter studies published between 1998 and 2021 reported that the lowest mean was measured in Egypt with a value of 0.00100 $\mu W/m^2$ in 2007 and the highest mean was measured in Belgium with a value of 285,000 $\mu W/m^2$ in 2019 [38]. A further systematic review that combined data from personal exposimeter studies and fixed site monitoring systems in European countries found no indication of an increase in personal RF-EMF exposure since 2002 and no distinct differences in exposure levels between countries [39]. This last part is important as some of the included countries (Belgium, Greece, Italy, Poland, Serbia, Slovenia) had RF-EMF limits more restrictive than the ICNIRP guidelines yet typical public exposure levels are the same as countries with the international limits.

3.4 5G Network RF-EMF Exposure Levels

Measurements have been reported for trial 5G sites and for commercial 5G networks [40–43], including 5G deployment using Dynamic Spectrum Sharing (DSS) [44]. The GSMA maintains an interactive map[4] summarizing the results of published 5G RF-EMF surveys. Overall, measurement surveys show similar RF-EMF levels to other mobile network technologies, with maximum levels less than 1% of the ICNIRP public limits. However, the traffic on 5G networks is currently low in many countries and the 5G level may increase in the future while remaining well below limits. Some researchers have also proposed EMF aware methods to further reduce the already very low RF-EMF levels [17] and further study is needed to determine potential impacts on service objectives.

4 6G Technology Characteristics

It is anticipated that 6G implementation will be an evolution of existing radio technologies with particular emphasis from the perspective of RF-EMF exposure on expanded use of active antennas and network densification [1, 6].

---

[4] https://www.gsma.com/publicpolicy/emf-and-health/safety-of-5g-networks/5g-emf-surveys. Accessed 6 January 2023.

4.1 Influence of Beamforming on 6G Network RF-EMF Exposure Levels

It is expected that 6G will continue the use of Massive MIMO (MaMIMO) antennas in both devices and networks. These technologies may be augmented by the use of intelligent reflecting surfaces to control the distribution of RF-EMF energy; cell-free MaMIMO that can assist with geographic coverage; and advanced power control to optimize use of radio resources [6].

MIMO antennas are in use in some 4G networks [45] but the deployment of beamforming antennas has expanded in 5G to provide both greater capacity and to overcome higher path losses at higher frequencies. Beamforming reduces the level of RF-EMF outside the location where connectivity is being provided [46]. Studies show that when total dose (uplink plus downlink) is considered, 5G networks may result in lower exposure than 4G networks [47, 48]. It is too soon to say whether that could also be the case with 6G.

4.2 Influence of 6G Network Densification on RF-EMF Exposure Levels

Another trend that will likely continue is greater network densification to deliver improved coverage or greater capacity. In the past, low-powered sites were termed microcells or femto cells or distributed antenna systems (DAS) but now they are generally grouped under the term small cells. The idea that small cells will be placed 'everywhere' is not supported by the practical needs to provide power and data connections for each antenna. The reality is that 5G small cells (and in the future 6G small cells) are used in locations such as airports, sports stadia and transport hubs to provide targeted service improvements and not across whole urban areas [2].

The number of visible antennas or antenna proximity are not good indicators of RF-EMF exposure levels [49]. In the case of 5G, network densification does not change RF-EMF levels for most of the population but there is some evidence of a reduction for those living closest to a base station [50].

Sharing of mobile network infrastructure, whether physical structures or radio technology resources, is a feature of deployment in many countries. Considering multi-technology sites as a type of network sharing, there is mixed evidence about the effect on RF-EMF levels in nearby areas [51, 97]. A relatively new development is open radio access network (Open RAN) technology, which aims to use generic hardware instead of proprietary equipment. This is being used to support RAN sharing in parts of rural Europe.[5] As Open RAN technology matures, it may reduce

---

[5] For example, Orange and Vodafone to cooperate on Open RAN network sharing in rural areas across Europe. Orange press release. 22 February 2023. Accessed at https://newsroom.orange.com/orange-and-vodafone-to-cooperate-on-open-ran-network-sharing-in-rural-areas-across-europe/ on 26 April 2023.

the barriers to network sharing, but this is unlikely to affect RF-EMF levels as the total amount of data to be transmitted is going to be similar for one shared versus two independent networks.

Other future mobile technologies that could influence RF-EMF levels include cell-free networks [52] and joint optimization of the uplink and downlink considering EMF exposure and communication performance [53].

From a down-link point of view, small cells RF-EMF exposure levels are of the same order as from the macro network [54]. When considering total RF-EMF exposure, 4G small cell installations may increase the downlink exposure by about a factor of two while reducing the uplink level by about 10-times (if not throughput limited) due to the shorter communications path [55]. Comparable data is not currently available for 5G or 6G but the results are likely to be similar.

4.3 Influence of New Use Cases and New Device Types on 6G Exposures

It is generally true that uplink (device) RF-EMF exposure dominates over downlink (network) exposures [28]. However, with 5G the exposure is becoming more complex and more dependent on the communication activity of the user, which should be considered in the exposure assessment protocol [56]. This will continue with 6G, with additional complexity due to non-smartphone formfactors such as headsets or emotion-sensing wearable devices [57]. Device-based applications were developed to track RF-EMF exposure for research purposes [58, 59] and it is conceivable that these could be adapted to consumer-oriented applications. However, the value of such approaches is unclear as the international RF-EMF limits are threshold based. Additionally, risk perception is not strongly influenced by information on individual exposure [60]. It would be more effective to continue the current practical advice that reducing use or increasing device-body separation are steps that persons who want to reduce RF-EMF exposure can take.

5 Technology Independent Exposure Limits

In the absence of mandatory international RF-EMF limits, both the WHO [61] and the ITU [62] encourage countries to base their policies on the ICNIRP guidelines. The ICNIRP RF-EMF guidelines are technology independent and designed to protect against all established health hazards of RF-EMF exposure. Therefore, the ICNIRP guidelines would apply to 6G RF-EMF exposures, whether from devices or networks.

However, frequency bands (for example, in the THz range) may require a review of the limits to ensure continuity across frequency ranges and to define new limits if new exposure situations are identified. A similar process was undertaken in the recent updates by ICNIRP that take account of the expanded use of mmWaves for 5G [7].

