HORIZON 2020 AND BEYOND

On the 5G Operating System for a True Digital Society

David Soldani and Antonio Manzalini

The advanced fifth-generation (5G) infrastructure will not only be a sheer evolution of the current network generations but, more significantly, a revolution in the information and communication technology (ICT) field. The 5G technology will efficiently enable new secure, dependable, ultrareliable, and delay-critical services to everyone and everything, such as cognitive objects and cyberphysical systems (CPSs). A fully immersive experience and anything as a service are the primary drivers for a global adoption and market uptake of new technology components, beyond today's client–server model, where the network has been reduced to a mere pipe of bits. The network will become the nervous system of the true digital society and economy. This article gives fundamental insight into how the 5G mobile communications system is being designed to be powerful and, especially, flexible enough, thus meeting the foreseen and unknown traffic scenarios and services requirements. We also present how a massive adoption and exploitation of mobile-edge computing (MEC), software-defined networking (SDN), network functions virtualization (NFV), and services virtualization will make the 5G operating system (OS) feasible and business viable.

Advanced 5G Infrastructure

The advanced 5G infrastructure, defined as the ubiquitous ultrabroadband network supporting the future Internet (FI), is not only about new releases of current network generations, but, more drastically, it is expected to be a revolution in the ICT field. The network will efficiently and effectively enable newfangled services to everyone and everything, such as cognitive objects and CPSs [1].

As shown in Figure 1, the network infrastructure is expected to become the nervous system of the actual digital society, and the 5G OS is the distributed software that runs on top and implements an OS mainly oriented to cognitive objects networking. For example, the 5G OS will allow a server in a data center (DC) to administer data, services, security, and any other networking function of connected objects (terminals, machine, robots, drones, and so on) and networks, as applications.

This challenge calls for a complete redesign of services, architectures, interfaces, functions, access and nonaccess strata protocols, and related procedures, as well as advanced algorithms, e.g., for identity and mobility management as well as the setup, maintenance, and reconfiguration of services and any type of ICT resource.

The expected transformation will be especially true at the edge, i.e., around the end user (or the prosumer), where the intelligence already started migrating a few years ago and where massive processing, memory, and storage capacity are gradually accumulating.

As of today, many challenges are still to be addressed to meet the expected key performance indicators, e.g., in terms of throughput (1,000 times more in aggregate and ten times more at link level), latency (1 ms for remote control of robots, or tactile Internet applications, and below 5 ms for a 2–8-K change in view at 30–50 Mb/s), coverage (seamless experience), battery lifetime (ten times longer), and quality of service (QoS).

This article introduces the fundamental peculiarities of the advanced 5G infrastructure and proposes a solution for unified connection, security, mobility, and routing management, end to end. We also detail how the massive adoption and utilization of SDN, services and NFV, and MEC will make the 5G infrastructure technically feasible and, especially, business viable.

Many research and innovation actions are still ongoing at a global level. Currently, most of the effort is being placed on research; after that, intensive standardization activities and large field test trials will follow to accelerate the industrial preadoption; commercial products will most probably be available on the market after 2020. This approximated road map applies to 5G networks and new devices for human and, especially, machine type of traffic (MTC).

Network, Services, and Service Capabilities

The 5G capabilities and recommendations from the global stakeholders are shown in Figure 2. International Mobile Telecommunications (IMT)-Advanced [fourth-generation (4G)] Third-Generation Partnership Project long-term evolution (3GPP LTE) was designed for improving the capacity, user data rates, spectrum usage, and latency with respect to IMT-2000 (3G, 3GPP Universal Mobile Telecommunications System). IMT for 2020 and beyond (future IMT, 5G) will be, above all, suitable for massive and mission-critical machine communication. As shown in Figure 2, the key performance metrics 5G improves are, especially, in terms of latency, reliability, speed, mobility, and spectrum usage. [2].

Looking at spectrum utilization (see Figure 3), frequencies below 6 GHz (cellular band) are mostly suitable for
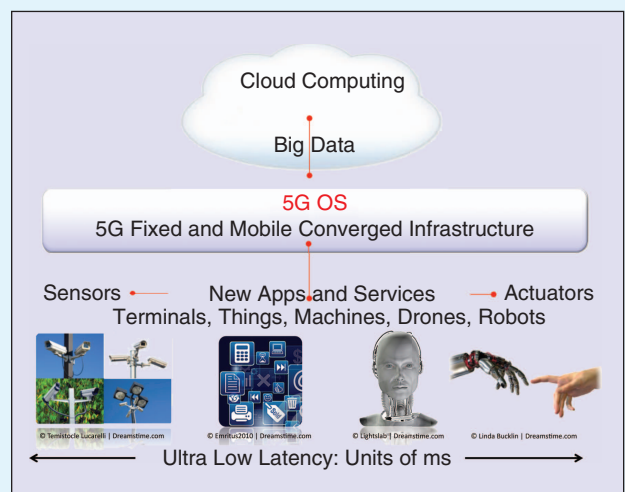

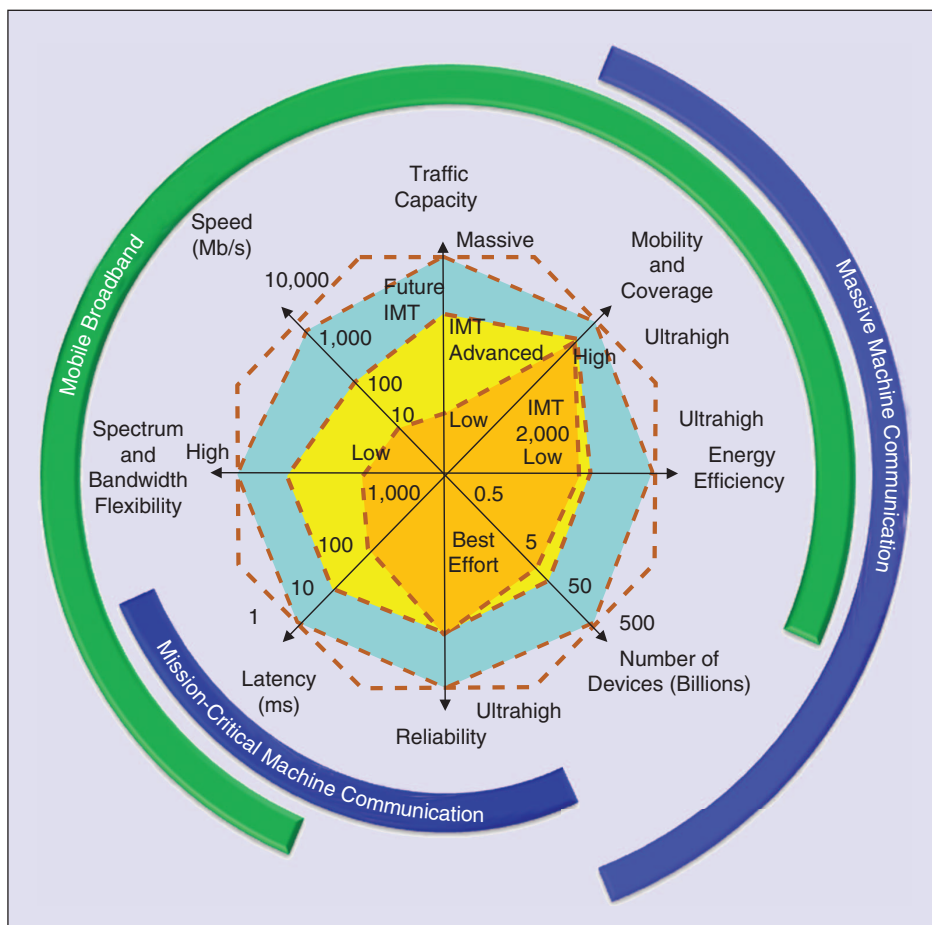**FIGURE 1** The new ecosystems based on 5G OS. (© Dreamstime.)

