3

New Waves of IoT Technologies Research – Transcending Intelligence and Senses at the Edge to Create Multi Experience Environments

Ovidiu Vermesan[1], Marcello Coppola[2], Mario Diaz Nava[2], Alessandro Capra[3], George Kornaros[4], Roy Bahr[1], Emmanuel C. Darmois[5], Martin Serrano[6], Patrick Guillemin[7], Konstantinos Loupos[8], Lazaros Karagiannidis[9] and Sean McGrath[10]

[1]SINTEF, Norway
[2]STMicroelectronics, France
[3]STMicroelectronics, Italy
[4]Hellenic Mediterranean University, Greece
[5]CommLedge, France
[6]Insight Centre for Data Analytics, NUI Galway, Ireland
[7]ETSI, France
[8]Inlecom Innovation, Greece
[9]Institute of Communication and Computer Systems, Greece
[10]University of Limerick, Ireland

Abstract

The next wave of Internet of Things (IoT) and Industrial Internet of Things (IIoT) brings new technological developments that incorporate radical advances in Artificial Intelligence (AI), edge computing processing, new sensing capabilities, more security protection and autonomous functions accelerating progress towards the ability for IoT systems to self-develop, self-maintain and self-optimise. The emergence of hyper autonomous IoT applications with enhanced sensing, distributed intelligence, edge processing and connectivity, combined with human augmentation, has the potential

to power the transformation and optimisation of industrial sectors and to change the innovation landscape. This chapter is reviewing the most recent advances in the next wave of the IoT by looking not only at the technology enabling the IoT but also at the platforms and smart data aspects that will bring intelligence, sustainability, dependability, autonomy, and will support human-centric solutions.

3.1 Next Wave of Internet of Things Technologies and Applications

IoT is defined [172] as "a dynamic global network infrastructure with self-configuring capabilities based on standard and interoperable communication protocols where physical and virtual things have identities, physical attributes, and virtual personalities using intelligent interfaces for seamless integration into the information network". In the IoT, things are expected "to become active participants in business, information and social processes where they interact and communicate among themselves and with the environment by exchanging data and information sensed about the environment, while reacting autonomously to the real/physical world events and influencing it by triggering actions and creating services with or without direct human intervention". Interfaces in the form of services facilitate interactions with these intelligent things over the Internet, query and change their state and any information associated with them, considering security and privacy issues.

The IoT is enabled by heterogeneous technologies used to sense, collect, act, process, infer, transmit, notify, manage, and store data. IoT technologies and applications are evolving from a network of objects towards intelligent social capabilities meant to address the interactions between humans and autonomous/automated systems.

The cognitive transformation of IoT applications allows the use of optimised solutions for individual applications and the integration of immersive technologies, i.e. virtual reality (VR), augmented reality (AR), and Digital Twin (DT). Such concepts transform the way individuals and robots interact with one another and with IoT platform systems.

The powerful combination of AI, IoT, Distributed Ledger Technologies (DLTs) can form the foundation of improved and – potentially – entirely new products and services. It also brings new challenges that require addressing, in a holistic manner distributed IoT architectures, decentralised security mechanisms and the evolution of IoT/IIoT devices towards intelligence-by-default and their deeper integration into platforms that aggregate the valuable data

involved in a variety of processes. Additionally, recent research outcomes in data processing and edge high-performance computing (eHPC) further support basic foundations for AI.

Artificial Intelligence of Things (AIoT) is seamlessly controlling and optimising the systems and their environment, analysing data from devices to applications. It is integrated into IoT/IIoT digital platforms that support artificial/augmented intelligence, and control and optimise the automation processes and thus support the next wave of innovation. This includes the adoption of federated learning concepts for the creation of low-latency intelligent IoT applications and new AI business models, by extracting common knowledge from participating IoT devices, while enabling reduced bandwidth use, localised personalisation of models, and granular data security and compliant privacy policy.

In the last ten years, new breakthrough technologies, such as IoT, AI and 5G have impacted our daily life at home, at work, in the public spaces, etc. They have allowed the emergence of more and more rich and complex applications developed and deployed in different market areas such as smart home and buildings, smart industry, smart mobility, smart city, smart healthcare, smart farming, and wearables. This evolution has been supported by progress done by the semiconductor industry in the design and fabrication of more complex integrated circuits, allowing the development of cost-effective smart devices, systems, and applications.

The new capabilities of 5G will lead to more complex value chains where more actors provide connectivity and bundle it with vertical-specific services. These actors, eager to benefit from new revenue streams, can include telecommunication companies looking for to diversification, equipment providers betting on small cell networks, over-the-top (OTT) players looking at unlicensed networks or purely vertical players integrating connectivity into their new services.

The IoT systems allow data gathering from the environment through different sensor types (according to the application) and data analysis to facilitate decision making by the end-user. Although the control-command concept used by IoT system is not revolutionary, IoT has allowed the gathering of a huge amount of data (big data), through wireless and Internet networks, and their storage in databases/data lakes/data spaces in the cloud for further analysis and the creation of valuable information.

At the end of 2019, there were 7.6 billion active IoT devices, a figure which is expected to grow to 24.1 billion in 2030, a compound annual growth rate (CAGR) of 11%, according to the research published by Transforma Insights [4].

Short range connectivity solutions, (e.g. Wi-Fi, Bluetooth and ZigBee), will dominate connections, accounting for 72% in 2030, largely unchanged compared to the 74% it accounts for today. Public networks, which are dominated by cellular networks, will grow from 1.2 billion connections to 4.7 billion in 2030, growing market share from 16% to 20%. Private networks account for the balance of connections, 10% in 2019 and 8% in 2030. Services, including connectivity, will account for 66% of spend, with the remainder accounted for by hardware, in the form of dedicated IoT devices, modules and gateways [4].

In IoT deployments, the collection and processing of a large amount of information at the edge of the network where the IoT end-devices are deployed and where information is captured and collected has been a major evolution. New IoT systems use smart solutions with embedded intelligence performing real-time analysis of information and connectivity at the edge. These new IoT systems are departing from centralised cloud-computing solutions towards distributed intelligent edge computing systems. Traditional centralised cloud computing solutions are generally fit for non-real-time applications that require high data rates, huge amounts of storage and processing power, that do not require low latency, and can be used for heavy data analytics and AI processing jobs. On the other hand, distributed edge solutions introduce computations at the edge of the network where information is generated supporting real-time services with very low latency (in the order of milliseconds) that can be used for simple to medium ultra-fast analytics. The collection, storage and processing of data at the edge of the network in a distributed way also contributes to the increased privacy of the user data since raw personal information no longer needs to be stored in backbone centralised servers and each user can retain the full control of the data. Even more, for some applications, data can be processed online without need of storage capabilities. However, the convergence between cloud computing, edge, and IoT requires smart and efficient management of resources, services, and data, whose elements can move across different heterogeneous infrastructures.

The densification of the mobile and associated services network strongly challenges the connection with the core network. Next-generation IoT networks need greater cloud systems flexibility and implement cloud utilisation mechanisms to maximise efficiency in terms of latency, security, energy efficiency and accessibility. Cloud technologies should combine software-defined networks and network function virtualisation to enable network flexibility in order to integrate new applications and adequately configure network resources (e.g. by sharing computing resources, splitting data traffic,

enforcing security rules, implementing QoS parameters, ensuring mobility). Global connection growth is mainly driven by IoRT devices on both the consumer side (e.g. smart homes) and the enterprise/business-to-business (B2B) side (e.g. connected machinery). These devices require an extension of the spectrum in the 10–100 GHz range and unlicensed bands and technologies, such as WiGig or 802.11ax, which are mature enough for massive deployment and can be used for cell backhaul, point-to-point or point-to-multipoint communication.

Edge processing requires more specialised hardware knowledge (e.g. security and privacy protection, data analysis through Machine Learning (ML) techniques and more precisely implementation of Deep Learning (DL) algorithms) to ensure that the IoT devices operate efficiently and capture data in ways supporting more complex and sophisticated processing at the edge (such as distributed ledgers, increased trust, and identity management, data analysis, etc.). These advances in IoT technologies allow the integration of IoT capability directly into modules or hardware so that, once the IoT device is deployed, it can automatically connect to any network available and start provisioning data – to an on-premise solution, an edge computing intelligent infrastructure or the hyperscale cloud services – that can be used by a rich set of applications.

In this context, the wireless connectivity for IoT deployments will continue to diversify with different communication protocols being used according to the needs and performance required by different IoT applications and services. At the same time, these communication technologies have opened the door to potential malicious cyber-attacks and threats, and the risk to stole personally identifiable information (PII). End-to-end security and privacy mechanisms have been developed over time to avoid these issues. However, standard solutions are needed. Furthermore, end-devices processing capabilities have been fostered by the maturation of AI, and the flexibility offered by ML and DL techniques regarding the place/level where the data analysis must be performed. Performing data analysis closer to the data sources has important benefits such as reduced overall system power consumption, reduced communication bandwidth and throughput, increased data security and privacy (personal data protection), and reduced processing latency. However, state of the art techniques and requirements impose for more data processing and “smartness” at the edge, including at sensor level, end-device (sensor nodes) level, server level with gateway capabilities, and other network components.

From the IoT system architecture perspective to obtain the mentioned benefits, this implies to move the data analysis (for many applications) from Cloud to Fog and Edge architectures. It also creates the need for greater interoperability between the different systems at cloud, edge and fog levels and a key role for the associated standards.

These changes are made possible especially if the devices involved in the edge domain (e.g. edge servers, end-devices, and sensors) are energy efficient and can support AI techniques at low power consumption to ensure high autonomy/longer battery life, system availability and reliability. This will become possible by developing a new generation of more performant Integrated Circuits (ICs) with the best trade-off between increased processing power vs ultra-low power consumption.

Table 3.1 gives an overview of the IoT evolution phases and indicates, the focus areas of edge devices evolution. Finally, it should be also considered the impact that 5G will have in the more intelligent edge devices for future applications requiring high throughputs such as vehicle to everything (V2X), augmented reality (AR), virtual reality (VR), Industrial 4.0, autonomous systems and e-health.