5.1 Exposure Limits Up to 300 GHz

The ICNIRP (2020) basic restrictions applying to energy absorption internal to the body for exposures longer than 6 min are summarized in Table 7. For frequencies above 30 MHz the main reference levels (external field quantities) averaged over 30 min and the whole-body are presented in Table 8.

The ICNIRP (2020) guidelines contain additional limits for pulsed-type exposures that may be relevant for 6G uses that result in short transmission times.

For countries and mobile network operators considering moving from the ICNIRP (1998) limits to ICNIRP (2020) it is useful to know that for base stations the changes are marginal and that base stations already compliant with the former will remain compliant with the latter [63].

Table 7 ICNIRP (2020) basic restrictions above 30 MHz

Population	Frequency range	Whole-body average SAR (W/kg)	Local head/torso SAR (W/kg)	Local limb SAR (W/kg)	Local S_{ab} (W/m^2)
Occupational	100 kHz to 6 GHz	0.4	10	20	NA
	>6 to 300 GHz	0.4	NA	NA	100
Public	100 kHz to 6 GHz	0.08	2	4	NA
	>6 to 300 GHz	0.08	NA	NA	20

Note: *NA* not applicable. Other ICNIRP notes not reproduced, please consult [7] for details

Table 8 ICNIRP (2020) whole-body reference levels above 30 MHz

Population	Frequency range	Incident E-field (V/m)	Incident H-field (A/m)	Incident power density (W/m^2)
Occupational	>30 to 400 MHz	61	0.16	10
	>400 to 2000 MHz	$3f_M^{0.5}$	$0.008f_M^{0.5}$	$f_M/40$
	>2 to 300 GHz	NA	NA	50
Public	>30 to 400 MHz	27.7	0.073	2
	>400 to 2000 MHz	$1.375f_M^{0.5}$	$0.0037f_M^{0.5}$	$f_M/200$
	>2 to 300 GHz	NA	NA	10

Note: f_M is frequency in MHz. *NA* not applicable. Other ICNIRP notes not reproduced, please consult [7] for details

5.2 Controversy About Biological Effects from Low-Level Exposures

The ICNIRP acknowledges that there is evidence of biological effects at low levels of exposure but as there are no proven consequences for human health these effects are not used to establish limit values [7]. The ICNIRP (2020) guidelines considered the evidence for adverse health effects regardless of the potential mechanism, stating:

> For the purpose of determining thresholds, evidence of adverse health effects arising from all radiofrequency EMF exposures is considered, including those referred to as ‘low-level’ and ‘non-thermal’, and including those where mechanisms have not been elucidated.

The ICNIRP (2020) limit values are based on RF-EMF health effects that ICNIRP assessed as being ‘*both harmful to human health and scientifically substantiated.*’ There are arguments, both theoretical and biological, against the possibility of mechanisms at nonthermal levels and in particular that signal modulation is not significant for health risk assessment [65, 66], so separate health evaluations are not needed for new RF technologies.

Some groups accuse ICNIRP of ignoring evidence of harm at RF-EMF levels below the international limit values [16]. One of the reasons for the differing interpretations of the scientific literature is the approach to study quality assessment. It is unfortunately true for RF-EMF research that the quality of the scientific literature is highly variable with several reviews finding that lower quality studies were more likely to report an association between RF-EMF exposure and a biological endpoint [67–69]. Industry funded studies tend to be of higher quality, which may explain why they are less likely to report health hazards rather than alternative conspiracy theories about influence on researchers [70, 71].

The 2011 classification of RF-EMF as a possible human carcinogen by a working group of the International Agency for Research on Cancer (IARC) is often misunderstood [72, 73]. The IARC process is a hazard assessment exercise and examines the strength of the evidence that exposure to an agent at any level is linked to carcinogenesis. It does not consider risk, which includes consideration of other evidence, such as actual population exposures. In the case of RF-EMF, the IARC working group concluded that there was limited evidence showing a link between wireless phone exposures and certain brain tumors. The evidence for environmental type exposures (broadcasting and mobile networks) and occupational exposures was judged inadequate.

It is generally accepted that the epidemiological evidence for a mobile phone link to brain cancer has weakened since the IARC classification but it is not possible to exclude the existence of a small risk [74, 75]. The results of US National Toxicology Program (NTP) rodent bioassays in 2019 appeared to link high-level RF-EMF exposures to cancer development in animals, however, interpretation has been difficult for several reasons including statistical considerations and the possibility of heat stress as an alternative explanation [76, 77].

An Advisory Group recommended that IARC revisit the RF-EMF classification in the second half of the period 2020–2024 [78]. As the initial IARC review of RF-EMF was only scheduled when sufficient new evidence was available it seems likely that IARC will wait for results of the international cohort study of mobile phone users (COSMOS) [79] and the Japan/South Korea NTP study replications [80] before calling for working group participants. The IARC Monographs programme has also changed procedures, with greater emphasis on mechanistic evidence and this could influence a future re-evaluation of RF-EMF [81].

Further clarity on the possibility of adverse health effects from RF-EMF exposures is expected from the WHO Task Group on Radiofrequency Fields and Health Risks[6] that is expected to complete work in 2024. A previous WHO risk assessment of RF-EMF was undertaken in 1993 and it has no mention of mobile phone technology. Inputs to the Task Group include ten systematic reviews[7] on RF-EMF health priorities identified by the WHO following a survey of EMF experts [82].

Health risk assessments often gave more weight to what is uncertain. However, it is also possible to be clear about what is known, as can be seen in the following statement from the 2015 SCENIHR review for the European Commission [83]:

> Several interaction mechanisms are well established. These enable extrapolation of scientific results to the entire frequency range and wide-band health risk assessment. They have been used to formulate guidelines limiting exposures to EMF in the entire frequency range from static fields to 300 GHz

It is important that the WHO Task Group is forward looking so that it can provide reassurance about the safety of not yet deployed RF technologies provided they comply with the international RF-EMF exposure guidelines.

5.3 RF-EMF Protection for Flora and Fauna

There are reports of low-level RF-EMF exposure affecting animals (especially insects and birds) and plants such as trees. Substantial changes in the environment, such as increased urbanization, agricultural practices and climate change, are more likely to explain reported changes in biodiversity [84, 85].