FIGURE 2 The 5G capabilities: a global perspective [2].

macrocoverage (0.5–2-km radius). The cellular band is expected to increase at least twice as much from what it is today, i.e., from 300 to 984 MHz should be allocated already at the 2015 World Radiocommunication Conference (WRC-15) [3]. In the range of 6–30 GHz, about 2.5 GHz could be made available for microcoverage (i.e., within a 50–100-m radius). Frequencies from 30 to 90 GHz (visible light) would be particularly suitable for fronthauling and backhauling as well as local deployments (i.e., within a 10-m radius). In this range, about 40 GHz could be allocated for massive machine

FIGURE 3 The spectra availability and possible usage for 5G deployment [3], [4].

communications. These carrier bandwidths and spectrum at higher frequencies need to be identified at WRC-18/19 [4].

Vision

Today, for the first time in history, we are witnessing the convergence of cloud computing, terminal computing power, and connectivity at high speed into a single point: the smartphone, which was realized over the bit pipe networks of the Internet service provider denoted in this article as the client–server model. As Google said in 2010, "the smartphone is the extension of what we do and what we are: the mobile is the answer to pretty much everything" [5].

It is now time to go beyond this model and design the next generation of ubiquitous ultrahigh-broadband infrastructure [6], denoted as 5G, which will support the FI and provide delay-critical and ultrareliable, secure, and dependable services to billions of smart objects and CPSs such as cars, robots and drones, and new mobile terminals.

This will bring about the fundamental shift in paradigm from the client–server model to the new concept of neural bearer or bearer graph—not limited to the single-dimensional case of end-to-end connectivity—efficiently enabled by the 5G open and flexible infrastructure (5G OS), where anything or everything may be offered as a service (XaaS). XaaS refers to those services—beyond the curr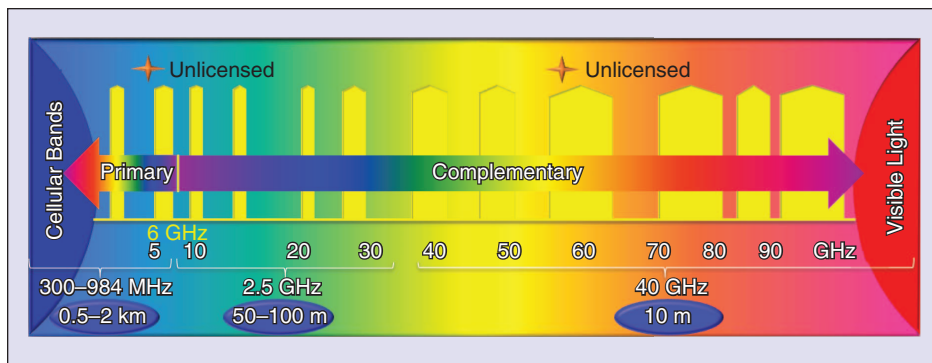ent software platform and infrastructure (SPI) model (software as a service, platform as a service, and infrastructure as a service) of cloud computing—such as data as a service, security as a service, knowledge as a service, machine as a service, robot as a service (RaaS), and so forth, that could be delivered over the 5G infrastructure without the need to own hardware, software, or the cognitive objects themselves [1], [7], [8].

With 5G, ICT will find application to generate new services at a low cost for improving our lives and not only for communicating. Most communication services will become commodities that are free of charge and, looking at the innovative business models, will monetize those applications, machines, and things that will be offered as a service, meaningful to what really matters to us for achieving a true information-oriented society, i.e., a digital society and economy [1].

A high-level vision of networks and services at the horizon 2020 and beyond is shown in Figure 4. Unlike the previous generations of wireless mobile networks, 5G will allow the use of any authorized spectrum and any access technology through a software-defined and virtualized architecture. In other words, the service delivery will not be limited to any specific frequency band; rather, it will follow the optimal delivery over the best available spectra. In such an innovative paradigm, all access technologies may be exploited as a single set of radio access capabilities. All of these capabilities will be enabled by the new regime of software-defined air interface supported by cloud computing-based and virtualized radio access networks. As shown in Figure 4, under the assumption of conventional network planning for a fully immersive and augmented reality, we envision the following wireless network capabilities:

1) 50-Gb/s macrocoverage
2) 100-Gb/s microlocal coverage
3) 80-Gb/s E-band back-/fronthauling links.

The redesign of the radio access nodes will require innovation in multiple areas of basic radio technologies, such as new air interfaces, new virtualized radio access networks, new radio-frequency transceiver architecture, and new device radio architecture. New radio backhaul and new fiber access for the fixed network to support 5G wireless are also required as an integral part of the solution. Most of the intelligence is expected to be placed and orchestrated at the edge of the network, i.e., in the aggregation and access segments up to the end user premises or wireless access points [9].

Furthermore, in the last mile, an enormous quantity of processing and storage capability will be accumulated, thus ensuring better levels of QoS. In other words, virtualized network functions and services and related states will be optimally located around the prosumer (machine or human being) to efficiently deliver the expected experience and meet the target performance, which is 1,000 times higher wireless area capacity and 10-Gb/s wireless link speed for a truly immersive experience; hyperconnectivity to trillions of things and, especially, five times lower end-to-end latency than 4G (1 ms is the target for tactile Internet); and 90% energy savings per provided service, as highlighted and expected in [6].