Table 3.1 IoT evolution and associated technologies

Main Features	**Phase 1 (2016–2018)**	**Phase 2 (2019–2020)**	**Phase 3 (2021–2025)**
Reference Architecture Domains: Cloud, Edge	Cloud Computing	Cloud and Fog Computing	Cloud, Fog, Edge Computing /Private Networks
Data analysis	On the Cloud	On the Cloud, starting in the Edge servers, Hybrid	Cloud, Edge (Server, End-Device and Sensor), Hybrid
Artificial Intelligence	No	Cloud, Edge Servers	Cloud, Edge (Server, End-Device and Sensor)
End-to-End security	No	Yes	Yes
End-to-End privacy	Not considered	Not yet supported every where	Yes
Connectivity between domains	Wired, 3G/4G	Wired, 3G/4G	Wired, 3G/4G, 5G

(Continued)

Table 3.1 Continued

Main Features	Phase 1 (2016–2018)	Phase 2 (2019–2020)	Phase 3 (2021–2025)
Connectivity edge domain	Wi-Fi, BT, Z-Wave, ZigBee, LoRa, Sigfox,	Wi-Fi, BT, Z-Wave, ZigBee, LoRa, Sigfox,NB-IoT, LTE-M	Wi-Fi, BT, Z-Wave, ZigBee, LoRa, Sigfox.NB-IoT, LTE-M, 5G
Edge devices	Gateway with routing capabilities	Edge Servers with Gateway and routing capabilities	Distributed Edge Servers with Gateway and routing capabilities
End devices	Smart Sensors	Smart Sensors Cyber-Physical Systems (Drones, Robots)	Smart Sensors, Smart Cameras, Cyber-Physical Systems (Drones, Robots, Autonomous cars)
Architecture	Sensor + Computing processing + Connectivity	Sensor + Computing processing + Connectivity + Security	Sensor + Computing processing + Connectivity + Security + Privacy + AI
Security	Not supported	Yes	Yes
Privacy	Not supported	Not supported	Yes
AI capabilities for data analysis and decision making	No supported	No supported	Yes
Data throughput	Low ($<$ 200 Kbytes)	Low ($<$ 500 Kbytes)	$>$ 500 Kbytes
Computing processing	Low	Medium	High
Sensor	Simple	Multiple sensors	Multiple-sensors, smart cameras
Actuator	Not supported	Not supported	Yes
Autonomy	Low	Medium	High Autonomy
Reliability	No	Improved	Mandatory
IoT Digital Twin (DT)	Status IoT DTs, simulation DTs,	Operational IoT DTs with events and simulations (domain specific)	Autonomous IoT DTs, collaborative twins, at the edge, cross-domain

The next waves of IoT developments are presented in Figure 3.2. IoT/IIoT research and innovation is building on advancements in semiconductors, photonics, sensing technologies AI techniques and embedded systems for the "things". The creation of next-generation information and communication platforms will be strongly driven by wireless, cellular (e.g. 5G and beyond) technologies, which will generate massive amounts of information as input for hyper computing processing capabilities and enabling new applications in services such as robotics, autonomous vehicles, AI and intelligent systems of systems.

3.2 Energy Efficient and Green IoT 3D Architectural Approach

Climate change and environmental degradation are an existential threat to Europe and, more globally, to the world. The European Green Deal roadmap for economic sustainability entails a new growth strategy to transform Europe "into a modern, resource-efficient and competitive economy where there are no net emissions of greenhouse gases by 2050, economic growth is decoupled from resource use" [15], and no persons and no place are left behind. This will be done by "turning climate and environmental challenges into opportunities across all policy areas and making the transition just and inclusive for all" [15]. The Green Deal is an integral part of the EC's strategy to implement the UN's 2030 Agenda and its sustainable development goals [16]. To implement this strategy, a circular economy action plan [17] has detailed measures to make sure that sustainable products are the norm in the EU, which means that products on the EU market are designed to last longer, are easier to reuse, repair, and recycle, and incorporate recycled material as much as possible, put a major focus on electronics, ICT and batteries, and drastically reduce waste by transforming it into high-quality secondary resources [17].

Digital technologies are a significant enabler for attaining the sustainability goals of the Green Deal in many different sectors. The EC is launching initiatives and taking measures to ensure that digital technologies such as AI, 5G, cloud and edge computing and the IoT can accelerate and maximise the impact of policies that deal with climate change and protect the environment. In digitalisation, IoT in particular offers new opportunities for remote monitoring of air and water pollution, or for monitoring and optimising how energy and natural resources are used [15].

ICT is consuming increasing amounts of energy and this is also true for IoT/IIoT systems and applications. Saving energy and fuel-costs by using energy-efficient demand-side technologies can lead to increased consumption of those services, an outcome known as the "rebound effect". For ICT, this means that if the demand for the services increases more than 11% due to rebound, all of the environmental or natural resource impacts considered will rise even if overall energy consumption declines [13].

Disruptive technological change can enable sustainable development with global benefits for closing the emissions gap but can also exacerbate unsustainable patterns of resource use. This is clearly evidenced by the promises and risks of the digital revolution and the ongoing advances in ICT, ML and AI, connectivity, IoT, additive manufacturing (e.g., 3D printing), virtual and augmented reality, blockchain, robotics and synthetic biology [12].

The next waves of IoT technologies and system architectures are designed to improve energy efficiency and circular economy performance of the IoT/IIoT systems and use the IoT applications to support all industrial sectors to increase the energy efficiency, reduce their CO_2 footprint and provide new innovative value chains, value networks and business models to achieve economic sustainability.

New innovative models are needed to accelerate circularity and the dematerialisation of the economy to make Europe less dependent on primary materials [17]. These models, based on better involvement of customers, information sharing, mass customisation, sharing and collaborative economy, will be powered by digital technologies, such as IoT, big data, blockchain and AI.

IoT must consider the increase of energy efficiency with an optimal use of resources and infrastructure supported by AI techniques with the use of real-time analytics related to energy and resources, the analysis of event and information streams across the architectural layers, the optimisation of the location of processing, and the transfer of intelligence where the application needs it (e.g. edge, IoT gateway, device, etc.). Energy-efficient, and green IoT requires a holistic end-to-end strategy across the information value chain through the IoT architectural layers. This requires the design of energy-efficient and green IoT components at each layer level and at the IoT applications level by combining AI technologies and optimising IoT capabilities across the architectural layers to unify the complete IoT and analytics life cycle, streaming the information, filtering, scoring, exchanging storing what's relevant, analysing and using the results to continuously improve and optimise the IoT system.

As it is defined in [21] the green IoT must address the study and practice of designing, using, manufacturing, and disposing of IoT devices, systems, subsystems, communication network systems, HW/SW platforms efficiently and effectively with minimal or no impact on the environment. In this context, a holistic approach is needed to follow the IoT green design (e.g. advanced and adequate semiconductor technologies, efficient design, energy efficient SW/HW platforms) for providing environmentally sound components at all IoT architectural layers and functions, energy efficient and low CO_2 footprint at IoT infrastructure and technical solutions levels (e.g. edge, cloud, data centre, AI-based learning/training, etc.), green manufacturing (e.g. manufacture IoT electronic components, HW/SW platforms, and IoT systems with minimal or no impact on the environment.), green use (e.g. reduce/minimise the energy consumption of IoT systems and support the monitoring and optimisation of the energy consumption for other systems in various industrial sector applications and use them in an environmentally sound manner) and green disposal (refurbish and reuse the IoT systems and properly recycle these systems and support through IoT applications by monitoring the implementation of solutions for recycling, and circular economy in different industrial sectors).

Green and energy efficient IoT strategies involve the use of IoT technologies and the design of applications in such a way that the IoT systems are optimised across all the IoT architectural layers. These strategies must include the embedded carbon generated by producing and moving the materials for the IoT electronics, producing the devices, installing the IoT systems, etc. In addition, IoT technologies and applications enable and leverage applications in different industrial sectors that support cutting CO_2 emissions increasing energy efficiency, monitoring supply chains for providing circular economy solutions, reducing resources consumption, preserving natural resources, minimising the technology impact on the environment and human health and reducing the costs.

In order to reach the previous goals, the use of the IoT 3D reference architecture provides the tool to optimise the IoT systems developments for energy efficiency and green design by identifying the specific (energy and green design) techniques that need to be applied to the design and implementation of various HW/SW components across the different IoT architectural layers as part of the systems implementation and IoT platforms used.

The architecture implementation depends on the application domains and their requirements, specific QoS factors, systems characteristics, and the cross-cutting functions involved. The architecture and application designers

can identify the major energy intensive components, nodes, protocols, middleware elements, SW algorithms, HW blocks/sub-systems in each layer and optimise them at each layer, across several layers and at the application level.

For this, CAD tools and design flows are required to optimise each individual component and the overall system. Furthermore, the analysis of the energy efficiency, the green methods and techniques to be applied, can be done by defining, at each layer, the main functions and components and identifying adequate optimisation parameters to be used for energy efficiency and green IoT design.

The approach utilises the 3D IoT layered architecture [62] based on eight IoT domain layers as illustrated in Figure 3.1. The architecture is composed of two other views that include the system properties and the cross-cutting functions.

The green attributes and energy efficiency are correlated with various tasks (e.g. sensing/actuating, communicating, processing, analysing, storing, transferring, learning, etc.) performed in the different architecture layers by various HW/SW components and algorithms integrated into IoT platforms running at the edge or on cloud infrastructure.

The implementation of green and energy efficient techniques and methods (e.g. optimisation, trade-off analyses among cross-cutting functions/system properties vs. energy, green IoT metrics, performance, measurement, test-beds, energy harvesting, wireless power transfer, etc.) depends on the

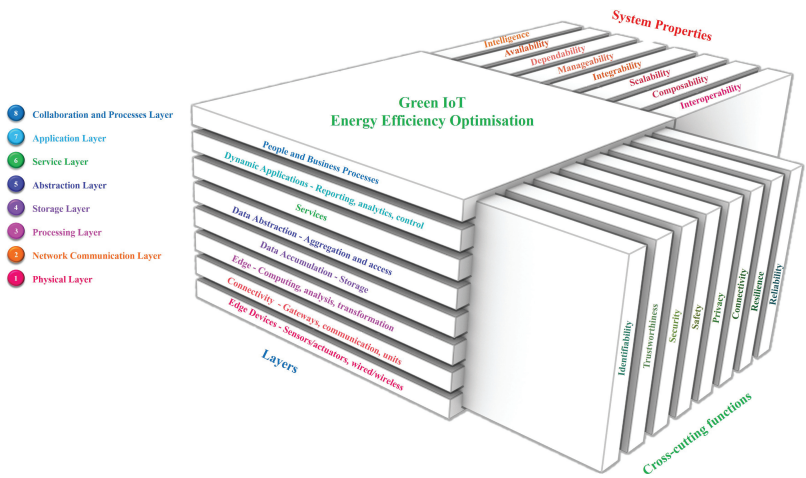

Figure 3.1 3D IoT layered architecture. Adapted from [62].

functions performed by different HW/SW components integrated in the IoT layers, which include energy management, wake-up scheduling mechanisms, selective sensing, HW-SW partitioning, energy efficient algorithms, communication techniques and distribution of tasks, minimisation of data path length, data buffer delivery, wireless communication, and processing trade-offs [18–20]. To optimise energy efficiency, CAD tools and design flows are required at each layer.

3.2.1 Physical Layer

The physical layer includes the sensors/actuators, processing (SW/HW), and connectivity for all IoT end-devices. The design of the HW/SW components, functional blocks, sub-systems define the relevant functions that should be implemented through a set of ICs on a board, a system-on-chip (SoC), a system-in-package (SiP) or an embedded system intelligent component. The intelligence and complexity of a physical object can vary from simple RFID tags with sensing capabilities to Cyber physical systems (CPS) such as intelligent robots, drones and autonomous vehicles. The requirements for these physical layer components are correlated with the elements from the other two dimensions of the IoT 3D architecture such as system properties (intelligence, availability, dependability, manageability, integrability, scalability, composability, interoperability) and cross-cutting functions (identifiability, trustworthiness, security, safety, privacy, connectivity, resilience, reliability). The IoT application requirements will include the optimisation of energy efficiency and green properties of the HW/SW components in the physical layer aligned with the degree of system properties embedded (e.g. intelligence, scalability, interoperability, etc.) and the stringent implementation of the cross-cutting functions (e.g. end-to-end security, safety, resilience, etc.). The energy efficiency and green properties optimisation, with the right tools at this level in the early design phase of the IoT systems developments offers flexibility in designing energy efficient and green IoT devices aligned with the requirements of the IoT use cases and applications. Furthermore, technology considerations are also key aspects to ensure high performances at low power (autonomy is critical for many IoT applications) and provide high integrated solutions at cost effective. They are important factors to accelerate the introduction of IoT devices on the market.