The ICNIRP (2020) guidelines for RF-EMF exposure are not intended to provide protection against environmental effects of RF-EMF exposure [7]. However, ICNIRP has established a project group to analyze whether the current human exposure guidelines are sufficiently protective for plants and animals in their

---

[6] https://www.who.int/teams/environment-climate-change-and-health/radiation-and-health/non-ionizing/emf/radiofrequency-fields. Accessed 7 January 2023.

[7] The systematic review protocols and outcomes are planned for publication as a special issue of *Environment International* accessible here: https://www.sciencedirect.com/journal/environment-international/special-issue/109J1SL7CXT. Accessed 8 January 2023.

natural environment. Australian scientists are undertaking a systematic map of the available evidence to inform authorities, identify knowledge gaps and develop research recommendations [86]. In the meantime, the results of a 2019 workshop on environmental effects of RF-EMF exposure are reassuring, while reiterating the need for additional higher quality studies [87]:

> The results presented at the workshop did not show any sound scientific evidence of adverse effects of low-level anthropogenic RF-EMFs at frequencies exceeding 100 MHz on animals or plants under realistic environmental conditions. Extrapolations from laboratory animal studies, often performed at higher exposure levels, do not allow conclusions on ecological effects of RF-EMFs at low levels. Field studies of an appropriate quality are scarce in both animals and plants and so far do not show clear evidence supporting adverse effects of RF-EMFs. Some correlations between RF-EMFs and adverse biological effects were observed, but bias and confounding factors cannot be excluded

It is well established that size influences the amount of RF-EMF energy absorbed by the body, with highest absorption when the wavelength of the incident field is comparable to the size of the organism. It is no surprise then that dosimetric studies show higher RF-EMF energy absorption for insects at higher frequencies [88]. However, the maximum RF-EMF power absorption in a honeybee is small at 0.7 nW. The metabolic power production of a flying honeybee is 495 mW at 25 °C [89], about 700 million times higher than the absorbed RF-EMF power suggesting that the RF-EMF would have no significant impact on bee behavior. At the THz frequencies discussed for 6G, RF-EMF absorption would be much smaller than the heating effect of incident solar IR.

5.4 Exposure Limits Above 300 GHz

There are ICNIRP exposure guidelines for laser and incoherent IR that apply above 300 GHz [8, 9]. The ICNIRP IR guidelines are based on limiting thermal damage to the eyes and skin. Power density (W/m^2), known as irradiance in the IR guidelines, is used to compare the limits in Fig. 3.

There is a factor of ten difference between the ICNIRP (2020) whole-body limit for the public and the ICNIRP (2013) limit for incoherent IR. The ICNIRP (2020) local limits for the public also differ but the margin is smaller. The gaps are even larger relative to the ICNIRP (2013) laser limit but as this is for a particular product class it is unclear how relevant it would for THz sources. It is important that ICNIRP address the apparent step change in limit values between the RF-EMF and IR guidelines in anticipation of 6G at THz frequencies.

Not all countries follow the ICNIRP guidelines, and some countries have in place significantly more restrictive limits that apply to fixed transmitters (there are fewer country specific RF-EMF exposure limit restrictions on devices, which are global products) [64]. In most cases the restrictive limits existed before the widespread deployment of mobile technologies and reflect long-held differing interpretations of EMF science. Harmonization of RF-EMF limits based on the ICNIRP guidelines

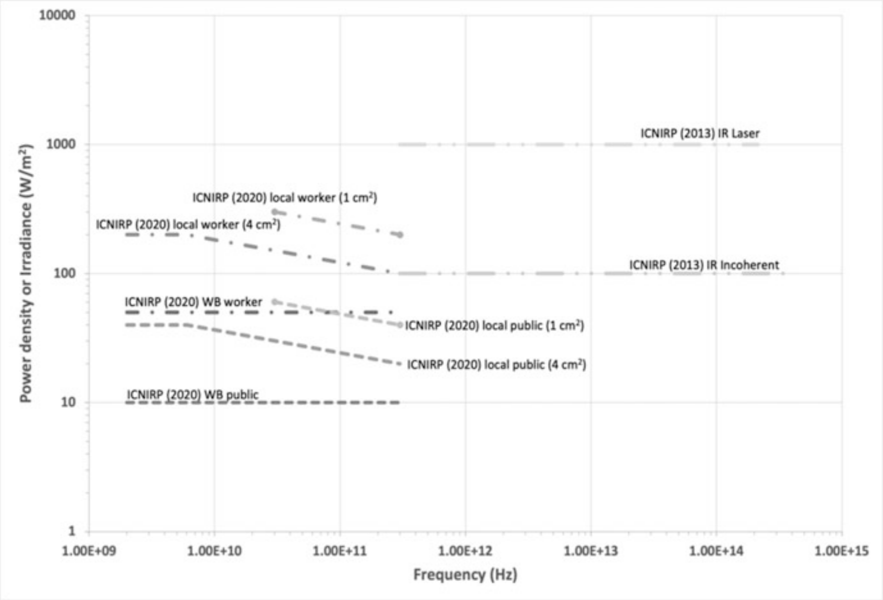


Fig. 3 Simplified comparison of exposure limits for RF-EMF and IR exposure guidelines from 2 GHz to 1 × 10^6 GHz. (Note: *WB* whole body. 1 cm^2 and 4 cm^2 refer to averaging areas in the ICNIRP (2020) guidelines. Details simplified for comparison purposes, refer to [7–9] for details)

is identified as a positive policy goal for countries [62]. Providing clarity on the application of ICNIRP limits to THz frequencies would eliminate a potential technical obstacle to adopting harmonized limits for all 6G frequencies.

6 Assessing RF-EMF Compliance

International technical standards [90] describe a range of methods for the accurate assessment of RF-EMF compliance boundaries at frequencies up to 300 GHz. Different calculation or measurement approaches may be applied depending on the circumstances.

For devices a similar situation exists with both computational approaches and measurement methods defined in the international technical standards. For mobile devices intended for use close to the body, depending on the frequency range the compliance assessment may be done relative to basic restrictions for specific absorption rate (SAR in W/kg) or power density (W/m^2).

For both mobile networks and mobile devices it will be necessary to consider how the methods can be extended to proposed THz frequencies for 6G.