The network services and service capabilities [10], as well as access stratum (AS) and non-AS (NAS) protocols [11], will, accordingly, evolve for carrying delay-critical, ultrareliable,
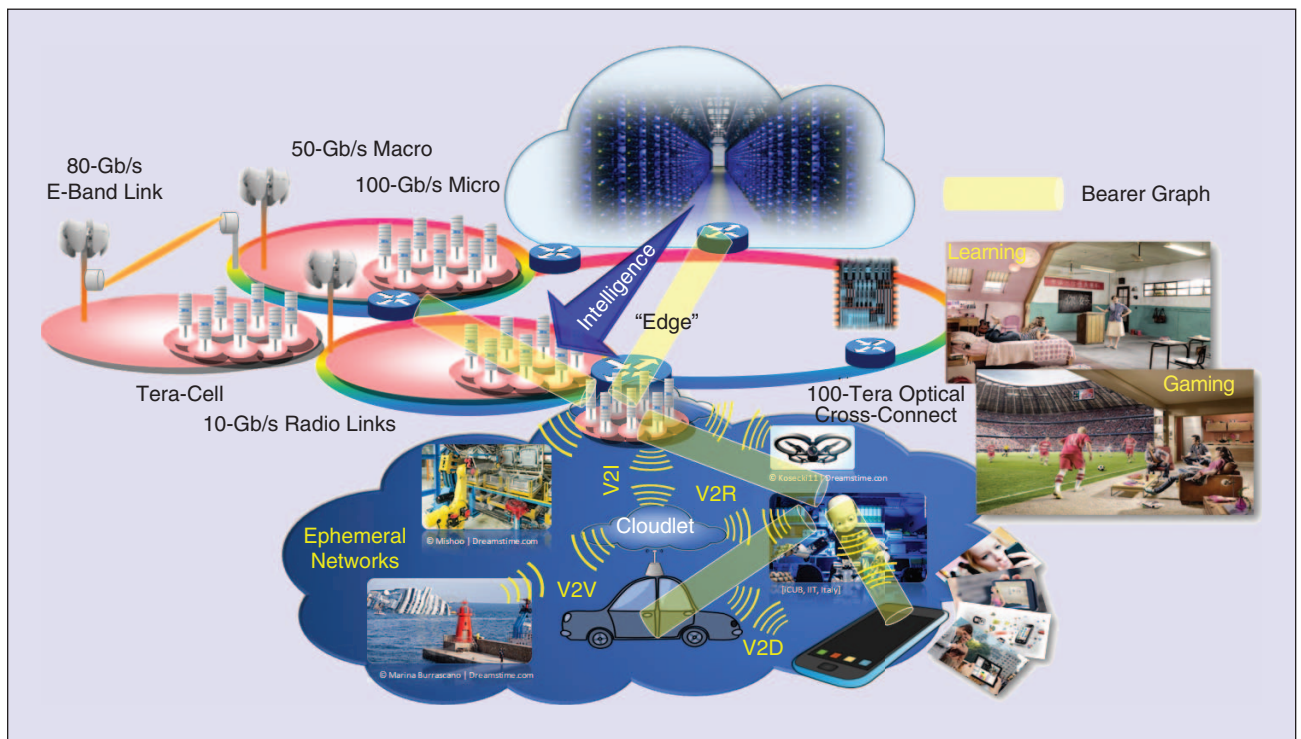

FIGURE 4 The network and services capabilities at horizon 2020 and beyond (V2D: vehicle to device and V2I: vehicle to infrastructure). (© Dreamstime.)

secure, privacy-preserving, and dependable services, as a fully immersive [three-dimensional (3-D)] experience and mission-critical vehicle-to-vehicle (V2V) communications.

The concept of the bearer service—a set of quality/capability parameters currently defining the single virtual pipe between protocol entities in the mobile terminal and related peers in the gateway to external packet data networks—will be replaced by a neural bearer or simply a bearer graph that will enable multidimensional carrier-grade communication paths, i.e., well beyond the scope of the current bidirectional communication chains, as shown in Figure 4. New technology enablers for establishing, maintaining, and reconfiguring the bearer graph connecting multiple cognitive mobile objects—forming ephemeral networks (with and without network assistance)—will allow operators to meet the performance targets, especially in terms of reliability, latency, and speed, which are in fact unreachable implementing today's and the upcoming 3GPP mobile services and SDN solutions as they are limited to the transport and network layers of the infrastructure [12].

As shown in Figure 4, the robot, as well as other cognitive objects, is expected to be identified and handled as a new device type with its own capabilities. Control entities for establishing, maintaining, and reconfiguring the new bearer service in dual-mode mobility are distributed in the network as well as in other vehicles around the robot in question, i.e., at the edge. In this case, the head of the robot is connected to a car, a device, and three peer entities in the network, including the Internet. The vehicle (robot)-to-X networking, denoting as X any other fixed or mobile node, consists of a local, opportunistic, and multihop communication with direct connections among objects whenever advantageous and possible.

The proposed concept goes well beyond the current 3GPP device-to-device (D2D) proximity services, which are limited to a single hop and typically rely on assistance from the network infrastructure, and the well-known mobile ad hoc networks, which are based on multihop routing with limited network performance [13], [14]. The proposed neural bearer may enable several carrier-grade communications, i.e., layer 1–7 links, simultaneously, through different radio interfaces with the availability of multiple transceivers, which is currently not possible in D2D, especially due to low battery consumption and hardware integration constraints. As shown in Figure 4, the protocols of the new NAS enable entities located in the different connected nodes to exchange information, knowledge, and any kind of service among them. In this sense, the notion of the bearer graph defines capabilities beyond what is currently possible with relay control, in in-coverage and out-of-coverage modes [13]. The service applications may run on each of the connected nodes and exchange services among themselves in a secure, reliable, dependable, and low-latency manner not limited to client–server bidirectional communications.

The 5G OS will be capable of unifying the control (connection, security, and mobility management) and forwarding planes (routing management) of the future fixed and mobile networks composed of heterogeneous devices, robots, machines, and nodes up to the cloud. The 5G OS will transform the advanced 5G infrastructure in the ultralow-latency nervous system supporting many innovative service paradigms, such as XaaS, looking at a true digital society and digital economy at horizon 2020 and beyond.

The socioeconomic and business implications of this vision are enormous. Today, the main control variables of our complex economy are still human intelligence, attention, effort, and time. Humans are the most productive part of the current economy. In fact, we are witnessing the migration of industries to regions where labor costs are much lower. This vision will help us to create a pervasive machine intelligence capable of reshaping this economy equation by taking over many cognitive tasks that humans can or cannot do and improving quality of life. Organizations should focus on harnessing and leveraging 5G technologies around machine intelligence, big data, and connected objects and at the same time create new different jobs. We believe the advanced 5G infrastructure to be the most important catalyst of the second machine age [15]: 5G will power intelligent machines to flood the landscape of new jobs. There will be a number of socioeconomic benefits, including the reduction of human efforts in jobs subject to computerization and robotization, aimed at bringing down operating costs with much higher product quality, worker safety, and improved operational conditions; increasing local production; reducing long-distance transportation; and optimizing many socioeconomic processes. As a result, human labor costs will no longer drive investments and the number of jobs created will be far greater than the numbers lost due to automation [16].