The tasks performed by the devices in this layer focus on collecting the information from the physical environment, self-organised sensing, processing, load balancing and preparing data to be exchanged with other layers.

The components of this layer include sensors, microcontrollers (MCUs), wireless modules, actuators, energy management, energy sources (batteries, solar panels, etc.), SW embedded components, operating systems, etc. which all together constitute the edge processing nodes or well known as end-points or end-devices. For intelligent edge devices, performing data analysis, more intelligence is added to the IoT edge devices and AI-based HW/SW modules or HW accelerations associated to MCUs are included in the design. For the intelligent devices, the overall CO_2 footprint including the green characteristic and energy efficiency of the AI inference at the edge must be considered over the lifetime of the device compare with the same function and the learning/training performed in the cloud or data centres. For the external infrastructure elements, the type of energy (e.g. renewables, fossil, etc.) used for powering the facilities must be included in the design optimisation. Moreover, the energy management techniques, the selection of materials, energy sources (e.g. batteries, solar panels, socket, etc.), energy harvesting, SW optimisation for data sensing, monitoring, filtering, prediction and compression with processing, sleep/wake-up techniques, energy efficient task scheduling algorithms, selection of quality of information (QoI), allocation of workload distribution at edge, wireless communication optimisation (send/receive), power down mechanisms, dynamic wireless network behaviour (e.g. IoT devices-move-in and IoT devices-move-out), cooperation/information exchange between the edge nodes while processing are elements that need to be consider to optimise the green and energy profile of the IoT devices. It should also be considered that the benefits brought by performing data analysis in the physical layer such as reduce bandwidth use have a direct impact in the energy consumption saving in the overall communication path, reduce the size of the data memory storage considering that the data were already processed by the physical device, finally reduce processing power and energy saving in the cloud/data centres. The energy and resources saving could be very important in an edge computing approach versus a cloud computing one.

3.2.2 Network Communication Layer

The network and communication layer plays a central role in all IoT applications and deployments that support the exchange of information between different layers and create value due to the interaction in real-time between IoT devices. The transmission of information between IoT devices, the network, across networks, and between the networks and high-level information

processing infrastructure (e.g. cloud, data centres, etc.) is one of the main features of IoT applications. The evolution from centralised IoT architectures to decentralised and distributed increase the importance of the continuous "6As" connectivity (anything, anyone, anywhere, anytime, any path/network, any service).

This layer provides communication between the physical layer (IoT devices) and the components of other layers including the interactions between different types of networks (e.g. wireless, wireless sensor networks, cellular, optical, satellite, etc.) that connect the IoT devices to gateways, edge, cloud, and data centres infrastructure. The selection of connectivity infrastructure constituted of hubs switches, gateways, routers, radio access network architecture, edge processing for software-defined and cognitive-based radios and network functions virtualisation, SW algorithms, remote transmitting, cloud servers and data centres architectures, define the CO_2 footprint, green, and energy profile of the network and communication layer IoT components. This layer plays the role of a dynamic channel for transferring/exchanging information from the physical layer to processing, storage layers and beyond base on more than 50 wireless and cellular communication protocols (e.g. Wireless Personal Area Network (WPAN), Wireless Local Area Network (WLAN), Wireless Neighbourhood Area Network (WNAN), Wireless Wide Area Networks (WMAN). Cellular WWAN, Low Power Wide Area Networks (LPWAN), etc.).

The optimisation of the network and communication layer components must consider the functionalities required for transmission of information, neighbourhood communication, retransmission and avoidance of packet collision, handshaking mechanisms, etc. For the new cellular technologies such as 5G and beyond the network densification via small cells need new models for optimising the green effect and energy efficiency for IoT applications as the 5G network infrastructure will increase the energy consumption. The 5G networks will have higher energy efficiency (less energy per transmitted bit) but will consume more power than 4G networks. The energy consumption is a result of the growth in mobile data consumption, deployment of additional, network elements (e.g. small cells, MIMO antennas) and the use of mobile edge computing, cloud computing and data centres processing.

A trade-off must be considered among multi-hop and cooperative multi-hop routing in wireless sensor networks communication to gateways and base stations, path optimisation to reduce energy consumption, network traffic control, dynamic downloading of packets using access points, scheduling of the communication tasks to reduce energy consumption in network

communication, etc. The optimisation techniques to be applied at the IoT network and communication layer support are achieving a much higher energy efficiency of the IoT network with limited bandwidth provisioning and low transmit power, by utilizing advanced capabilities of IoT (e.g. in-network storage and caching, offload the IoT data to release the traffic scale in the cellular networks, provide low-latency IoT services in an energy efficient manner, etc.), designing lightweight context-aware security schemes (e.g. encryption, authorisation, etc.) to reduce the energy consumption of secure IoT networks.

Last but not least, the IoT system architecture choice between cloud vs fog vs edge computing will be fundamental in the energy savings and determinant in the complexity of the Network Communication Layer.

3.2.3 Processing Layer

This layer addresses edge computing, information element analysis and transformation, analytics, mining, machine learning, in a pervasive manner considering that autonomic services must be provided through ubiquitous machines in both "autonomic" and "smart" way. The processing layer provides the ability to process and act upon events created by the edge devices and store the data into a database in the storage layer. The processing layer can be closely interlinked with data analytics platform(s) based on edge-/cloud-scalable platform that supports information processing technologies. The green IoT and energy efficiency optimisation depend on the type of information procession, location, real-time requirements, AI components, platforms, and system architecture (e.g. centralised, decentralised, distributed). Complex event processing can be supported for edge, cloud, data centre solutions to initiate near real-time activities and actions based on data from the edge devices and from the rest of the system. The requirements for the processing layer are connected to the need for highly scalable, column-based data storage for storing events, map-reduce for long-running batch-oriented processing of data and complex event processing for fast in-memory processing and near real-time reaction and autonomic actions based on the data and activity of devices and the interconnected systems.

Edge computing performs data analysis at sensors, end-devices, the gateways, micro servers, edge/embedded high-performance computing levels and the enterprise applications leverage edge devices data in end-to-end value streams involving edge devices and people within digitised processes.

Developments of a new range of IoT applications is enhancing the role of this layer that is being used to analyse, process and in many cases store the information at the edge level. This layer provides support to the lower layers, allocates resources for efficient storage in virtual and physical machines and converts the data into the required form. The components of this layer include HW/SW processing elements/units (e.g., operating systems, embedded software, MCUs, CPUs, GPUs, FPGAs, other AI-based accelerators, etc.), semantic-based and service-based middleware, databases virtual machines, resource allocators, information convertors and translators.

Processing the information at the edge where data is being generated, and providing insights from the information kept at the network and processing level has the benefits of reducing both latency and the demand on network bandwidth, preventing downtime, reducing CO_2 emissions, and improving the overall system efficiency. The integration of ML and AI methods support the optimisation of edge computing-based green and energy efficient functions providing solutions for moving optimally the processing from cloud to the edge and decarbonising the whole value chain of IoT information.

3.2.4 Storage Layer

The storage layer addresses the efficient storage and organisation of data and the continuous update with new information, as it is made available through the capturing and processing of IoT channels. The functions implemented include archiving the raw and processed data for offline long-term storage and the storage of information that is not needed for the IoT system's real-time operations. Centralised storage considers the deployment of storage structures that adapt to the various data types and the frequency of data capture. Decentralised and distributed storage is implemented with the functions in the processing layer. Relational database management systems that involves the organisation of data into table schemas with predefined interrelationships and metadata for efficient retrieval for later use and processing are used for different IoT systems.

For many autonomous IoT applications, the storage is decentralised, and data is kept at the edge or at the device that generates it and is not sent up across the other top layers. This model is used in conjunctions with the mobile edge computing and fog computing implementations and offer energy, computing, and connectivity resource optimisation advantages. Scalable storage platforms can be used to process very large data sets (including the ones for AI operations) across many computing nodes that operate in parallel,

which provides a cost-effective storage solution for large data volumes with no format requirements and used by several IoT solutions and connected to different existing IoT platforms.

3.2.5 Abstraction Layer

The abstraction layer provides the interfaces, the event and action management through simple rules engine to allow mapping of low-level sensor events to high level events and actions, while assuring the basic analytics for data normalisation, reformatting, cleansing and simple statistics. It provides the visualisation of processed data in the form of intelligent tasks. The layer incorporates the components and SW/HW elements to implement these functions and allow the IoT systems to scale using multiple storage systems to accommodate IoT device information and data from traditional Enterprise Resource Planning (ERP), Customer Relationship Management (CRM), Supplier Relationship Management (SRM), Supply Chain Management (SCM), Product Lifecycle Management (PLM) and other systems that interact with the IoT platforms and applications. The data abstraction functions are rendering data and its storage in ways that enable developing simpler, performance-enhanced applications. Green and energy efficient HW/SW components at data abstraction layer must ensure the data processing optimisation and reconcile multiple data formats from different sources, assuring consistent semantics of data across sources, confirming that data is complete to the higher-level application, consolidating data, providing access to multiple data stores through data virtualisation, normalising or denormalising and indexing data to provide fast application access, protecting data with appropriate authentication and authorisation.

3.2.6 Service Layer

The service layer integrates the middleware on top of networks and IoT device streams and provides data management and data analytics that are key functions for IoT systems where large amounts of sensor generated data and events must be logged, stored and processed to generate new insights or events, AI-based processing on which business decisions can be made.

The service layer functions are provided at the edge or cloud level. Several cloud service providers are extending the offerings to implement IoT solutions by using the existing “infrastructure as a service (IaaS)” ecosystems to provide new IoT services. In many cases these systems are not optimised

for energy efficiency and green usage and an evaluation must be performed to analyse the CO_2 footprint and appropriateness to be used for specific IoT applications.

Various IoT platforms are providing full stack solution for ingesting data from IoT devices and linking them to edge-/cloud-based storage and processing services. The platform as a service (PaaS) offering integrate the functions of the IoT service layer and are not always optimised for low energy consumption and CO_2 footprint.

3.2.7 Application Layer

The application layer is offering the software platforms that are suited to deliver the key components for implementing various IoT applications that are connecting users, business partners, devices, machines, and enterprise systems with each other to provide information interpretation to different applications. The SW/HW components at this layer interact with the service layer, while the software applications are based on vertical markets, the nature of device data, and business needs. At this layer, many applications are addressed such as mission-critical business applications, specialised industry solutions, mobile applications, analytic applications that interpret data for business decisions, etc. The optimisation for energy efficiency and green IoT requires the use of federation and orchestrations techniques that creates dynamic and distributed energy control frameworks for IoT applications that need large capacity, higher delivery efficiency, reduced energy consumption and low costs. The implementation of energy efficient search engines, the cooling systems and use of energy harvesting techniques and of renewables must be considered when the HW/SW components of the IoT application layer are evaluated. New IoT applications including AR/VR, digital twins, virtual simulations, real-time searching engines and discovery services will bring new challenges to optimising the energy efficiency.