6.1 Network Assessment Methods

In order to achieve energy efficiency targets, 5G has reduced signaling overhead and this created some challenges in the development of in situ 5G RF-EMF measurement methods. Methods to measure RF-EMF exposure levels from 5G networks have been reported [91], including for mmWave frequencies [92], and incorporated into international technical standards [90]. The measurement methods for 3G, 4G and 5G rely on the measurement of a stable reference signal within the communication frame. This level can then be extrapolated to the total exposure at a location at particular point-in-time or, with additional information, the maximum possible exposure at a location. When the 6G signaling scheme is being defined, it would be helpful to consider how RF-EMF assessments can be conducted.

It is known that the maximum time-averaged power per beam direction for 5G networks is well-below the theoretical maximum for a site [93]. If theoretical values are used as the basis of determining RF-EMF levels, the estimates are likely to be less accurate than approaches using the actual time-averaged base station transmissions and real effective antenna gain [94, 95].

This has led to the definition of the actual maximum approach to assess RF-EMF compliance boundaries [90]. A consequence of using the actual maximum approach with power control is that there is a potential for some loss of site throughput, with [96] estimating a 10% throughput loss when the average power threshold is set to 25% of the maximum transmit power. It should be noted that the actual maximum approach is applicable to other generations of mobile technologies, whether current or future.

As a related point, for base station sites with multiple technologies, approaches in some national regulations of requiring that RF-EMF assessments are conducted by assuming that all transmitters transmit at the same time at their maximum power over-estimates the real exposure and should be revisited [97].

6.2 Device Assessment Methods

There are a family of standards from the International Electrotechnical Commission (IEC) for the RF-EMF compliance assessment of mobile devices operating close to the head or body as summarized in Table 9. Additional IEC standards are available for low-power and other devices.

As was shown in Table 5, mobile devices typically operate at levels much lower than their maximum rated power, yet the compliance standards are based on simulating maximum possible exposure conditions. The device compliance test methods have their origin in compliance of mobile phones used against the head. Current usage is much more varied, with voice calls usually a minority of the time of use. The test methods may need to evolve to cater for new 6G device formfactors.

Table 9 IEC standards for assessing mobile device RF-EMF compliance

Standard	Frequency range	Assessment quantity	Method
IEC/IEEE 63195-1:2022	6–300 GHz	Power density	Measurement
IEC/IEEE 63195-2:2022	6–300 GHz	Power density	Computation
IEC/IEEE 62209-1528:2020	0.004–10 GHz	SAR	Measurement
IEC 62209-3:2019	0.6–6 GHz	SAR	Measurement (vector)
IEC/IEEE 62704-3:2017	0.03–6 GHz	SAR	Computation (FDTD)
IEC/IEEE 62704-4:2020	0.03–6 GHz	SAR	Computation (FEM)

7 Information and Misinformation

The deployment of 4G was associated with relatively low levels of public concern, probably in a part because deployment could largely reuse existing base station sites. In some countries, the situation with 5G was quite different and, the political response to social concerns resulted in restrictions on 5G until additional advice was provided to policymakers. Five possible reasons for the different public response follow.

Claimed novelty of 5G technology – though applications for small cells, active antennas and mmWave frequencies can be found in earlier mobile technology generations, their expanded use for 5G provided convenient hooks for concerned groups to generate wider public concern. This was especially the case in the USA, where legislative efforts to limit municipality authorization of small cell deployments could be characterized as attempts to bypass local decision making.

Authority information vacuum – following the relatively low-key public interest in the rollout of 4G, both authorities and the mobile industry were slower to react to campaign group claims about 5G. Many authorities took the position that 5G was covered by existing RF-EMF exposure limits and there was no need for technology specific information. As a result there were few 5G specific RF-EMF health statements by national or international health authorities when the first public appeal to stop 5G in Europe was launched [98].

Revolutionary claims about 5G – commentary by industry groups, think-tanks and government bodies described 5G as revolutionary and capable of transforming societies, economically, socially and environmentally. This acted to amplify both the positive aspects of 5G and noise about unsubstantiated 5G health claims.

Social media – the contribution of social media to health misinformation is well known and there are online communities concerned about RF-EMF exposures [99]. These channels are a much more important source of information than was the case when 4G launched in 2009.

Wider societal concerns – 5G is the first mobile technology where national security concerns about network equipment suppliers has been a significant discussion. In our climate change impacted world questions about energy efficiency and e-waste have also been raised about the transition to 5G. Other social concerns, such as online privacy and work insecurity, were also linked to 5G [100].

Table 10 Six points for communication about 6G RF-EMF exposure and standards

1	6G will use radio waves
2	6G will use frequencies similar to other mobile technologies and new radio spectrum
3	6G will comply with national and international safety limits
4	6G signals are expected to be similar to typical levels from other radio technologies
5	6G levels in public areas when combined with other radio sources must comply with limits
6	6G test methods will evolve to cover new applications

Looking forward to 6G, it is important that trusted authorities monitor the technology development so that they are in the position to make clear statements about what is known about RF-EMF exposures and the limits that protect people and the environment. Some communication points that might be made are listed in Table 10.

Social media is a challenging environment in which to convey information. Before engaging, EMF policymakers and scientists are encouraged to consult research-based guidance on successful debunking of misinformation [101] and communicating fact checks online [102].

7.1 COVID-19 and 5G Misinformation

Misinformation reached a peak in early 2020 during the COVID-19 pandemic when claims linking 5G to the disease were widely shared on social media and associated with attacks on telecommunications infrastructure in several parts of the world [17, 99, 103]. The April 2020 decision by the WHO[8] to include 5G in its COVID-19 mythbusters series with the clear message that '*5G mobile networks DO NOT spread COVID-19*' marked a step-up in activities by national authorities on 5G misinformation more generally. The WHO position was echoed by fact checking organizations, governments, and mainstream media. Anecdotally, a positive development is that mainstream media continues to be less willing to include coverage of unsubstantiated claims around 5G and health risks.

8 Preparing for the Future of 6G and RF-EMF

While we do know that 6G devices and mobile network antennas will need to comply with national and international guidelines for RF-EMF exposures, there is still work to be done to prepare for future deployments. Based on the experience

---

[8] https://www.who.int/multi-media/details/covid-19-5g-mobile-networks-do-not-spread-covid-19 (emphasis in the original). Accessed 7 January 2023.

of previous mobile technology generations here are six areas where additional preparatory work is required.

Extension of RF-EMF limit values to 3 THz – current international RF-EMF exposure limit guidelines apply up to 300 GHz but 6G may use bands up to 3 THz. ICNIRP should commence work to extend the frequency range of the existing RF-EMF guidelines and clarify the limit transition to the IR guidelines.