The network will be transformed from a sterile pipe to the nervous system of the digital society and economy. Network operators, leveraging their edge assets, will have a chance to play new business models and will be in a much stronger position in the value chain, e.g., opposed to the over the top, which will remain most likely confined to the client–server model.

Last, but not least, to rapidly profit from 5G, first the environment around cognitive mobile objects needs to evolve to host them. For instance, a robot does not need to recognize that there is an object in front of it. The entity in the proximity may alert the robot to its presence instead, and the robot may retrieve all necessary information on it from the edge cloud. This largely simplifies the machine-learning process as robots would just need to compute the geometry of the physical environment around them. This is shown in Figure 5, where the 5G fundamental technology components are shown:

1) *Sensing*: massive amount of sensors.
2) *Rendering/actuating*: new types of devices.

3) *Edge computing*: storage, memories, and computational power for a zero-latency cyberworld, where objects can be digitally described and meet each other.

4) *Control and orchestration*: intelligent functions at the edge, pervasive and ephemeral.

5) *Networking*: high-speed connectivity via new waveforms and optical fiber carrying digital/modulated signals.

Ultimately, as it was soundly presented in [17] for car-to-car communications, we also argue that cooperation among many stakeholders (automotive, robotics, vendors, carriers, and related suppliers) is definitely required to successfully introduce the new 5G enabling technologies.

Key Enabling Technologies

The new air interface is expected to be flexible and suitable for a broad range of frequencies. Latency, reliability, speed, and interference control of the access stratum protocols need to be substantially improved to support the envisioned scenarios and other unforeseen performance requirements. Transmission time interval (TTI) and subcarrier spacing should adapt flexibly to different multipath channel conditions (e.g., delay and Doppler) at a high frequency and speed, especially in the case of V2V communications and dual-mode mobility. The TTI is expected to be less than 1 ms (the current minimal length) with modulation and coding schemes to cope with short packets and fragmented spectra. At the time of writing, filter bank-based multicarrier (FBMC) and universal-filtered multicarrier (UFMC) are good examples of enabling technologies for both out-of-band interference control and efficient utilization of narrow frequency bands. They allow partitioning of spectrum into independent bands with relaxed requirements for synchronization and present excellent capabilities for coexistence of services in the same frequency band and spectrum sharing [12].

Other relevant techniques from [12] are: 1) nonorthogonal multiple access (NOMA) and 2) sparse code multiple access (SCMA). Both methods allow for spectrum overload, where a number of devices may be served simultaneously as they are no longer bounded to a set of orthogonal resources. NOMA and SCMA are particularly suitable for uplink massive connectivity, wireless backhauling for moving networks, and ultra-dense networks.

Beyond this, beam-forming and multiple-input, multiple-output (MIMO) techniques will certainly improve coverage and reduce interference. For instance, 3-D massive MIMO—consisting of large planar antenna arrays—supports beam-forming and spatial combining capabilities at the transmitter and receiver. Furthermore, the MIMO principle may be applied to distributed systems of cooperative antennas with the possibility of superimposing waveforms and creating spatial interference patterns, thus achieving much higher gains from beam forming.

A possible architecture for implementing the bearer graph concept introduced in the "Vision" section is shown in Figure 6. Three levels of controllers are defined in the proposed model.

■ *Device controller*: It is responsible for the selection of fixed and mobile radio access technologies that the terminal supports subject to the limitation imposed at the edge or by one of the controllers in the upper layers. It is located locally in the device.

■ *Edge controller*: It performs delay-constraint tasks, such as radio resources management and adaptive layer 1–2 functions, essentially (hybrid automatic repeat request) and modulation and coding schemes selection. Wireless packet scheduling and handover control (active mode mobility) are also performed at this level. It handles end-to-end QoS provisioning, session/route establishment, and service chaining, mobility management, policy, and charging. It is responsible for idle and active mode mobility management, including handoffs, and layer 1–3 routing and forwarding (establishing, maintaining, and improving the routes among nodes for the entire communication duration). In the case of out-of-coverage and/or in relay control, it also performs
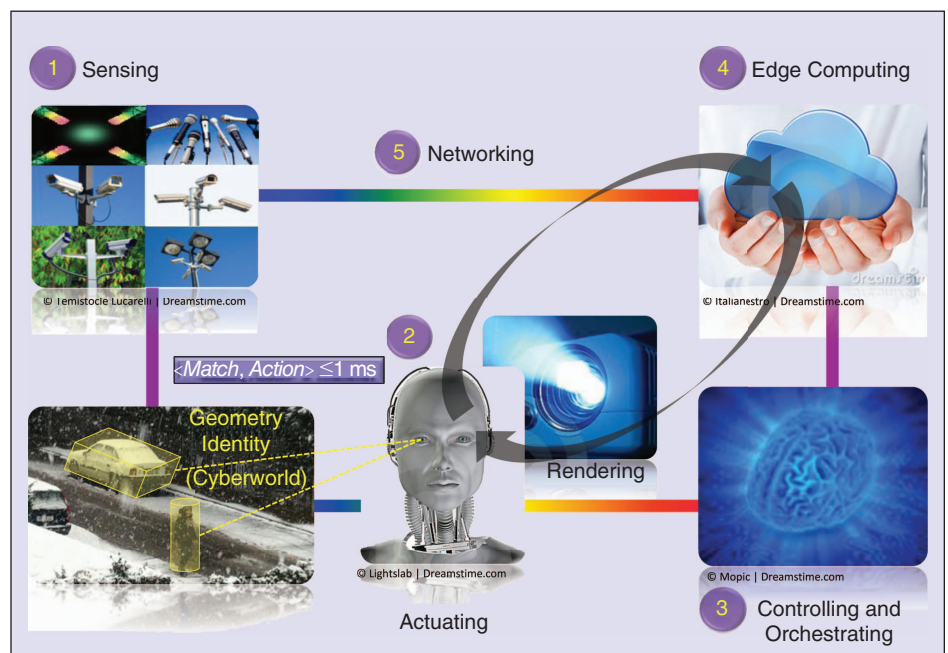

FIGURE 5 The fundamental technology components. (© Dreamstime.)

cluster-head tasks in transparent and nontransparent mode (decentralized control). It includes device access and nonaccess strata functions for identity, session, and mobility control. It may be located at the edge fabric, specifically, minidata centers (cloudlets) up to aggregation networks, such as smart kiosks, stores, stations, airports, and cars.