3.2.8 Collaboration and Processes Layer

The collaboration and processes layer includes the enterprise systems and large data platforms and the exchange of information among these platforms based on high-level collaborative processes. This layer addresses the processes that involves assets, people and organisations that use IoT applications and associated information for their specific needs or for a range of different purposes, to provide the right information, at the right time, to perform the

right tasks. The energy efficiency and green IoT functions have to consider end-to-end optimisation for each layer and as the data is moved across the layers to secure the optimal communication, exchange of information and energy consumption per functions at each level and between layers.

The overall optimisation required to address the energy efficiency and green IoT capabilities must be addressed at the architecture level by considering the aggregation, over the technology stack, of the functions that are required to fulfil a given IoT task. This includes estimating the energy used for learning/training of different algorithms implemented in various IoT architectural layers by employing large data sets from different databases.

3.3 IoT Strategic Research and Innovation at the Horizon

The IERC brings together EU-funded projects with the aim of defining a common vision for IoT technology and addressing European research challenges. The rationale is to leverage the large potential for IoT-based capabilities and promote the use of the results of existing projects to encourage the convergence of ongoing work; ultimately, the endpoints are to tackle the most important deployment issues, transfer research and knowledge to products and services, and apply these to real IoT applications.

The objectives of IERC are to provide information on research and innovation trends, to present state-of-the-art IoT technologies and societal trends, to apply the developments to IoT-funded projects and to market applications and EU policies. The main goal is to test and develop innovative and interoperable IoT solutions in areas of industrial and public interest. The IERC objectives are addressed as an IoT continuum of research, innovation, development, deployment, and adoption.

The IERC launches every year the Strategic Research and Innovation Agenda (SRIA), which is the outcome of discussion involving the projects and stakeholders involved in IERC activities.

Enabled by the activities of the IERC, IoT is bridging physical, digital, virtual, and human spheres through networks, connected processes, and data, and turning them into knowledge and action, so that everything is connected in a large, distributed network. New technological trends bring intelligence and cognition to IoT technologies, protocols, standards, architectures, data acquisition, processing approaches, and analysis, all with a societal, industrial, business, and/or human purpose in mind. The IoT technological

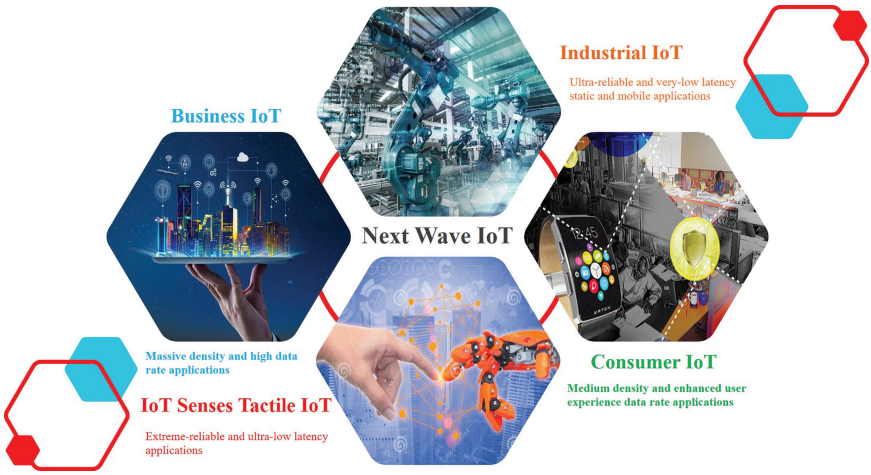

Figure 3.2 Next wave IoT technology developments.

trends are presented in the context of integration; hyperconnectivity; digital transformation; and actionable data, information, and knowledge.

The IERC SRIA addresses these IoT technologies and covers in a logical manner the vision, technological trends, applications, technological enablers, research agendas, timelines, and priorities, and finally summarises in two tables future technological developments and research needs.

The IoT technologies and applications will bring fundamental changes in individuals' and society's views of how technology and business work in the world. The IERC supports the expansion of common European IoT, Industrial IoT (IIoT) and data/knowledge ecosystems, strengthen the impact of research and innovation in developing, supporting and implementing the policies address global challenges, including climate change and the Sustainable Development Goals. The next wave of human-centred IoT/IIoT green technologies, applications and the development of competitive European ecosystems are aligned with the priorities such as the European Green Deal, an economy that works for people, and a stronger Europe fit for the Digital Age. This has an important impact on the research activities that need to be accelerated without compromising the thoroughness, rigorous testing and time required for commercialisation, business impact and economic growth.

Knowledge-driven technologies may bring future risks, that can potentially include erosion of individual privacy, misinformation, misuse of data, increase of cybersecurity threats, and increase the digital divide.

Developing research and innovation agendas that blends the priorities to foster technology excellence while encouraging the respect to rights and values of citizens become of paramount relevance. This balancing act involves costs and legitimate private, national, public interests and the rights and interests of producing and using data.

Research and development are tightly coupled. Thus, the IoT research topics should support the further development of IoT ecosystems, partnerships, stakeholders networking and implementation of secured and trusted IoT solutions based on reference implementations for smart devices into self-adaptive, robust, safe, interconnected smart network and service platforms, proof-of-concept, demonstrations driven by realistic use cases in different sectors. A hyperconnected society is converging with a consumer-industrial-business Internet that is based on hyperconnected IoT environments. The latter requires new IoT systems architectures that are integrated with network architecture (a knowledge-centric network for IoT), system design and horizontal interoperable platforms that manage more intelligent things that are digital, automated and connected, functioning in real-time, having remote access and being controlled based on Internet-enabled tools.

The next-generation IoT/IIoT developments, including human-centred approaches, are interlinked with the evolution of enabling technologies (AI, connectivity, security, etc.) that require strengthening trustworthiness with electronic identities, service and data/knowledge portability across applications and IoT platforms. This ensures an evolution towards distributed IoT architectures with better efficiency, scalability, end-to-end security, privacy, and resilience. The virtualisation of functions and rule-based policies will allow for free, fair flow of data and sharing of data and knowledge, while protecting the confidentiality, integrity, and privacy of data. Vertical industry stakeholders will become more and more integrated into the connectivity-network value chain. Moreover, unified, heterogeneous, and distributed applications, combining information and operation technologies (IT and OT), will expose the network to more diverse and specific demands.

Intelligent/cognitive connectivity networks provide multiple functionalities, including physical connectivity that supports the transfer of information and adaptive features that adapt to user needs (context and content). These networks can efficiently exploit network-generated data and functionality in real-time and can be dynamically instantiated close to where data are generated and needed. The dynamically instantiated functions are based on intelligent algorithms that enable the network to adapt and evolve to meet changing requirements and scenarios and to provide context- and

content-suitable services to users. The intelligence embedded in the network allows the functions of IoT platforms to be embedded within the network infrastructure and data, and the knowledge generated by the intelligent connectivity network and by the users/things can be used by the network itself. This knowledge can be taken advantage of in applications outside of the network.

The connectivity networks for next-generation IoT/IIoT are transforming into intelligent platform infrastructures that will provide multiple functionalities and will be ubiquitous, pervasive, and more integrated, further embedding telephone/cellular, Internet/data and knowledge networks.

Advanced technologies are required for the Next Generation Internet (NGI) to provide the energy-efficient, intelligent, scalable, high-capacity and high-connectivity performance required for the intelligent and dynamically adaptable infrastructure to provide digital services – experiences that can be developed and deployed by humans and things. In this context, the connectivity networks provide energy efficiency and high performance as well as the edge-network intelligence infrastructure using AI, ML, NNs and other techniques for decentralised and automated network management, data analytics and shared contexts and knowledge.

New solutions are needed for designing products to support multiple IoT standards or ecosystems and further research is required to investigate new standards and related APIs.

Summarising, although huge efforts have been made within the IERC community for the design and development of IoT technologies, the continuously changing IoT landscape and the introduction of new requirements and technologies create new challenges or raise the need to revisit existing well-acknowledged solutions. Thus, below is a list of the main open research challenges for the future of IoT:

- New IoT architectures that consider the requirements of distributed intelligence at the edge, cognition, AI, context awareness, tactile applications, heterogeneous devices, end-to-end security, privacy, proven trustworthiness, and reliability.
- Development of digital twin concepts, technologies and standards that enable cross-domain collaboration and deployment.
- Augmented reality and virtual reality IoT applications.
- Autonomics in IoT towards the Internet of Autonomous Things.
- Inclusion of robotics in the IoT towards the Internet of Robotic Things.
- AI and ML mechanisms for automating IoT processes.

- Distributed IoT systems using securely interconnected and synchronised mobile edge IoT clouds.
- IoT systems architectures integrated with network architecture forming a knowledge-centric network for IoT.
- Intelligence and context awareness at the IoT edge, using advanced distributed predictive analytics.
- IoT applications that anticipate human and machine behaviours for social support including contextual behaviour and understanding.
- Stronger distributed and end-to-end holistic security solutions for IoT, preventing the exploitation of IoT devices for launching cyber-attacks, i.e., remotely controlling IoT devices for launching Distributed Denial of Service (DDoS) attacks.
- Pre-normative and standardisation activities to address IoT end-to-end security.
- Stronger privacy solutions, considering the requirements of the new General Data Protection Regulation (GDPR) [184] for protecting the users' personal data from unauthorised access, employing protective measures (such as Privacy Enhancing Technologies – PETs) as closer to the data generation source to perform anonymisation and thus guarantee the personal data protection.
- Cross-layer optimisation of networking, analytics, security, communication, and intelligence.
- IoT-specific heterogeneous networking technologies that consider the diverse requirements of IoT applications, mobile IoT devices, delay tolerant networks, energy consumption, bidirectional communication interfaces that dynamically change characteristics to adapt to application needs, dynamic spectrum access for wireless devices, and multi-radio IoT devices.
- Adaptation of software-defined radio and software-defined networking technologies in the IoT.
- Trusted software validation approaches supporting advanced firmware analysis and validation approaches before and during the firmware update process involving industrial IoT devices.
- Application-specific mechanisms installed on IoT devices to support quicker algorithmic implementations (encryption, etc.) and hardware solutions supporting distributed trust (e.g. Physical Unclonable Functions, etc.) to cope with limited processing power on IoT devices.
- Tactile IoT applications and supportive technologies.

3.3.1 Digitisation

Digitisation is entering many fields, and the influence of digital approaches and techniques is becoming more apparent as time passes. Buildings and cities are becoming smarter, vehicles are becoming self-driving, design processes are becoming highly efficient and objects and spaces can be visualised before being materialised. Devices with embedded sensors featuring complex logic are scattered everywhere to measure light, noise, sound, humidity, and temperature, and they are empowered to communicate with each other, forming an IoT ecosystem.

Digital transformation is pushing all market sectors to level up their digital capabilities to better serve customers and meet their expectations, requiring to quickly transform insights into the best user experience. This scenario is only feasible through the inherent core value of the information provided by IoT/IIoT applications and services.