Updated RF-EMF assessment methods for 6G – new frequency ranges and new bandwidths, combined with expanded use of antenna technologies like MaMIMO will require the updating of assessment methods, both calculation and measurement.

Futureproof the WHO RF-EMF risk assessment – it would be helpful to 6G (and other new applications of RF-EMF) if the WHO Task Group could provide reassurance about the safety of not yet deployed RF technologies provided they comply with the international safety guidelines.

Real-world 6G RF-EMF exposure levels – information on the real-world exposures from 6G devices and networks will be important to public communication, even though based on the past two-decades of expansion of mobile technology services it is expected that the overall level of RF-EMF will remain largely unchanged.

Plan for effective science communication – with hindsight the mobile industry, governments and other stakeholders did not prepare effectively to address questions about compliance of 5G with exposure limits. Greater anticipation of the concerns that may arise and the availability of information at the time of 6G deployment will be important.

6G RF-EMF research needs – as past research at related frequencies will be relevant to 6G, it is unclear what value future biological studies will deliver in the absence of an identified mechanism other than heating. If biological studies are contemplated they should be of high quality in order to be informative to agencies undertaking health risk assessment [68].

9 Conclusion

Many of the performance targets for 6G are not yet fully defined but it aims to deliver data rates, reduced latency and increased connection density beyond the capabilities of 5G. 6G will use frequencies similar to existing mobile networks from a few hundred MHz to about 100 GHz. It may also use higher frequencies up to 3 THz but there are technical challenges. From an RF-EMF point-of-view regardless of the frequencies that are selected, 6G will need to comply with technology independent exposure guidelines developed by ICNIRP (or national equivalents), which provide protection for all persons against all established health hazards. Existing international RF-EMF exposure guidelines and assessment methods are specified to 300 GHz. With 6G expected to launch in the 2030s, there is time for the responsible bodies to undertake the studies necessary to extend the frequency ranges. In the last two decades of mobile network evolution measurement studies consistently show that the typical level of exposure from mobile networks is a small fraction of the international guidelines and that environmental levels have not changed by very much. It is expected that 6G will be a similar story, even with increased densification of base stations and IoT devices.

There is work to be done to prepare for questions about RF-EMF exposure and 6G. Authorities should continue to work towards the harmonization of national RF-EMF exposure limits and assessment methods with international guidance. It is also important that the WHO health risk assessment for RF-EMF is future looking and can provide reassurance about the safety of not yet deployed RF technologies so long as they comply with the international exposure guidelines. Industry and government stakeholders can learn from the experience of 5G and ensure that information on real-world 6G RF-EMF levels is available as well as clear and specific statements on 6G safety from trusted agencies.

Conflict of Interest The author provides consultancy services to companies in the mobile telecommunications industry. No funding was provided for this work. Observations and analyses reported here are solely those of the author.

References

1. Akyildiz et al (2020) 6G and beyond: the future of wireless communications systems. IEEE Access 8:133995–134030. https://doi.org/10.1109/ACCESS.2020.3010896
2. Chiaraviglio et al (2021) Health risks associated with 5G exposure: a view from the communications engineering perspective. IEEE Open J Commun Soc 2:2131–2179. https://doi.org/10.1109/OJCOMS.2021.3106052
3. TS 36.101: Evolved universal terrestrial radio access (E-UTRA); User equipment (UE) radio transmission and reception (17.7.0 ed). 3GPP 359. 26 Sept 2022
4. TS 38.104: New Radio (NR); Base station (BS) radio transmission and reception (Release 17). 3GPP. Sept 2022
5. Rappaport et al (2019) Wireless communications and applications above 100 GHz: opportunities and challenges for 6G and beyond. IEEE Access 7:78729–78757. https://doi.org/10.1109/ACCESS.2019.2921522
6. Matthaiou et al (2021) The road to 6G: ten physical layer challenges for communications engineers. IEEE Commun Mag 59(1):64–69. https://doi.org/10.1109/MCOM.001.2000208
7. International Commission on Non-Ionizing Radiation Protection (ICNIRP) (2020) Guidelines for limiting exposure to electromagnetic fields (100 kHz to 300 GHz). Health Phys 118(5):483–524. https://doi.org/10.1097/hp.0000000000001210
8. ICNIRP guidelines on limits of exposure to laser radiation of wavelengths between 180 nm and 1,000 μm. (ICNIRP) (2013) Health Phys 105(3):271–295. https://doi.org/10.1097/hp.0b013e3182983fd4
9. Guidelines on limits of exposure to incoherent visible and infrared radiation. (ICNIRP) (2013) Health Phys 105(1):74–96. https://doi.org/10.1097/HP.0b013e318289a611
10. Airport passenger screening using millimeter wave machines: compliance with guidelines (2017) Natl Acad Sci Eng Med. https://doi.org/10.17226/24936. Accessed 2 Jan 2023