- *Orchestration controller*: It enables coordination of cloud computing resources and distributed networking. A set of centralized control logics is responsible for efficient resources allocation and service composition at multiple levels of substrate abstraction. Performed tasks may go from resource requests to embedding of links, memory, storage, and computational capacity to optimization of locations of network functions and corresponding states across virtual machines (VMs). In practice, it handles the different steps involved in the provisioning of virtual functions and services, such as creating, moving, and shutting down VMs in the virtual distributed infrastructure as well as installing, configuring, monitoring, running, and stopping software in the VMs. It is located in the cloud.

The control of multihop paths may be based on tables, following the <*Match, Action*> principle, analog to SDN using OpenFlow [9]. The translation function may be performed at the edge: tell me what you want, not how to do it, which can be done later in a proactive or reactive manner. For example, in the case of mobility, handover decisions are pushed down to lower-level controllers, and, at the same time, the associated route between the moving entities is updated. Here, routing and location tracking updates may be reactive or proactive, depending on the desired grade of service, as proposed in [18]. When reactive, the controller establishes a new route for the flow only after a handover is completed. In the proactive mode, routing decisions may be formulated a priori using intelligent algorithms, which may consider the history of the journey, context information, and long-term mobility behaviors at that network edge. The internal packet forwarding may be optimized through the orchestration function [19].

The three levels of controllers present interfaces to the outer world for any type of service or network application. In Figure 6, for instance, sensors located in mobile objects are used for the sound and light field and allow a 3-D experience to be reproduced at 30–50 Mb/s/view.

To achieve the target performance improvements, the proposed logical architecture may run over a converged optical-wireless infrastructure for network backhauling and fronthauling and flexible spectrum usage carrying digital and even modulated radio signals over fibers [20].

Another possible approach would be to introduce an active remote node between the wireless access points and core network that would bypass the metropolitan area network through adaptive ultralong-reach links, as proposed in [21].

Virtualization of resources will allow the coexistence of a number of logical architectures, dynamically self-adapting for fitting business demands and operational objectives on the same physical core infrastructure.

In the following paragraphs, we elaborate briefly on how this architecture would support the necessary capabilities of a machine in the case of RaaS.

The main problem is how to design the decision-making processes in the robot, i.e., the brain, which is actually an issue for any intelligent machine providing any service. Complex and articulated tasks cannot be performed locally using the processing power of the robot because it would immediately drain the battery. It is therefore necessary to complement the local processing power with the computing and storage capabilities of data centers, where big data (e.g., data captured by sensors) are stored and then analyzed to infer decisions to local actuators.
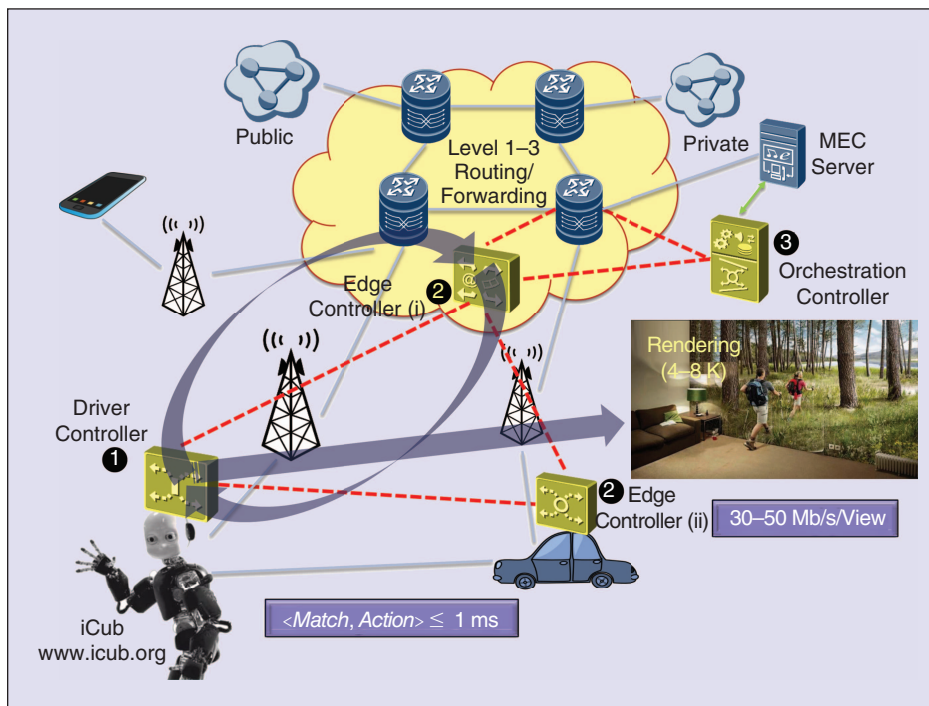

FIGURE 6 A proposed logical architecture for meeting the key 5G performance requirements, especially in terms of reliability and latency.

Therefore, we can immediately realize that a key requirement will be that the network latency, in interconnecting the local robot to the data center, should be extremely low. This is especially true for motion control beyond computer vision, as the network connection needs to respond with a very short dormancy (order of units of ms) for the robot or, in general, the intelligent vehicle to react in time. The brain of the robot could be modeled using three different levels of reactive intelligence [8]

- *automatic reactions*: fast and predefined actions for specific local contexts, which could be designed and implemented, e.g., by means of simple rules *<Match, Action>* deployed into local processing units
- *autonomic reactions*: actions cascaded by means of data analytics systems/methods, which imply the elaboration of a large amount of local information and may be enforced via unsupervised learning capabilities; these intelligent tasks will be performed at the edge data centers
- *orchestrated behavior*: actions mainly based on data analytics systems and methods and with the elaboration of a very large amount of significant data coming from different local contexts; these capabilities also reside in the cloud.

In this example, the three levels of intelligence of a robot match perfectly to the three levels of control (device, edge, and orchestrator) previously described. Hence, the 5G network will become the de facto nervous system of the pervasive artificial intelligence of the second machine age.

In this evolution, software will be the true challenge. In fact, the future 5G infrastructures will rely more on software, which will accelerate the pace of innovation as in the computing and storage domains, and dramatically reduce costs. This trend, together with the giant economic drive given by myriad new players entering the market, will yield the shift in paradigm capable of leading investment outside of the network infrastructure boundaries and stimulate the advent of the new paradigms proposed by this vision.

Global Scenario

Many initiatives on 5G are currently ongoing globally, and only some of them are shown in Figure 7. For example, in the United States, the three main activities carried out on 5G are: 1) the Intel Strategic Research Alliance, 2) 4G Americas, and 3) New York University's Wireless Research Center. In China, these initiatives are ongoing with the National High-Tech Research and Development 863 Research Program (Chinese: 863计划) and the IMT-2020 (5G) Promotion Group. In Japan, the 2020 and Beyond Ad-Hoc Group is under the Association of Radio Industries and Businesses's (ARIB's) Advance Wireless Communications Study Committee. In Korea, the main activity is the 5G Forum. The most important initiatives in the European Union (EU) are the 5G Private Public Partnership (5G PPP) and the 5G Innovation Centre (5GIC) at the University of Surrey, United Kingdom.