The IoT technologies and applications are enablers and accelerators of digital transformation with IoT devices deployed in consumer sector and many several industrial domains. Of the enterprise segment in 2030, 34% of devices will be accounted for by "cross-vertical" use cases such as generic track-and-trace, office equipment and fleet vehicles, 31% by utilities, most prominently smart meters, 5% by transport and logistics, 4% by government, 4% for agriculture, and 3% each for financial services and retail/wholesale. The consumer Internet and media devices will account for one third of all devices in 2030. Smart Grid, including smart meters, will represent 14% of connections and connected vehicles, will represent 7% of the global installed base [4].

In the very near future, every company will either buy or sell information as this resource continues to increase its relevance in the value chain, creating strategic advantages in business models and empowering technology strategies for companies.

Digital technology is rapidly transforming the socio-economic fabric and the codes of ethics that rule algorithms-writing, access and exploitation of information have a critical impact on how the societies of the future evolve.

A common element in all these developments is that digitisation creates a great amount of information. A considerable part of this information reveals how objects work internally and as elements of more complex setups. Accordingly, many innovative technological installations offer creative solutions concerning how to collect and process this information and how to take necessary action.

The challenge with this information is related to how things interact with each other and with the environment while exhibiting behaviour that is often like human behaviour (also related to contextual processing). This behaviour cannot be accurately handled by robots, drones, etc., so this is where technologies, such as swarm logic and AI, come into play.

Security-perceived threats almost always trigger interactive installations equipped to sense and react to surrounding parameters. Changes in these parameters can be visualised, increasing the chances of real threats being detected and asserted.

Due to advanced visualisation techniques, the threat landscape is better defined and digested by the involved actors (being users, LEAs, or other involved agencies). While security used to be primarily about securing information, the landscape has widened considerably. The timely transfer of information, threat identification, isolation and correct and traceable actions all rely on security protection.

IoT ecosystems evolve together with security strategies, which have to account for the layered architecture, where all things, encryptions, communications and actions must be protected against a growing number of diverse attacks, whether via hardware, software or physical tampering.

The IoT system is a group of agents with non-coordinated individual actions that can collectively use local information to derive new knowledge as a basis for some global actions. The intelligence lies both in agents (AI) and in their interactions (collective intelligence). At the core of swarm logic is the sharing of information and interactions with each other and the surroundings to derive new information. However, this collective intelligence is prone to several attacks, especially related to malicious nodes sending false information to influence the decision-making system. Thus, reputation and trust management systems should be in place to be able to identify malicious or misbehaving system agents/nodes and remove them from the system until they behave normally again. These types of attacks can be easily identified and corrected at the edge of the network without having to move all the information to the cloud. Swarm agents can locate and isolate the threat and then converge towards a common point of processing. This is visualised by depicting the real-time state of the agent's movement.

Swarm-designed security is inspired by nature; hence, if IoT can uncover behaviour patterns (of birds, ants, etc.), it may also be capable of meeting security challenges with well-functioning solutions.

3.3.2 IoT/IIoT Platforms Evolution

An IoT platform is defined in [55] as an intelligent layer that connects the things to the network and abstract applications from the things to enable the development of services. The IoT platforms achieve a number of main objectives such as flexibility (being able to deploy things in different contexts), usability (being able to make the user experience easy) and productivity (enabling service creation to improve efficiency, but also enabling new service development). An IoT platform facilitates communication, data flow, device management, and the functionality of applications. The goal is to build IoT applications within an IoT platform framework. The IoT platform allows applications to connect machines, devices, applications, and people to data and control centres [55]. The functionality of IoT platforms covers the digital value chain of an end-to-end IoT system, from sensors/actuators, hardware to connectivity, edge, cloud, and applications. The platforms enable communication between IoT devices, manage the data flows, support application development and provide basic analytics for connected devices [56].

There are 620 IoT Platform companies on the world market, up from 450 IoT Platforms companies in 2017 according to [56]. The report analysis is based on the evaluation of 1,414 IoT projects used for tracking IoT platforms which highlights the importance and pervasiveness of IoT platforms in bringing IoT solutions to market.

The IoT platforms are classified based on their development origin in device-, connectivity-, edge-, cloud-, industrial-, or open-source centric IoT platforms.

The IoT device-centric platforms implement the functions for provisioning tasks to ensure connected devices are deployed, configured, and kept up to date with regular firmware/software updates. The IoT connectivity-centric platforms implement capabilities and solutions for connecting the IoT device, managing and orchestrating connectivity, and provisioning communication services for connected IoT devices.

The edge-based IoT platforms provide the computational and analytics AI-based capabilities close to the edge devices where data is generated. IoT edge platforms deliver the management capabilities required to provide data from IoT devices to applications while ensuring the availability of management services for the devices over their lifetimes. Several IoT edge platforms are used for supporting developers to create, test, and deploy IoT applications or services. The edge IoT platforms need to use hardware-agnostic scalable architectures to support the deployment of sets of functionalities across various types of IoT hardware without modifications.

Cloud-based IoT platforms are offered by cloud providers to support developers to build IoT solutions on their clouds. Infrastructure as a Service (IaaS) providers and Platform as a Service (PaaS) providers have solutions for IoT developers covering different application areas.

The advanced analytics AI-based platforms were initially developed separately to offer tools including ML and other AI techniques and streaming analytics capabilities to extract actionable insights from IoT data. The new IoT platforms include more and more the functions provided by the analytics platforms. However, the advanced analytics AI-based platforms are advancing towards providing AI training/learning tools for different AI-based algorithms.

The IIoT platforms have been developed based on the industrial manufacturing requirement to monitor IoT edge devices and event streams, support and/or translate a variety of manufacturer and industry proprietary protocols, analyse data at the edge and in the cloud [57]. The IIoT platforms integrate and combine IT and OT systems in data sharing and consumption, enhance and supplement OT functions for improved asset management life cycle strategies and processes and enable the application development and deployment.

The IIoT platforms embed AI-based components at different architectural layers and support the security, safety-, and mission-critical requirements associated with industrial assets and their operating environments.

Recent innovations in the field of IIoT, AI, and edge computing have accelerated the developments of IIoT platforms that provide a collection of functions for edge device management, IoT data AI-based analytics, modern sensor technologies and connectivity solutions that enhance industrial equipment and industrial operations with remote monitoring, predictive maintenance, and extensive information analytics. The IIoT platforms support the creation of intelligent, self-optimising industrial equipment and production facilities through the technology functions that are provided. These functions are device management that includes SW/HW that enables manual and automated tasks (e.g. to create, provision, configure, troubleshoot, and manage securely fleets of IoT devices and gateways remotely, in bulk or individually); integration that provides SW/HW, IIoT devices (e.g. communications modules, Supervisory Control and Data Acquisition (SCADA) systems, Programmable Logic Controllers (PLCs), etc.) tools and technologies (e.g. communications protocols, APIs and application adapters, etc.) addressing the processing, enterprise applications and IIoT ecosystem integration requirements at the edge, across cloud and on-premises implementations for

end-to-end IIoT solutions. The IIoT platforms provide functions for data management that include capabilities to support the ingesting of IoT edge device data, storing data from the edge to enterprise platforms, data accessibility (e.g. for IIoT devices, IT and OT systems, external systems, etc.), tracking the flow of data and implementing enforcing mechanisms for data and analytics governance policies to ensure the quality, security, privacy and value of data.

The evolution of IIoT devices at the edge enable a new phase in the digital transformation providing more intelligent functions, seamlessly controlling and optimising the environment and systems, gathering data at every stage, together with being integrated with IoT/IIoT digital platforms, process automation and artificial/augmented intelligence. Gartner estimates that by 2025, 75% of data generated and processed by enterprises will exist at the edge rather than in the traditional centralised data centre or cloud with the capabilities of edge computing solutions ranging from event filtering to complex-event processing or batch processing that creates a need for enterprises to deploy computing power and storage capabilities at the network edge, or edge computing [3].

The new generation of IIoT platforms include AI-based analytics functions that integrate processing of data streams, (e.g. device, enterprise and contextual data), to deliver insights into asset state by monitoring use, providing indicators, tracking patterns and optimising asset use. Various AI techniques, rule engines, event stream processing, data visualisation and ML are applied and implemented to enhances the analytics capabilities.

The IIoT platforms have to integrate both the IT and OT security measures, and features and the IIoT platform security function includes the software, tools and practices facilitated to audit and ensure compliance, for establishing and executing preventive, detective and corrective controls and actions to assure privacy and the security of information across the IIoT solution on both IT and OT domains.

The evolution of edge computing for IIoT applications needs to address the secure data storage, efficient data retrieval and dynamic data collection. A data processing framework for IIoT by integrating the functions of data preprocessing, storage and retrieval based on both the fog computing and cloud computing was presented in [31]. The data processing system of IIoT consists of five entities (e.g. IIoT, edge server, proxy server, cloud server and data users). The IIoT continuously collects data from physical environments and then sends the data to the edge server. The time-sensitive data are extracted and processed by the edge server, and then the data is passed in the cloud to

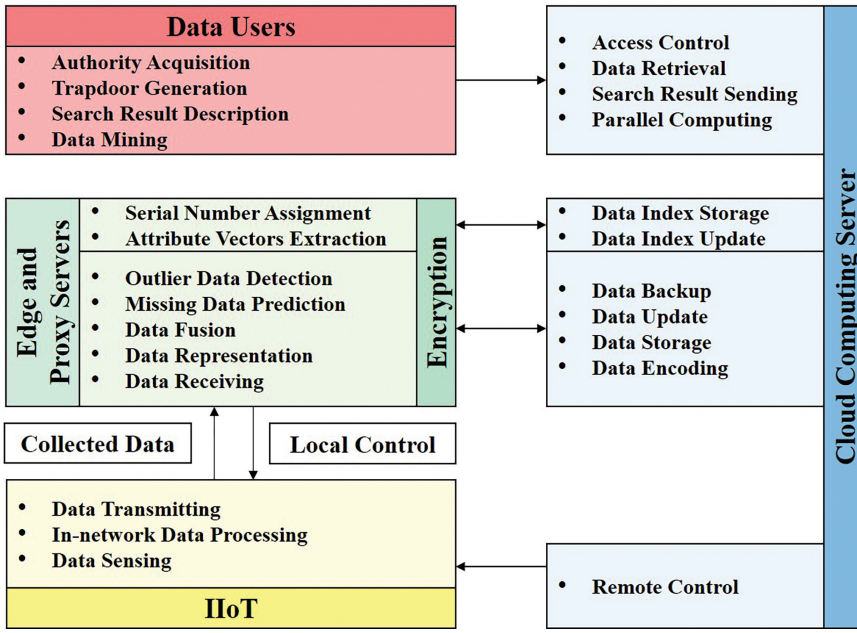

Figure 3.3 Framework of data collection, storage, retrieval, and mining [31].

be organised dynamically. A framework of data collection, storage, retrieval, and mining is presented in Figure 3.3 [31].

The IIoT devices are collecting data that is aggregated and fused at the data level and delivered to the edge server. After receiving the information, the edge server transforms the data into a unified representation framework and fuse the information at the feature level for convenient storage. The false data and missing data are also processed. To improve search efficiency, and support multiple search patterns, corresponding index structures need to be constructed. The data users and IIoT can communicate with the cloud server to execute specific instructions. Due to the amount of IIoT data, the data users likely employ the cloud server to mine the data. The cloud is used for its parallel computing capabilities.