11. Mason et al (2001) Lack of effect of 94 GHz radio frequency radiation exposure in an animal model of skin carcinogenesis. Carcinogenesis 22(10):1701–1708. https://doi.org/10.1093/carcin/22.10.1701
12. Alekseev et al (2018) Biological effects of millimetre and submillimeter waves. In: Biological and medical aspects of electromagnetic fields. https://doi.org/10.1201/9781315186641
13. Arens et al (2006) The skin's role in human thermoregulation and comfort. In: Thermal and moisture transport in fibrous materials, pp 560–602. https://doi.org/10.1533/9781845692261.3.560
14. Christ et al (2020) RF-induced temperature increase in a stratified model of the skin for plane-wave exposure at 6–100 GHz. Radiat Prot Dosim 188(3):350–360. https://doi.org/10.1093/rpd/ncz293
15. Sasaki et al (2017) Monte Carlo simulations of skin exposure to electromagnetic field from 10 GHz to 1 THz. Phys Med Biol 62(17):6993. https://doi.org/10.1088/1361-6560/aa81fc
16. de Vocht et al (2022) The population health effects from 5G: controlling the narrative. Front Public Health 10. https://doi.org/10.3389/fpubh.2022.1082031
17. Elzanaty et al (2021) 5G and EMF exposure: misinformation, open questions, and potential solutions. Front Commun Netw 2(5):19. https://doi.org/10.3389/frcmn.2021.635716
18. Karipidis et al (2021) 5G mobile networks and health – a state-of-the-science review of the research into low-level RF fields above 6 GHz. J Expo Sci Environ Epidemiol 31:585–605. https://doi.org/10.1038/s41370-021-00297-6
19. Expositions aux champs électromagnétiques liées au déploiement de la technologie de communication « 5G » et effets sanitaires éventuels associés. Anses January 2022. https://www.anses.fr/fr/system/files/AP2019SA0006RA-2.pdf. Accessed 2 Jan 2023
20. 5G and health. Health Council of the Netherlands No. 2020/16e: 2 September 2020. https://www.healthcouncil.nl/documents/advisory-reports/2020/09/02/5g-and-health. Accessed 4 Jan 2023
21. Gallerano et al (2004) THz-BRIDGE: an European project for the study of the interaction of terahertz radiation with biological systems. In: Infrared and millimeter waves, conference digest of the 2004 Joint 29th international conference on 2004 and 12th international conference on terahertz electronics, pp 817–818. https://doi.org/10.1109/ICIMW.2004.1422345
22. Wilmink et al (2011) Current state of research on biological effects of terahertz radiation. J Infrared Millimeter Terahertz Waves 32(10):1074–1122. https://doi.org/10.1007/s10762-011-9794-5
23. Kleine-Ostmann et al (2014) Field exposure and dosimetry in the THz frequency range. IEEE Trans Terahertz Sci Technol 4(1):12–25. https://doi.org/10.1109/TTHZ.2013.2293115
24. Spathmann et al (2015) Numerical computation of temperature elevation in human skin due to electromagnetic exposure in the THz frequency range. IEEE Trans Terahertz Sci Technol 5(6):978–989. https://doi.org/10.1109/TTHZ.2015.2476962
25. Lajevardipour et al (2016) Estimation of dielectric values for tissue water in the terahertz range. Bioelectromagnetics 37(8):563–567. https://doi.org/10.1002/bem.22010
26. Zheng et al (2021) An overview of probe-based millimeter-wave/terahertz far-field antenna measurement setups. IEEE Antenna Propag Mag 63(2):63–118. https://doi.org/10.1109/MAP.2021.3054017
27. Gati et al (2010) Duality between uplink local and downlink whole-body exposures in operating networks. IEEE Trans Electromagn Compat 52(4):829–836. https://doi.org/10.1109/TEMC.2010.2066978
28. Liorni et al (2020) Evaluation of specific absorption rate in the far-field, near-to-far field and near-field regions for integrative radiofrequency exposure assessment. Radiat Prot Dosim 190(4):459–472. https://doi.org/10.1093/rpd/ncaa127
29. ITU-T K series, Supplement 13: radiofrequency electromagnetic field (RF-EMF) exposure levels from mobile and portable devices during different conditions of use. ITU December 2021. https://handle.itu.int/11.1002/1000/14881
30. Hadjem et al (2010) Analysis of power absorbed by children's head as a result of new usages of mobile phone. IEEE Trans Electromagn Compat 52(4):812–819. https://doi.org/10.1109/TEMC.2010.2052810

31. TS 38.101-2: NR; user equipment (UE) radio transmission and reception. Part 2: Range 2 standalone (17.7.0 ed). 3GPP. 28 Sept 2022
32. IEEE Guide for EMF exposure assessment of Internet of Things (IoT) technologies and devices. IEEE Std 1528.7–2020. 1–90. 2021. https://doi.org/10.1109/IEEESTD.2021.9319817
33. Aerts et al (2019) Emissions from smart meters and other residential radiofrequency sources. Health Phys 116(6):776–788. https://doi.org/10.1097/hp.0000000000001032
34. Iakovidis et al (2022) Electromagnetic fields exposure assessment in Europe utilizing publicly available data. Sensors 22(21):8481. https://doi.org/10.3390/s22218481
35. Rowley et al (2016) Observations from national Italian fixed radiofrequency monitoring network. Bioelectromagnetics 37(2):136–139. https://doi.org/10.1002/bem.21958
36. Joyner et al (2014) National surveys of radiofrequency field strengths from radio base stations in Africa. Radiat Prot Dosim 158(3):251–262. https://doi.org/10.1093/rpd/nct222
37. Rowley et al (2012) Comparative international analysis of radiofrequency exposure surveys of mobile communication radio base stations. J Exposure Sci Environ Epidemiol 22(3):304–315. https://doi.org/10.1038/jes.2012.13
38. Ramirez-Vazquez et al (2023) Measurement studies of personal exposure to radiofrequency electromagnetic fields: a systematic review. Environ Res 218(114979). https://doi.org/10.1016/j.envres.2022.114979
39. Jalilian et al. Public exposure to radiofrequency electromagnetic fields in everyday microenvironments: an updated systematic review for Europe. Environ Res. 176(108517). https://doi.org/10.1016/j.envres.2019.05.048
40. Chiaraviglio et al (2022) Six months in the life of a cellular tower: is 5G exposure higher than pre-5G one? In: 2022 IEEE international symposium on measurements & networking (M&N). 18–20 July 2022. https://doi.org/10.1109/MN55117.2022.9887670
41. Selmaoui et al (2021) Exposure of South Korean population to 5G mobile phone networks (3.4–3.8 GHz). Bioelectromagnetics 42(5):407–414. https://doi.org/10.1002/bem.22345
42. Conil et al (2021) In-situ evaluation of exposure induced by 5G antennas in the 3.4–3.8 GHz band. Comptes Rendus Physique 22(S1):3–13. https://doi.org/10.5802/crphys.65
43. Aerts et al (2021) In situ assessment of 5G NR massive MIMO base station exposure in a commercial network in Bern, Switzerland. Appl Sci 11(8):3592. https://doi.org/10.3390/app11083592
44. Adda et al (2022) Assessment of population exposure to electromagnetic fields due to systems that implement dynamic spectrum sharing between 4G and 5G: definition of methods and on-site measurements. In: 2022 IEEE international workshop on metrology for living environment (MetroLivEn), pp 263–268. https://doi.org/10.1109/MetroLivEnv54405.2022.9826981
45. Werner et al (2019) A comparison between measured and computed assessments of the RF exposure compliance boundary of an in-situ radio base station massive MIMO antenna. IEEE Access 7(170682–170689). https://doi.org/10.1109/ACCESS.2019.2955715
46. Chiaraviglio et al (2021) "Pencil beamforming increases human exposure to electromagnetic fields": true or false? IEEE Access 9(25158–25171). https://doi.org/10.1109/ACCESS.2021.3057237
47. Kuehn et al (2020) Total local dose in hypothetical 5G mobile networks for varied topologies and user scenarios. Appl Sci 10(17):5971. https://doi.org/10.3390/app10175971
48. Castellanos et al (2022) Multi-objective optimisation of human exposure for various 5G network topologies in Switzerland. Comput Netw 216(109255). https://doi.org/10.1016/j.comnet.2022.109255
49. Bornkessel et al (2007) Determination of the general public exposure around GSM and UMTS base stations. Radiat Prot Dosim 124(1):40–47. https://doi.org/10.1093/rpd/ncm373
50. Chiaraviglio et al (2021) Do dense 5G networks increase exposure to electromagnetic fields? Proc IEEE 109(12):1880–1887. https://doi.org/10.1109/JPROC.2021.3125528
51. Elbasheir et al (2023) Multi-technology multi-operator site sharing: compliance distance analysis for EMF exposure. Sensors 23(3):1588. https://doi.org/10.3390/s23031588