The 5G PPP is within the EU Horizon 2020—The EU framework Programme for Research and Innovation—under one of the most important EU industrial leadership challenges: ICT-14 Advanced 5G network infrastructure [6]. Within this research and innovation framework, the European Commission, under the approval of the European Parliament, has already committed €700 million of public funds over six years (2015–2021). The investment of the private parties [e.g., industry, small and medium enterprises (SMEs) and research institutes] is expected to be two to ten times higher than the public funds [6].

As shown in Figure 8, most efforts are currently being placed on research work; after that, intensive standardization activities, field tests, and large-scale trials will take place to accelerate industrial preadoption; commercial products will likely be available on the market around 2020 and beyond [4]. These approximate milestones apply to infrastructures and devices for human type of traffic and, in particular, MTC.

In summary, in the previous sections we argued strongly that ultrahigh-bandwidth connectivity, in orchestration with a high-performance and off-the-shelf hardware (processing power), has already enabled a hyperconnected world where software-defined (programmable) networks and services are the true challenge, especially for mission-critical machine communications. It is, of course, intended that future networks, relying more on software, will accelerate the pace of innovation, similarly to what has been happening in computing and storage domains, and reduce costs while maintaining, or even improving, carriers' class levels of performance. Technology is going to become accessible worldwide to all enterprises on an equal basis in any part of the world, reducing most of today's competitive advantages of hardware. This will reduce the thresholds for new players to enter the market, leading to a shift in paradigm where the main value proposition will be in software solutions.

Most likely, future transport networks will become less hierarchical, with a limited number of core optical nodes, interconnecting metropolitan areas via fibers. In turn, these areas at the edge are going to be densely populated by thousands of small aggregation nodes (with processing and storage capabilities) and by an incredible number of devices and machines assembling around users. The number of devices connected to the network is thus expected to grow exponentially. Today, every smartphone already integrates several sensors, and by the end of this decade, there will be more than 100 sensors in each device, dramatically increasing the number of connections.

Still, by far, operators tend to keep a distinction between the network and what connects to it: the mobile terminals. This distinction will gradually fade away as more and more tasks are being performed either in the network and/or in the terminals. As shown in Figure 4, at the edges, dynamic (ephemeral) virtual networks will be created out
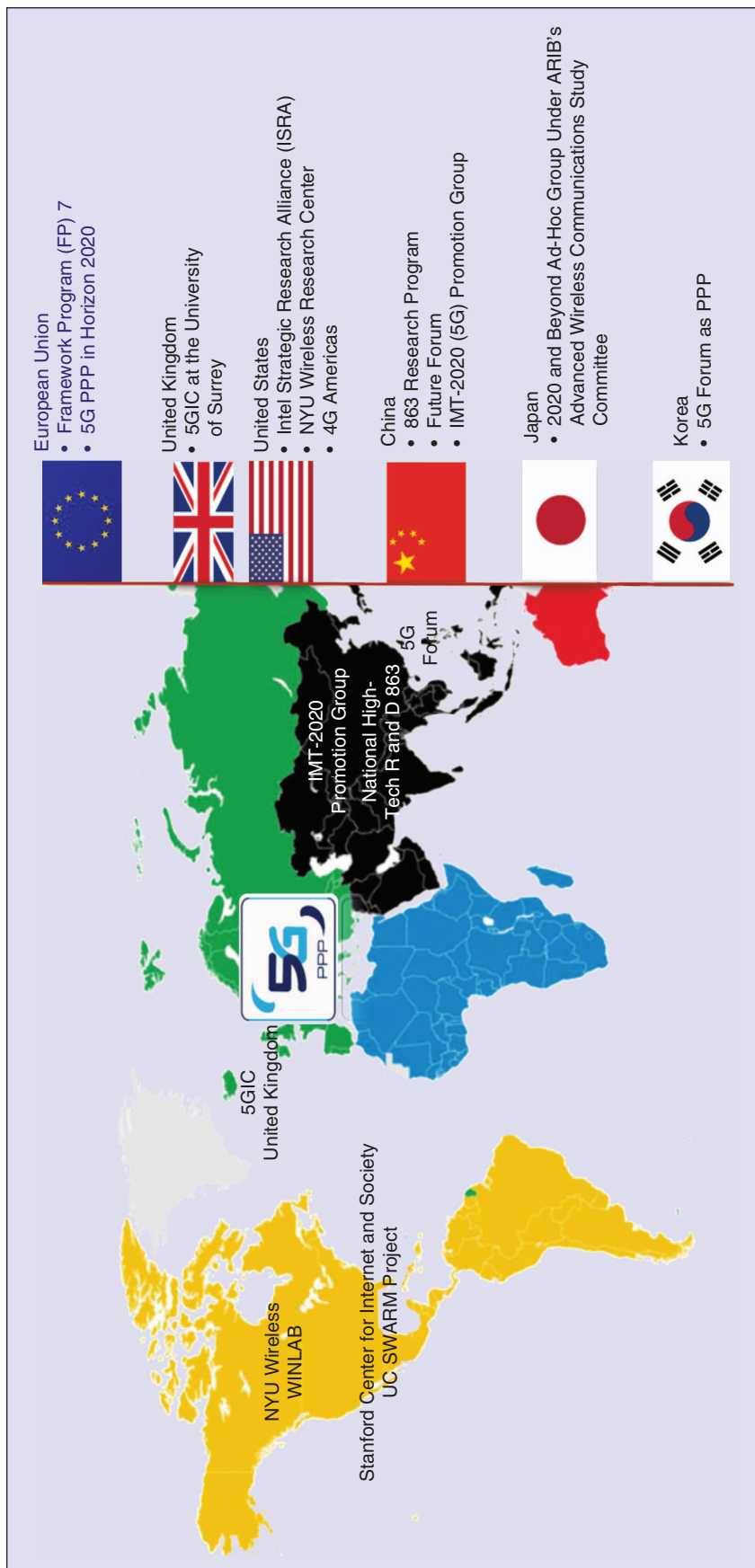
European Union
• Framework Program (FP) 7
• 5G PPP in Horizon 2020

United Kingdom
• 5GIC at the University of Surrey

United States
• Intel Strategic Research Alliance (ISRA)
• NYU Wireless Research Center
• 4G Americas

China
• 863 Research Program
• Future Forum
• IMT-2020 (5G) Promotion Group

Japan
• 2020 and Beyond Ad-Hoc Group Under ARIB's Advanced Wireless Communications Study Committee

Korea
• 5G Forum as PPP

IMT-2020 Promotion Group

National High-Tech R and D 863

5G Forum

5G PPP

5GIC United Kingdom

NYU Wireless WINLAB

Stanford Center for Internet and Society

UC SWARM Project

FIGURE 7 The main research initiatives on 5G ongoing globally.

of a variety of aggregation nodes, devices, drones, robots, and so on. These elements need to be considered as an integral part of the infrastructure. In other words, such an enormous number of nodes and elements aggregated in an application-driven way will be the edge fabric enabling the shift from the client–server paradigm to the new services paradigms, such as a 3-D fully collaborative and immersive experience without the need to wear glasses or binaural receivers, enriched by context information and anything as a service.