In industrial platforms, IoT devices continuously monitor event-triggered information, which is further transmitted to a remote server, so apprehending monitoring of the industrial outcome.

Combining edge and cloud computing with IIoT data analytics can optimise the network traffic and latencies for ML tasks by employing the edge nodes and the evaluation of the degree of data reduction that can be achieved

on edge without a significant impact on the ML task accuracy. Edge nodes act as intermediaries between IoT devices and the cloud, reducing the quantity of data sent to the cloud.

The edge computing in IIoT is focusing on deploying edge computing into different IIoT scenarios to reduce network traffic and decision-making delay. Therefore, the reference architecture of edge computing in IIoT needs to be improved and refined from the existing edge computing reference architectures.

Edge intelligence is applied on the edge of IIoT to enable edge devices and servers to perform more complex tasks with a higher data processing performance and lower latency. An AI model can be trained to perform predictions and make decisions with high-accuracy, using large amounts of training and verification data. For edge IIoT devices, training and leverage the AI model are hard due to the limited computing and storage resources. To resolve the conflict between the limited resources of edge devices and the high complexity of the AI model two basic approaches are used through enhancing the computing power of IIoT edge devices and simplifying or partitioning the AI model deployed on IIoT edge devices. AI at the edge enhances the range and computational speed of IoT-based devices in industries [34].

One of the advantages of IIoT is the massive amount of real-time data from multiple devices, sites, and infrastructures. Mining data values and making multi-dimensional business decisions improve significantly industrial production efficiency [32].

The traditional IIoT systems are represented in many cases by vertical, closed applications, focusing on maintaining the proper functioning of a machines/equipment or site, and the IIoT systems used constantly creates data islands. Adding edge computing to IIoT could support aggregation the data at lower processing levels and enhance its flexibility.

The complexity of the data security sharing problem is increased in IIoT applications. Opening data islands and sharing the same real-time data securely with any type and number of special applications and stakeholders is required in most of the edge-computing-based IIoT systems. There are two challenges that have to be addressed in edge data sharing: the inevitable increase of data interfaces that can lead to critical consequences (e.g. intrusion and destruction) and the limited performance of IIoT edge devices (e.g. strong/robust security algorithms are difficult to run on resource-constrained IIoT edge devices). The introduction of blockchain in edge computing in IIoT brings new challenges and opportunities for the secure sharing of data [33].

IIoT provides the network infrastructure for connecting IoT devices so that the monitoring and control of industrial manufacturing systems can be supported. From a cyber-physical system perspective, it is composed of both the physical subsystem and the cyber subsystem, which interact with each other so that the manufacturing process can be monitored and controlled with the aid of advanced information communication techniques.

By interacting with computing and networked objects in the physical subsystem, IoT devices (sensors, actuators, etc.) collect data, utilise the network subsystem to transmit the data to the operation centre, in which the data will be further analysed to assist system decision making and receive data to conduct actuation and modification of physical assets.

In IIoT, as numerous applications are time-sensitive, network performance is the key factor that affects the performance of IIoT applications. Nonetheless, to support automation and intelligence for IIoT applications, a large amount of data will be collected and analysed. While more data can provide better intelligence to IIoT applications, transmitting massive data through the network could lead to network congestion and further affect the monitoring and control performance of IIoT applications.

The integration of the IIoT platforms with other platforms in the industrial domains require addressing the application, collaboration and processes layers in the IoT architecture through the implementation of application enablement and management platform function that incorporates software/algorithms that enables business applications in any deployment model to analyse data and accomplish IoT-related business functions, interoperability and the flow of information across end-to-end systems. Application software components manage the operating system, standard input and output or file systems to enable other software components of the platform. The IIoT platforms with the integration of this function introduce application-enabling infrastructure components, application development, runtime management and digital twins that allows users to achieve edge and cloud federation while providing scalability and reliability to deploy and deliver IoT solutions seamlessly and in real-time. The future developments of IIoT platforms will accelerate the integration of AI-based components at all IoT architectural layers through a set of key capabilities, (e.g. data ingestion, processing and transmission) that optimise edge analytics by placing data and secure computing infrastructure closer to the factory floor that leads to improved industrial product quality, operational performance, prediction of downtime and automated operational flows.

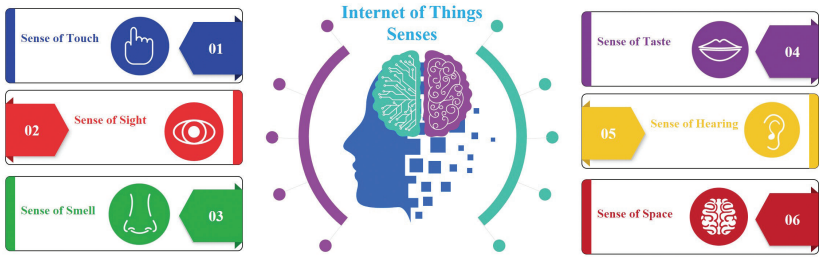

Figure 3.4 Internet of Things Senses (IoTS) overview.

3.3.3 Internet of Things Senses (IoTS)

Internet of Things Senses (IoTS) is part of the IoT concept involving new sensing technologies to reproduce over the Internet the senses of sight, hearing, taste, smell, and touch, enabled by AI, VR/AR, intelligent connectivity, and automation. The IoTS developments are key for the IoT considering that the cognitive decision-making capabilities of the devices can be implemented by AI algorithms implemented into the intelligent IoT devices (e.g. robotic things) at the edge.

The IoTS complements and extends the capabilities of many IoT applications by including other senses (e.g. vision, hearing, touch, taste, smell, pain, mechanoreception-balance, temperature, etc.) and providing new perceptions and experiences by integrating augmented intelligence and information across senses, time, and space. A list of possible human senses is presented in [46].

The IoTS is bringing the technological advances for developing new sensing solutions as part of the intelligent things to address the implementation over the Internet of the senses of touch, sight, smell, taste, hearing, space.

Touch is discussed as part of tactile IoT and consists of several distinct sensations (e.g. pressure, temperature, light touch, vibration, pain, etc.). For humans, the sensations are part of the touch sense and are attributed to different receptors in the skin. For robotic and other types of things, the sensations are generated by sensors that mimic the reaction to pressure, temperature, touch, vibration, etc. Touch sense transmits different other messages in the case of human interaction, which can be transferred as well to things.

Sight is the capability of perceiving things through the vision system that can be represented by different types of cameras, AR glasses that support the navigation, searching for routes, identify places, recognising objects, persons, and sceneries. The sight will improve the capabilities of robotic things and enhance their functional abilities.

The smell is the ability to detect the different odours/scents/aromas (up to 1 trillion scents), that can be represented by different sensors that are sensitive to various odours. The remote smell integrated as online experience for humans and things can improve the capabilities of the robotic things to smell scents in different remote environments, provide new services and improve the perception in these environments.

Taste is the capability to sense different tastes like salty, sweet, sour, bitter, and savoury. The different tastes can be detected by various sensors that can provide a palette of tastes with a determined ranking. Information fusion from the different types of taste sensors is significant to experience a flavour. In many cases, other factors are needed to build the perception of taste in the cognition system of things.

Hearing is the sense to recognise the sound and decode the sound waves and vibrations. The detection of sound using different types of microphones, vibration sensors, etc. allows for enhancing the capabilities of the robotic things by developing techniques for voice control, automatically translate languages, voice biometrics, etc.

The sense of space is based on the information fusion from multiple sensors types and cognition process to comprehend where the things are in space. This is part of the proprioception includes the sense of movement and position of the movable parts of the items. This sense is critical in the future for autonomous intelligent robotic things operating in fleets in different environments. Additional senses are used to sense movement to control balance and tilt the body of an object, sense the direction, acceleration, to attain and maintain equilibrium. Bio-and chemical sensors can be used to detect chemical substances and biologic materials (e.g. viruses, bacteria, etc.).

An Ericsson report [81] found that by the next decade digital sound and vision, complemented by touch, taste, smell and more, will transform the current screen based experiences into multi-sensory ones that are practically inseparable from physical reality. The report explores what that could mean for consumers, with AR glasses as the entrance point and presents what the consumers envisage as future developments driven by IoT sensory connectivity through AI, VR, AR, 5G and automation.

3.3.4 The Evolution of Tactile IoT/IIoT

The Tactile IoT/IIoT is a shift in the collaborative paradigm, adding sensing/actuating capabilities transported over the network to communications modalities so that people and machines no longer need to be physically close to the systems they operate or interact with as they can be controlled remotely.

Tactile IoT/IIoT combines ultra-low latency with extremely high availability, reliability and security and enables humans and machines to interact with their environment, in real-time, using haptic interaction with visual feedback, while on the move and within a certain spatial communication range.

Faster internet connections and increased bandwidth vastly expand the information garnered from onsite sensors within the industrial IoT network. This requires new software and hardware for managing storing, analysing, and accessing the extra data quickly and seamlessly through a Tactile IoT/IIoT applications. Hyperconnectivity is needed to take VR and AR to the next level for uniform video streaming and remote control/tactile internet (low latency).

The TIoT/TIIoT edge encompasses the sensors, actuators, computing, and communication resources deployed at the remote site where the tactile operation is controlled.

The TIoT/TIIoT can be classified based on the nature of the controlled environment (physical, digital, virtual) and based on the type of operator-teleoperator combination (e.g. machines, robotic things, humans).

The TIoT/TIIoT teleoperation in a physical environment (e.g. remote disaster management, cooperative autonomous systems, telesurgery) can be based on the nature of operator and teleoperator a 1) human operator-machine teleoperator, 2) machine operator-machine teleoperator (e.g. typical TIoT/TIIoT with minimal human intervention) or 3) human operator-human teleoperator (e.g. a human operator performs a physical, non-life-critical operation through a human teleoperator for the case of maintenance with the use of haptic-audio-video feedback from the controlled domain).

The TIoT/TIIoT teleoperation in a virtual/augmented reality involves controlling remote things in virtual environments using real-time interactions or simulations by employing digital twins. The teleoperators for these types of applications are robotic things and their digital/virtual representations. Robotic things and machines interacting remotely in virtual environments are used for testing and evaluation of different operational scenarios in the real world and optimise the design parameters for future physical devices. The case of human operator-machine teleoperator addresses the interactions between human operators with a virtual object via force feedback (e.g. immersive, multi-player, networked VR gaming, telemedicine, physiotherapy, etc.).

In the future, coworking with robots in IoT applications will favour geographical clusters of local production ("inshoring") and will require human

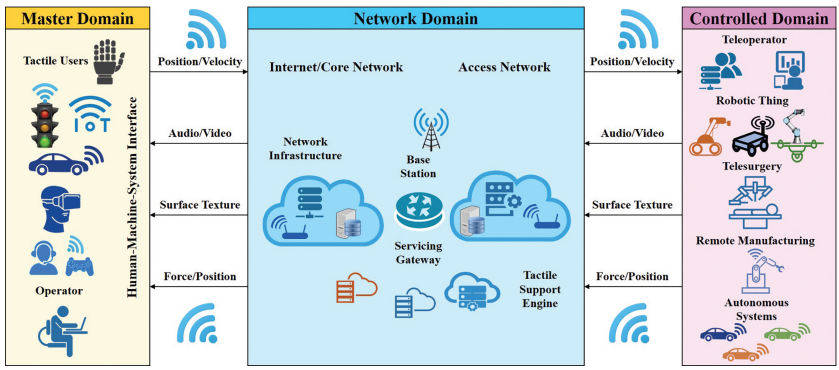

Figure 3.5 Tactile Internet of Things.

expertise in the coordination of the human-robot symbiosis with the purpose of inventing new jobs humans can hardly imagine or do not even know they want to be done. Fibre-wireless (Fi-Wi) enabled human-to-robot communications may be a stepping stone to merging mobile IoT/IIoT, and advanced robotics with the automation of knowledge work and cloud technologies, which together represent the five technologies with the highest estimated potential economic impact in 2025 [168, 169].