52. Malandrino et al (2022) Performance and EMF exposure trade-offs in human-centric cell-free networks. In: 20th international symposium on modeling and optimization in mobile, ad hoc, and wireless networks (WiOpt), pp 377–382. 19–23 Sept 2022. https://doi.org/10.23919/WiOpt56218.2022.9930537
53. Chen et al (2023) Joint uplink and downlink EMF exposure: performance analysis and design insights. IEEE Trans Wirel Commun. https://doi.org/10.1109/TWC.2023.3244155
54. Van Wyk et al (2019) Measurement of EMF exposure around small cell base station sites. Radiat Prot Dosim 184(2):211–215. https://doi.org/10.1093/rpd/ncy201
55. Mazloum et al (2019) RF-EMF exposure induced by mobile phones operating in LTE small cells in two different urban cities. Ann Telecommun 74(35–42). https://doi.org/10.1007/s12243-018-0680-1
56. Velghe et al (2021) Protocol for personal RF-EMF exposure measurement studies in 5th generation telecommunication networks. Environ Health 20(1):36. https://doi.org/10.1186/s12940-021-00719-w
57. Tataria et al (2021) 6G wireless systems: vision, requirements, challenges, insights, and opportunities. Proc IEEE 109(107):1166–1199. https://doi.org/10.1109/JPROC.2021.3061701
58. Mazloum et al (2020) XmobiSensePlus: an updated application for the assessment of human exposure to RF-EMFs. In: 2020 XXXIIIrd general assembly and scientific symposium of the international union of radio science. 29 August 2020–5 September 2020. https://doi.org/10.23919/URSIGASS49373.2020.9232310
59. Boussad et al (2022) Longitudinal study of exposure to radio frequencies at population scale. Environ Int 162:107144. https://doi.org/10.1016/j.envint.2022.107144
60. Zeleke et al (2019) Radiofrequency electromagnetic field exposure and risk perception: a pilot experimental study. Environ Res 170:493–499. https://doi.org/10.1016/j.envres.2018.12.058
61. Framework for developing health-based EMF standards. World Health Organization (WHO). 14 April 2006. https://www.who.int/publications/i/item/9241594330. Accessed 2 Jan 2023
62. Policies, guidelines, regulations and assessments of human exposure to radio- frequency electromagnetic fields: output report on ITU-D question 7/2 for the study period 2018–2021. International Telecommunications Union (ITU) 71. 2021. https://www.itu.int/en/myitu/Publications/2021/10/01/08/44/Policies-guidelines-of-human-exposure-to-radio-frequency-electromagnetic-fields. Accessed 2 Jan 2023
63. Colombi et al (2022) Implications of ICNIRP 2020 exposure guidelines on the RF EMF compliance boundary of base stations. Front Commun Netw 3. https://doi.org/10.3389/frcmn.2022.744528
64. Dhungel et al (2015) Risk management policies and practices regarding radio frequency electromagnetic fields: results from a WHO survey. Radiat Prot Dosim 164(1–2):22–27. https://doi.org/10.1093/rpd/ncu324
65. Sheppard et al (2008) Quantitative evaluations of mechanisms of radiofrequency interactions with biological molecules and processes. Health Phys 95(4):365–396. https://doi.org/10.1097/01.HP.0000319903.20660.37
66. Kowalczuk et al (2010) Absence of nonlinear responses in cells and tissues exposed to RF energy at mobile phone frequencies using a doubly resonant cavity. Bioelectromagnetics 31(7):556–565. https://doi.org/10.1002/bem.20597
67. Dongus et al (2022) Health effects of WiFi radiation: a review based on systematic quality evaluation. Crit Rev Environ Sci Technol 52(19):3547–3566. https://doi.org/10.1080/10643389.2021.1951549
68. Vijayalaxmi et al (2021) Improving the quality of radiofrequency bioeffects research: the need for a carrot and a stick. Radiat Res 196(4):417–2022. https://doi.org/10.1667/RADE-21-00079.1
69. Simkó et al (2016) Quality matters: systematic analysis of endpoints related to “cellular life” in vitro data of radiofrequency electromagnetic field exposure. Int J Environ Res Public Health 13(7):701. https://doi.org/10.3390/ijerph13070701