In particular, the edge fabric will be controlled by a few hierarchical entities (edge controllers) and actuated locally because of decisions made by autonomous entities (device controllers). For example, an edge controller could be instantiated in a mini data center located in a kiosk, and the ambient around the kiosk could be defined by consistent surroundings of dynamic collections of devices creating locally a well-defined context. All of these consistent surroundings of devices are interconnected through the 5G infrastructure (i.e., via optical/radio communication pipes), forming a much larger and sometimes overlapping new ambient, such as a smart city, a campus, or a network of malls or schools. This transformation will present a number of business implications with some risks and plenty of opportunities (disruptions) for incumbent industry players. For instance, what will happen if this transition accelerates and this technology disruption is adopted widely (even on another continent)?

A Scenario for Europe's ICT Future

Let us imagine a scenario where in each EU member state (MS), there are two or more software-defined operators (SDOs) leasing hardware resources from two or more physical infrastructure providers (PIPs).

The operators will own virtual networks and services in software platforms, i.e., network services and functions (from layer 2 to 7) fully developed, implemented, and operated in software executed in DCs. PIPs will own and make available to SDOs basic radiating elements, fibers, and raw hardware, i.e., layer 0–1 of the networks.

As a result, under this assumption, the SDOs will undeniably observe dramatic cost reductions (e.g., 40–50% savings in energy), improve efficiency in overall operations, and reduce the time to market substantially when deploying new services.

FIGURE 8 A high-level road map toward the IMT for 2020 and beyond.

In this scenario, let us now picture what would happen if SDOs provided anything as a service to all EU citizens within an EU single market through any terminal, device, machine, smart thing, robot or drone, etc., or whatever they have connected at the edge by means of ultrafast low-cost access. This would create several industrial opportunities and new fast-growing socioeconomic developments in the EU. This would be about an unparalleled industry transformation underpinning on the ongoing softwarization trend. The big incumbent players as well as small and medium enterprises (SMEs) and citizens would be positively affected, and all of European society would have the chance to benefit from the new technological landscape. Citizens would see an improvement in their quality of life, and there would be plenty of opportunities for SMEs, e.g., looking at a concrete development of smart cities, villages, houses, and so on. There would also be a massive adoption of services and applications using intelligent machines to cope with most of our current and future societal challenges, such as aging citizens' needs, energy, and pollution; the digitization of Europe's cultural heritage; or even for enabling new high-quality agri-food ecosystems, which is truly becoming a vital priority for our society.

Conclusions

In this article, we argue that the 5G OS will transform the advanced 5G infrastructure into an ultralow-latency nervous system supporting a 3-D fully collaborative and immersive experience and anything as a service for a true digital society and economy at the horizon 2020 and beyond.

Mission-critical machine communication and massive MTC are the main drivers for the wide and rapid adoption of the new 5G technologies. Latency and ultrahigh reliability are the most stringent performance indicators to support a pervasive artificial intelligence capable of offloading humans from many cognitive tasks that they
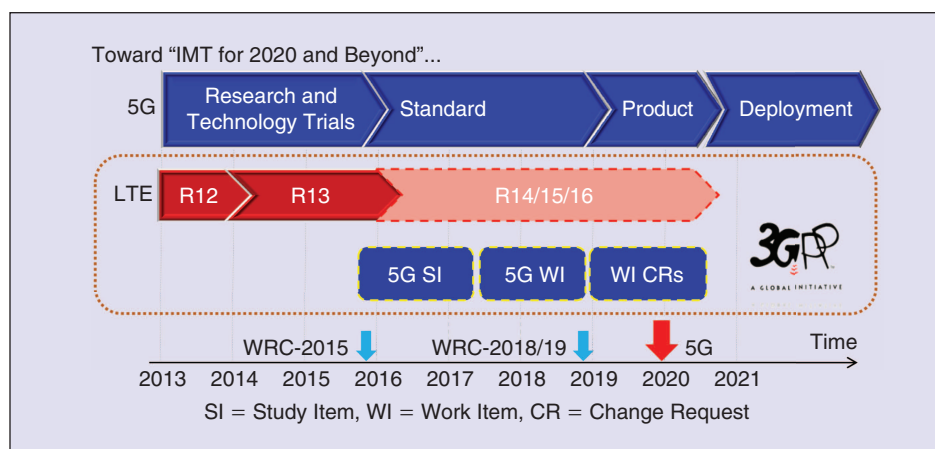
can or cannot perform and, above all, improving considerably our quality of life.

Communication services will be free of charge in most of the scenarios, they will provide a full immersive experience, and anything as a service will be the main source of revenue for SDOs, harnessing and leveraging new 5G technologies around machine intelligence, big data, and connected vehicles (mobile cognitive objects). Hence, the advanced 5G infrastructure is expected to become the nervous system of the true digital society and economy.

This hyperconnected knowledge economy will become similar to a complex system from many perspectives. Socioeconomic variables interconnect systemically in a nonlinear way. A linear analysis of these variables will no longer be valid in predicting and controlling the overall dynamics. As known, complex systems cannot easily be described by means of global rules, and their characteristics are not reducible to one level of description, even more when observing the systems from a meso- or macroscale. Complex systems exhibit properties (e.g., self-adaptation) that emerge (even quite suddenly) from local rules and nonlinear interactions between their microcomponents. New instruments of analysis will be required.

Eventually, due to this evolution, several economists as well as technologists have started to wonder whether the usual representation of relationships among myriad players in a certain area can still be modeled on the basis of value chains. There is a growing consensus that value chains modeling needs to be complemented by a broader view considering business ecosystems. The current regulations should evolve to support this new complex market and make it sustainable.

We believe that Europe can make this vision happen though crucial investments in 5G technologies and related measures and actions to strengthen the know-how in each MS, ensure EU leadership in the field of ubiquitous ultrafast-broadband infrastructure, re-enforcing privacy and data protection, and ultimately supporting the most

important plausible scenarios and valuable use cases expected at horizon 2020 and beyond.

Ultimately, we would like to state clearly that the views expressed herein are solely those of the authors and do not necessarily represent those of their affiliates.