The tactile IoT/IIoT (TIoT/TIIoT) enables the real-time remote control and physical (haptic) experiences, and TIoT/TIIoT capabilities support the creation of a spatial safety zone that can interact with other nearby objects connected to robotic things that are part of different IoT applications.

The current network infrastructures are not able to support the emerging TIoT/TIIoT applications in terms of reliability, latency, sensors/actuators, access networks, system architecture and mobile edge-clouds. The design requirements of TIoT/TIIoT systems/devices to achieve real-time interactions are still dependent on the monitoring of the underlying system/environment based on human (or human-like) senses limited by the perception processes. TIoT/TIIoT applications need to adapt the feedback of the system to human reaction time.

TIoT/TIIoT applications have ultra-low end-to-end latency and ultra-high reliability design requirements and need to ensured data security, availability and dependability of systems without violating the latency requirements and considering the encryption delays and the end-to-end processing loop. The centralised architectures cannot meet these requirements, and more decentralised and distributed network architectures based on mobile-edge computing and cloudlets need to be developed to bring the TIoT/TIIoT

applications closer to the end-users [170]. The wireless access networks used for these type of applications need to provide novel resource allocation techniques, feedback mechanisms, interference management and medium access control techniques in order to meet the stringent reliability and latency requirements of TIoT/TIIoT applications [157].

The enhancement in various aspects of physical and Medium Access Control (MAC) layers, emerging network technologies including Software Defined Networking (SDN), Network Function Virtualisation (NFV), network coding and edge/fog computing are part of the technologies that are supporting the TIoT/TIIoT applications in future connectivity networks.

The Tactile Internet standardisation is addressed in the IEEE Pl918.X "Tactile Internet" [165] with the scope of defining a framework, incorporating the descriptions of its definitions and terminology, including the necessary functions and technical assumptions, as well as the application scenarios [162]. Within this framework, IEEE P1918.X defines the architecture technology and assumptions in Tactile Internet systems with IEEE P1918.X.1 dedicated for Codecs for the Tactile Internet, IEEE P1918.X.2 focussed on Al for Tactile Internet and IEEE Pl918.X.3 addressing the MAC for Tactile Internet. The Industrial Internet Consortium (IIC) is working on developing a standard for low-latency TIoT/TIIoT for different smart cyber-physical systems including smart transportation systems, smart manufacturing, and smart healthcare systems [166, 167].

In the next generation of Internet of Robotic Things (IoRT) applications, it is expected that more network intelligence will reside closer to the robotic things. Several functions of IoRT applications can be implemented using TIoT/TIIoT ultra-low end-to-end latency and ultra-high reliability design to ensured data security, availability and dependability of IoRT systems across the end-to-end processing loop. This will lead to the rise of edge cloud/fog and Mobile Edge Computing (MEC) distributed architectures, as most data will be too noisy, latency-sensitive, or expensive to be transferred to the cloud. IoRT next intelligent generation requires to address the issues of unstable and intermittent data transmission via wireless and mobile links, efficient distribution and management of data storage and computing, interfacing between edge computing and cloud computing to provide scalable services and, finally, mechanisms to secure IoRT applications. To ensure the development of intelligent robotic things, it is necessary to reduce the amount of data processed and sent to the cloud. This requires to use a set of data functions for quality filtering and aggregation, and to fusion more functions into intelligent devices and gateways closer to the edge.

Moving from data centres and cloud computing into distributed edge computing infrastructures based on high-performance HW/SW platforms is opening the way for achieving the required latency constraint for TIoT/TIIoT implementation. The processing at the edge reduces the round-trip latency of transferred information, provides an efficient method for offloading information delivered to the core network, provides high bandwidth, offers new services and applications by accessing the network context information.

The master domain in TIoT/TIIoT includes the human and a human–system interface, together with machine and machine-system interface that transforms the human and machine operation into the control information by sensing technologies. The controlled domain consists of the teleoperators (e.g. remote things, such as robotic things) in remote environments that can be controlled by the master domain.

The network domain is an intermediary between the human/machine in the master domain and the remote environment. The sense of touch is generated by imposing an operation on the ambience and feeling the environment by a change or reaction force through the bidirectional haptic communication.

In TIoT/TIIoT, the human/machine delivers the command information to the remote thing. Then the remote thing interacts with the remote environment and feeds back the haptic information to the human/machine in the master domain. The haptic feedback information can contain two types of information (e.g. kinesthetic and tactile feedback information).

The kinesthetic feedback information is employed by the "things" in the master domain to estimate the force, location, speed, and torque in the target remote environment. The tactile feedback information is utilised to determine the roughness parameters of surface texture and friction information in remote environments.

The feedback information supports closing the global control over human/machine, communication network and remote environments allowing the master and controlled domains to exchange energy and tasks with each other based on the "real-time" interaction of commands and feedbacks.

The connectivity requirements of the TIoT/TIIoT/IoRT are matched by the capabilities of cellular networks. For example, 4G and 5G networks with dedicated radio base stations can be used to ensure that traffic remains local. In this case, on-site cellular network deployment with local data breakout ensures that critical production data do not leave the premises using quality of service (QoS) mechanisms to fulfil use case requirements and optimise reliability and latency. Critical applications can be executed locally and independently of the macro network through cellular network deployment with

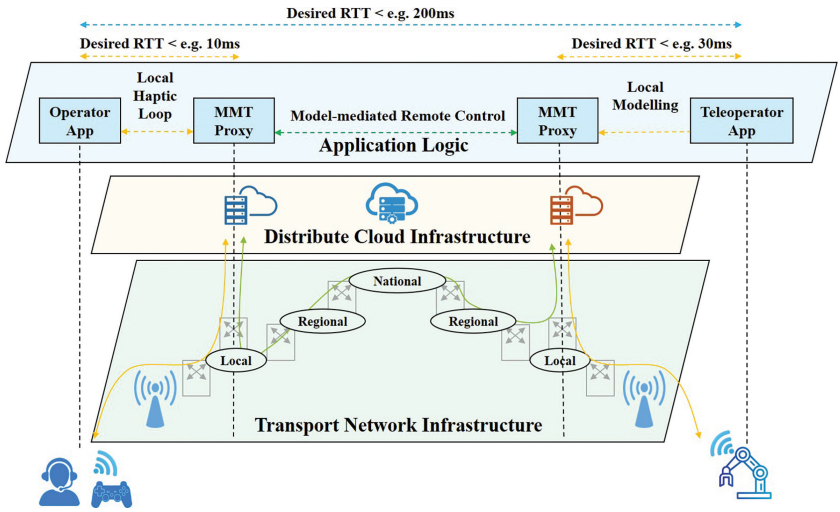

Figure 3.6 Long-distance haptic feedback loops through a national 5G network including MMT proxies deployed close to the operator and teleoperator [47].

edge computing. Edge computing, multi-access edge computing, processes data created around the network and the edge robotic things devices enable entry into core networks and computation is largely or completely performed on distributed device nodes, rather than primarily taking place in a centralised cloud environment.

An evaluation of the 5G ultrareliable and low-latency communications (URLLC) radio configurations in a reference urban macro network deployment, with 500 m distance in-between base stations sites for tactile internet devices located outdoors, was performed in [47]. The full outdoor coverage that was expected to be provided for latency was in the range 1–2 ms and 4–6 ms for the investigated NR-based and LTE-based URLLC configurations, respectively. The end-to-end latency for the tactile internet services depends on the service requirement, the latency components introduced by different domains (e.g. master, network, including multidomain network orchestration, controlled domain, etc.) along the end-to-end information path. Tuning and adapting the parameters to determine the optimal end-to-end latency can be implemented in an automated manner, to enable flexible, scalable, and cost-efficient service deployments.

The solution proposed in [47] uses a model-based approach for the service specification and configuration. It includes haptic control proxies instantiated within the 5G communication system tactile internet services

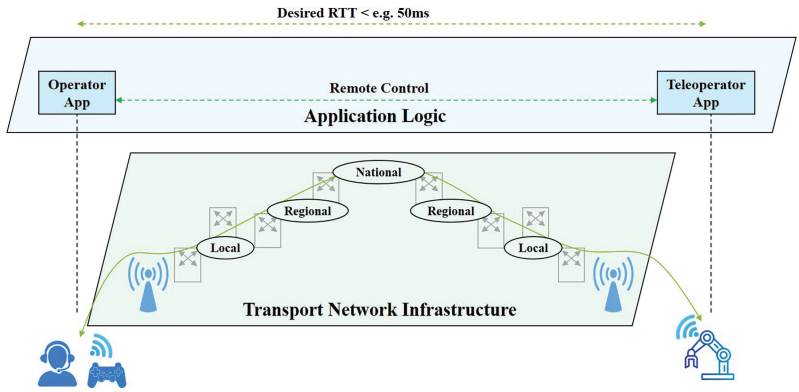

Figure 3.7 Long-distance haptic feedback loop through a national 5G network for the TDPA scheme, without MMT proxy [47].

when a specific end-to-end latency is exceeded. The haptics communication flow for a long-distance remote operation on a national level is illustrated in Figure 3.6. Long-distance haptic feedback loops through a national 5G network including Model-Mediated Teleoperation (MMT) proxies deployed close to the operator and teleoperator [47]. for teleoperating a drone-mounted robotic arm.

For this case, the MMT-proxy software is running on distributed cloud edge computing infrastructure located in vicinity to the remote "thing" (e.g. drone-mounted robotic arm) and operator. The corresponding haptics loop flow through the network for the time-domain passivity approach (TDPA) method (without a proxy) is shown in Figure 3.7. The authors in [47] used MMT as preferred method over the TDPA method only in case of a round trip time (RTT) that is larger than some threshold (e.g., 50 ms), implying that the selection of the method to use (and whether to use an MMT proxy or not) should ideally depend on the achievable RTT [47].

There are several challenges to be addressed for realising haptic communications over the wireless TI and TIoT/TIIoT as listed below [48–50].

Haptic Edge Sensors/Actuators – These edge devices are essential for collecting/capturing and transferring the information into correct actions. The haptic sensors sense the tactile information by interacting with the environment and are mounted both at the master domain (e.g. machines) and at the teleoperators end (e.g. remote things, such as robotic things) in remote environments. The sensed information is transmitted to the master domain in the form of the haptic feedback by the haptic actuators (also called haptic feedback devices). In the case of TIoT/TIIoT, the haptic sensors are mainly

pressure sensors to detect the underlying pressure (e.g. capacitive, resistive pressure sensors). For IoTS, other types of sensors are used for mimicking the human and beyond senses.