70. Huss et al (2007) Source of funding and results of studies of health effects of mobile phone use: systematic review of experimental studies. Environ Health Perspect 115(1):1–4. https://doi.org/10.1289/ehp.9149
71. Vijayalaxmi et al (2019) Funding source, quality of publications and outcome in genetic damage in mammalian cells exposed to non-ionizing radiofrequency fields. Radiat Res 192(4):353–362. https://doi.org/10.1667/RR15364.1
72. Baan et al (2011) Carcinogenicity of radiofrequency electromagnetic fields. Lancet Oncol 12(7):624–626. https://doi.org/10.1016/S1470-2045(11)70147-4
73. Wiedemann et al (2014) Do people understand IARC's 2B categorization of RF fields from cell phones? Bioelectromagnetics 35(5):373–378. https://doi.org/10.1002/bem.21851
74. Röösli et al (2019) Brain and salivary gland tumors and mobile phone use: evaluating the evidence from various epidemiological study designs. Annu Rev Public Health 40(1):221–238. https://doi.org/10.1146/annurev-publhealth-040218-044037
75. Deltour et al (2022) Time trends in mobile phone use and glioma incidence among males in the Nordic countries, 1979–2016. Environ Int 168(107487). https://doi.org/10.1016/j.envint.2022.107487
76. International Commission on Non-Ionizing Radiation Protection (ICNIRP) (2020) ICNIRP note: critical evaluation of two radiofrequency electromagnetic field animal carcinogenicity studies published in 2018. Health Phys 118(5):525–532. https://doi.org/10.1097/hp.0000000000001137
77. Kuhne et al (2020) Thermoregulatory stress as potential mediating factor in the NTP cell phone tumor study. Bioelectromagnetics 41(6):471–479. https://doi.org/10.1002/bem.22284
78. Marques et al (2019) Advisory Group recommendations on priorities for the IARC monographs. Lancet Oncol 20(6):P763–P764. https://doi.org/10.1016/S1470-2045(19)30246-3
79. Toledano et al (2018) An international prospective cohort study of mobile phone users and health (COSMOS): factors affecting validity of self-reported mobile phone use. Int J Hyg Environ Health 221(1):1–8. https://doi.org/10.1016/j.ijheh.2017.09.008
80. Ahn et al (2022) An international collaborative animal study of the carcinogenicity of mobile phone radiofrequency radiation: considerations for preparation of a global project. Bioelectromagnetics 43(4):218–224. https://doi.org/10.1002/bem.22407
81. Samet et al (2020) The IARC monographs: Updated procedures for modern and transparent evidence synthesis in cancer hazard identification. JNCI 112(1):30–37. https://doi.org/10.1093/jnci/djz169
82. Verbeek et al (2021) Prioritizing health outcomes when assessing the effects of exposure to radiofrequency electromagnetic fields: a survey among experts. Environ Int 146:106300. https://doi.org/10.1016/j.envint.2020.106300
83. Opinion on potential health effects of exposure to electromagnetic fields (EMF). Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR) 288. 27 January 2015. Available at: https://ec.europa.eu/health/scientific_committees/consultations/public_consultations/scenihr_consultation_19_en. Accessed 7 Jan 2023
84. Outhwaite et al (2022) Agriculture and climate change are reshaping insect biodiversity worldwide. Nature 605(97–102). https://doi.org/10.1038/s41586-022-04644-x
85. Nath et al (2022) Correlation does not imply causation: decline of house sparrow overshadowed by electromagnetic radiation. Urban Ecosyst 25:1279–1295. https://doi.org/10.1007/s11252-022-01227-6
86. Karipidis et al (2021) What evidence exists on the impact of anthropogenic radiofrequency electromagnetic fields on animals and plants in the environment? A systematic map protocol. Environ Evid 10(1):39. https://doi.org/10.1186/s13750-021-00252-w
87. Pophof et al (2023) Biological effects of radiofrequency electromagnetic fields above 100 MHz on fauna and flora: workshop report. Health Phys 124(1):31–38. https://doi.org/10.1097/HP.0000000000001625
88. Thielens et al (2020) Radio-frequency electromagnetic field exposure of Western Honey Bees. Sci Rep 10(1):461. https://doi.org/10.1038/s41598-019-56948-0

89. Woods et al (2005) Honeybee flight metabolic rate: does it depend upon air temperature? J Exp Biol 208(6):1161–1173. https://doi.org/10.1242/jeb.01510
90. IEC 62232:2022 ed3.0 determination of RF field strength and SAR in the vicinity of radiocommunication base stations for the purpose of evaluating human exposure. International Electrotechnical Commission (IEC) 342. 14 October 2022. https://webstore.iec.ch/publication/64934
91. Keller (2019) On the assessment of human exposure to electromagnetic fields transmitted by 5G NR base stations. Health Phys 117(5):541–545. https://doi.org/10.1097/hp.0000000000001089
92. Colombi et al (2022) Methods and practices for *in situ* measurements of RF EMF exposure from 5G millimeter wave base stations. IEEE Trans Electromagn Compat 64(6):1986–9113. https://doi.org/10.1109/TEMC.2022.3212543
93. Colombi et al (2020) Analysis of the actual power and EMF exposure from base stations in a commercial 5G network. Appl Sci 10(15):5280. https://doi.org/10.3390/app10155280
94. Thors et al (2017) Time-averaged realistic maximum power levels for the assessment of radio frequency exposure for 5G radio base stations using massive MIMO. IEEE Access. https://doi.org/10.1109/ACCESS.2017.2753459
95. Bechta et al (2022) Analysis of 5G base station RF EMF exposure evaluation methods in scattering environments. IEEE Access:7196–7206. https://doi.org/10.1109/ACCESS.2022.3142673
96. Törnevik et al (2020) Time averaged power control of a 4G or a 5G radio base station for RF EMF compliance. IEEE Access 8:211937–211950. https://doi.org/10.1109/ACCESS.2020.3039365
97. Mahfouz et al (2012) Influence of traffic variations on exposure to wireless signals in realistic environments. Bioelectromagnetics 33(4):288–297. https://doi.org/10.1002/bem.20705
98. Hardell et al (2020) Appeals that matter or not on a moratorium on the deployment of the fifth generation, 5G, for microwave radiation. Mol Clin Oncol 12(3):247–257. https://doi.org/10.3892/mco.2020.1984
99. Bruns et al (2020) ‘Corona? 5G? Or both?’: the dynamics of COVID-19/5G conspiracy theories on Facebook. Media Int Aust 177(1):12–29. https://doi.org/10.1177/1329878X20946113
100. Adversarial narratives: a new model for disinformation. Global Disinformation Index (GDI) 24. August 2019. https://www.disinformationindex.org/blog/2019-8-1-adversarial-narratives-are-the-new-model-for-disinformation/. Accessed 7 Jan 2023
101. Lewandowsky et al (2020) The Debunking handbook 2020. https://sksto/db2020. Accessed 26 Apr 2023
102. Communicating fact checks online. Vicol. Africa Check, Chequeado and Full Fact. February 2020. https://fullfact.org/about/research/. Accessed 26 Apr 2023
103. Jolley et al (2020) Pylons ablaze: examining the role of 5G COVID-19 conspiracy beliefs and support for violence. Br J Soc Psychol 59(3):628–640. https://doi.org/10.1111/bjso.12394