Acknowledgments

Author Information

David Soldani (david.soldani@huawei.com) received his M.Sc. degree with the maximum score and cum laude approbatur in electronic engineering from the University of Florence, Italy, in 1994 and his D.Sc. degree in technology with distinction from Aalto University, Finland, in 2006. In 2014, he was appointed a visiting professor at the University of Surrey, United Kingdom. He is one of the top experts in multidisciplinary, transformative frontier research. He has been active in the information and communication technology field for more than 20 years, successfully working on more than 150 research and development projects for four generations of networks [second, third, fourth, and fifth generation (5G)], generating an original contribution to any quality deliverable: from strategic research and innovation to modeling, simulations, emulations, and innovative proof of concepts with stakeholders. He is currently the vice president of Huawei European Research Centre and the head of the Central Research Institute in the European Union. His areas of responsibility include, but are not limited to, future wireless, networks, Internet of Things, and multimedia technologies. He represents Huawei on the board of the 5G Infrastructure Association in Brussels and the steering board of the NetWorld2020 European Technology Platform in Europe.

Antonio Manzalini received his M.Sc. degree in electronic engineering from the Politecnico of Turin, Italy. In 1990, he joined the Centro Studi e Laboratori Telecomunicazioni, which then became the Telecom Italia Lab. He started with activities on the technologies and architectures of future optical transport networks. He has been awarded five patents on network and systems. He is the author of a book on network synchronization (for synchronous digital hierarchy), and his research and technological development results have been published in many papers. He has been active in the International Telecommunication Union (ITU) as chair of two ITU-T questions. He has been actively involved in many European projects, leading some of them in the area of future networks and autonomic computing. In 2008, he was awarded the international certification of project manager by the Project Management Institute. He is currently a senior manager at the Future Centre of Telecom Italia. His current research and technological development interests are mainly in the area of software-defined networks and virtualization, primarily for the development of new fifth-generation networks, products, and services.

References

[1] D. Soldani. (2014, Feb. 19). Report on the 5G European Summit (5G ES 2014). European Parliament. [Online]. Available: https://www.eifonline.org/events/414-next-generation-network-infrastructures.html

[2] ITU. (2015, June). IMT for 2020 and beyond. Draft new Recommendation ITU-R M.[IMT.VISION]. [Online]. Available: http://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5d/imt-2020/Pages/default.aspx

[3] (2013, Aug.). The full spectrum possibilities. A manufacturer's perspective. Brussels. [Online]. Available: http://www.huawei.eu/research-and-innovation/spectrum

[4] *5G: A Technology Vision*, White Paper, Huawei, Shenzhen, China, Feb. 2014.

[5] E. Smith, "Google keynote speech," in *Proc. Mobile World Congr.*, Barcelona, Spain, Feb. 2010, p. 1.

[6] European Commission, "Leadership in enabling industrial technologies," Horizon 2020 Work Programme, 2014–2015.

[7] A. Manzalini and R. Saracco, "Software networks at the edge: A shift of paradigm," in *Proc. IEEE SDN4 Future Networks Services*, Trento, Italy, Nov. 2013, pp. 1–6.

[8] A. Manzalini and A. Stavdas, "The network is the robot," *Commun. Strategies*, ERN Economics of Networks e-Journal, no. 96, 2014.

[9] A. Manzalini, R. Saracco, C. Buyukkoc, P. Chemouil, S. Kukliński, A. Gladisch, M. Fukui, W. Shen, E. Dekel, D. Soldani, M. Ulema, W. Cerroni, G. Schembra, V. Riccobene, A. Galis, and J. Mueller. (2014, Jan. 29). Software-defined networks for future networks and services: Main technical challenges and business implications. White Paper. [Online]. Available: http://www.researchgate.net/publication/261854995_Software-Defined_Networks_for_Future_Networks_and_Services_Main_Technical_Challenges_and_Business_Implications

[10] 3GPP, "Services and service capabilities," 3GPP TS 22.105, Rel-12, 2014.

[11] 3GPP, "Non-access-stratum protocol for evolved packet system," 3GPP TS 24.301, Rel-12, 2014.

[12] P. Popovski, V. Braun, H.-P. Mayer, P. Fertl, Z. Ren, D. Gonzales-Serrano, E. Ström, T. Svensson, H. Taoka, P. Agyapong, A. Benjebbour, G. Zimmermann, J. Meinilä, J. Ylitalo, T. Jämsä, P. Kyösti, K. Dimou, M. Fallgren, Y. Selén, B. Timus, H. Tullberg, M. Schellmann, Y. Wu, M. Schubert, D. H. Kang, J. I. Markendahl, C. Beckman, M. Uusitalo, O. Yilmaz, C. Wijting, Z. Li, P. Marsch, K. Pawlak, J. Vihriala, A. Gouraud, S. Jeux, M. Boldi, G. M. Dell'aera, B. Melis, H. Schotten, P. Spapis, A. Kaloxylos, and K. Chatzikokalakis. (2013, May 1). Scenarios, requirements and KPIs for 5G mobile and wireless system. [Online]. Available: https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D1.1_v1.pdf

[13] X. Lin, J. G. Andrews, A. Ghosh, and R. Ratasuk, "An overview on 3GPP device-to-device proximity services," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 40–48, Sept. 2013.

[14] 3GPP, "Proximity-based services," 3GPP TS 23.303, Rel-12, 2014.

[15] E. Brynjolfsson and A. Mcafee, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. New York: W. W. Norton and Company, Inc., 2014.

[16] Capgemini Consulting. (2014, Jan.). Digital transformation review: Gearing up for digital operations, N. 5. [Online]. Available: http://www.uk.capgemini.com/resource-file-access/resource/pdf/dtr5_29thnov_final.pdf

[17] K. Matheus, R. Morich, and A. Lübke, "Economic background of car-to-car communication," in *Proc. 2nd Braunschweiger Symp. Informationssysteme fuer Mobile Anwendungen*, Braunschweig, Germany, Oct. 2004.

[18] V. Yazici, U. C. Kozat, and M. O. Sunay, "A new control plane for 5G network architecture with a case study on unified handoff, mobility and routing management," *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 76–85, Nov. 2014.

[19] R. Guerzoni, I. Vaishnavi, R. Trivisonno, A. Hecker, S. Beker, Z. Despotovic, and D. Soldani, "A novel approach to virtual networks embedding for SDN management and orchestration," in *Proc. IEEE Network Operations Management Symp.*, Krakow, May 2014, pp. 1–7.

[20] X. Ge, H. Cheng, M. Guizani, and T. Han, "5G wireless backhaul networks: Challenges and research advances," *IEEE Network*, vol. 28, no. 6, pp. 6–11, Nov. 24, 2014.

[21] T. G. Orphanoudakis, C. Matrakidis, and A. Stavdas, "Next generation optical network architecture featuring distributed aggregation, network processing and information routing," in *Proc. European Conf. Networks Communications*, June 2013, pp. 1–5.

VT