The quality of the sensors, precision, range, sensitivity, response time, spatial resolution, cost, placement, temperature dependence and complexity are important design parameters for different TIoT/TIIoT use-cases in different environments. The information received from the sensors is converted in the different actions (that provide a feeling of touch similar to the feeling the receiver would get in the real-world context), by using haptic actuators (e.g. cutaneous – muscle type for force tension, kinaesthetic – skin type for vibration, pressure, pain, temperature, etc.). The implementation of lightweight energy-efficient, fast response time, low-cost, actuators providing capabilities of both the cutaneous and kinaesthetic feedback is one of the main challenges for TIoT/TIIoT deployments.

Surface Sensing and Actuation – The TIoT/TIIoT use-cases can require more than a single-point contact for the tactile and kinesthetic feedback as the master and controlled domains address applications requiring touch-based sensations across the surfaces (e.g. palm of the hand or other areas for humans or robotic things). Surface sensing implies the use of multiple arrays of sensors and techniques to identify the forces across the surface. Reading the sensing surface-based or distributed sensing and actuation requires technologies that increase the energy use and latency and impact the communication requirements due to the increased data rates and a different perception for the case of an information loss.

Multi-Modal Sensory and Information Fusion: Different types of sensing/actuation edge devices are used for collecting/providing the tactile information and other perception data (e.g. sound, visual, etc.) to increase the perception capabilities and performances. The use of multi-modal sensory and the fusion of the data from these sensors increases the various requirements for TIoT/TIIoT in terms of latency, fusion, transmission rate and sampling rate. Multiplexing and fusion the information from different types of sensing devices is another challenge for the implementation of TIoT/TIIoT real-time multiplexing schemes across different protocol layers for integrating the various modalities in dynamically varying wireless environments.

Collaborative Multi-User Haptic Communications - TIoT/TIIoT use cases integrate a multitude of edge humans/things that interact and collaborate in a shared remote environment, requiring the creation of peer-to-peer

overlay to enable collaboration among multiple users. This overlay formation step brings new challenges in terms of meeting the requirements of TIoT/TIIoT applications as the overlay routing, and IP-level routing may further increase the end-to-end latency [49].

Efficient Resource Allocation – The different tasks in the master and controlled domains need to get specific communication channel parameters in the network domain based on the TIoT/TIIoT use case requirements. The efficient radio resource allocation in wireless/cellular networks brings new implementation challenges (e.g. the resources need to be shared among haptic, human-to-human, machine-to-machine, machine-to-human communications that have various and sometimes conflicting service requirements) due to the incorporation of haptic communications. Besides, as the haptic communications are bidirectional, symmetric resource allocation with the guarantee of minimum constant rate in both the uplink and the downlink need to be ensured. These requirements bring new issues for managing and orchestrating the wireless/cellular networks parameters to provide priority for resources based on QoS, safety-, mission-critical features. Flexible resource allocation techniques across different protocols layers, including adaptive management and network slicing with on-demand functionality, are needed for future deployments.

Haptic Codecs – The transmission of the hepatic information in digital form requires sampling the signals at rates of 1 kHz [48], information that is then compressed and transmitted across the wireless/cellular networks. The future challenges are developing standardised groups of haptic codecs that can be integrated into the kinesthetic and tactile information, that are energy efficient and able to perform effectively in time-varying wireless environments.

Ultra-High Reliability – The TIoT/TIIoT applications require ultra-high reliability for wireless/cellular networks. Several factors, such as the lack of resources, uncontrollable interference, reduced signal strength and equipment failure, impact on network reliability. To provide the reliability requirements for the wireless/cellular networks the different layers of the protocol stack including the MAC layer, transport layer and session layer need to be reconsidered to enable ultra-high reliability in haptic communications. Trade-offs between reliability, latency, packet header to the payload ratio must be considered [48–50].

Ultra-Low-Latency – The future TIoT/TIIoT applications will require end-to-end latency below 1 ms across different protocols-layers, air interface, backhaul, hardware and core Internet. Optimising the transmission

parameters at each layer and across the layers is required to achieve the overall end-to-end latency (e.g. upper bound fixed by the speed of light). At the physical layer providing shorter Transmission Time Interval (TTI) can lower the over-the-air latency, at the expense of higher needed bandwidth.

Stability for Haptic Control – The TIoT/TIIoT global control loop across the master, network and controlled domains need to be stable to avoid the degradation of the "continuum experience" to the remote environment. This is a challenge as the global loop integrates the humans/machines, the communication network, the remote environment, and the energy/tasks exchange among these components takes place via various commands and feedback signals [54]. The wireless environments can have time-varying delays and packet losses, and new techniques must be developed to reduce the instability due to communication channels parameters variations.

Performance Metrics – The TIoT/TIIoT requires evaluation methods beyond Technology Readiness Level (TRL) that include new QoS metrics, Quality-of-Experience (QoE) and Quality-of-Task (QoT), which ultimately will lead to Experience Readiness Level (ERL) classification. This will allow identifying new suitable performance metrics for analysing and comparing the performance (e.g. information fusion, connectivity features, data processing techniques, data reduction/control, compression, etc.) of various haptic systems over the TIoT/TIIoT [47]. The introduction of experience evaluation allows to analyse the difference of the physical interaction across a network and the same manipulation carried out locally and measure the accuracy by which a tactile user can perform a task [47].

The research challenges across different domains of TIoT/TIIoT are summarised in Table 3.2.

A generalised framework for TI beyond the 5G era that can be applied to TIoT/TIIoT is presented in Figure 3.8 [51].

The basic architecture of TIoT/TIIoT is composed of a master domain, a network domain and a controlled domain and this is reflected in the generalised framework [51] comprising various aspects of TI including key technical requirements, main application domains, a basic architecture and enabling technologies. The key technical requirements of TI include ultra-responsive connectivity, ultra-reliable connectivity, intelligence at the edge network, efficient transmission and low-complexity processing of tactile data as presented in the previous subsections.

Table 3.2 Research challenges across different domains of TIoT/TIIoT

TIoT/TIIoT Domain	Research Challenges
Connectivity	• Ultra-low latency ($<$ 1ms) • Ultra-high reliability ($>$ 99,999%) • Very-high data rates (Gbps-Tbps) • Very-high backhaul bandwidth • Support for communication overhead with cloud/edge/fog networking infrastructures
Computation	• Online processing of haptic feedback for near real-time interactions • In field processing to reduce ingress transmission • Support high-computation AI processing and training/learning at the tactile edge
Artificial Intelligence	• Predict reliable end-to-end communication • Movement and action prediction to compensate physical limitations of remote latency • Inferring movements and actions techniques • Acquisition and replication of skills for optimising the TIoT/TIIoT global loop
Haptics	• Codecs to acquire remote interactions • Methods to vary modalities for interactions • Machine-to-machine and machine-to-human coordination mechanisms and interfaces • Techniques for increasing the stability of haptic control • Surface-based sensing and actuation • Sensing fusion methods (e.g. multiplexing of multi-modal sensory information) • Performance metrics and ERL methods
AR/VR	• Localisation and tracking precision and efficiency • Scalability and heterogeneity • Quality-data rate-latency trade-offs • In-network vs. in-VR (edge) computation • Information processing theoretical advancements
Collaborative autonomous mobility systems	• Dynamic rout selection for autonomous systems • Traffic management for IoRT • AI/ML/DL techniques for prediction in real-time movements/actions • Multi-autonomous things perception/control for safety (fail-operational), traffic/movements efficiency • Ultra-low latency and ultra-high reliability design techniques for connecting (V2X) multiple collaborative autonomous mobility systems in real-time.

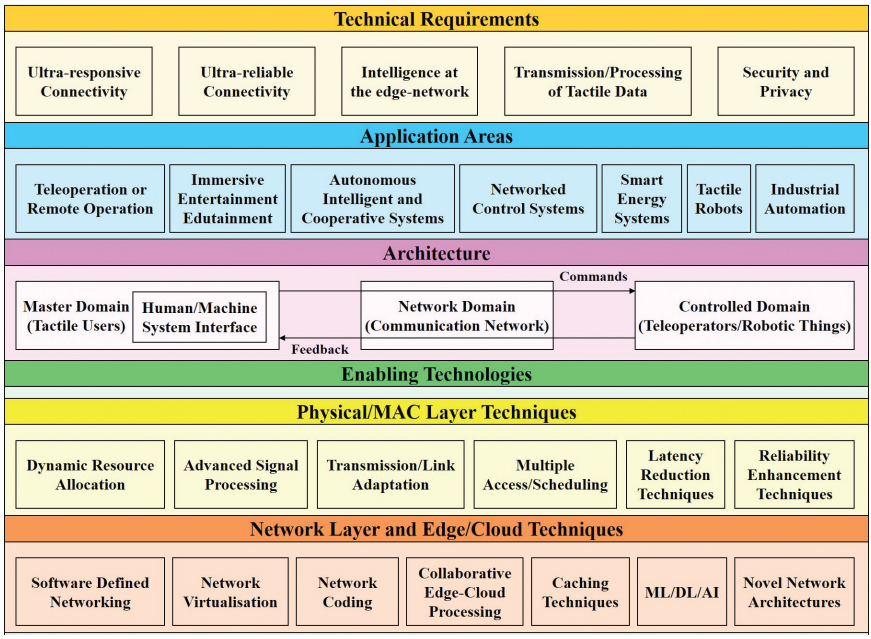

Figure 3.8 A generalised framework for TI. Adapted from [51].

3.3.5 IoT Digital Twins

Digital twins (DTs) are virtual representations of physical assets and things across their life cycle using real-time sensor data. The DTs can be utilised to expose a set of services allowing to execute certain operations and produce data describing the physical thing activity.

The concept was introduced by NASA as part of its spacecraft monitoring mission and is now a mainstream approach as technologies such as sensors, data analytics and edge computing fuel a new generation of IoT/IIoT devices, industrial assets, buildings and intelligent city infrastructure. In IIoT applications, DTs can be used in digital factories that consist of multi-layered integration of the information related to various activities along with the factory and associated resources.

A digital twin is comprised of a virtual object representation of a real-world item in which the virtual is mapped to physical things in the real-world such as equipment, robots, or virtually any connected business asset. This mapping in the digital world is facilitated by IoT platforms and software that is leveraged to create a digital representation of the physical asset. The

digital twin of a physical asset can provide data about its status, such as its physical state and disposition. Conversely, a digital object may be used to manipulate and control a real-world asset by way of teleoperation of DT modelling solution.

Digital twins use AI, ML and software analytics with data to render real-time digital simulation models that can update and change as their real, physical counterparts, or "twins" change. These IIoT systems can be used to optimise the operation and maintenance of physical assets, systems, and processes in real-time. Operational intelligence can be used towards building digital twins. The insights produced by real-time intelligence enable operators to understand the performance of distributed infrastructure, make predictions, improve efficiency, and even prevent disasters. Operational intelligence is one catalyst for the DT concept as it supports to digitise infrastructure, monitor operations in real-time, predict events, take actions based on intelligence, and engage with different stakeholders.

There are different definitions for Digital Twins across industry and academia. As described by the Digital Twin Consortium [138], a digital twin is an abstraction of something in the real world. It may be physical (a device, product, system, or other assets) or conceptual (a service, process, or notion). A digital twin captures the behaviour and attributes of its physical sibling with data and life cycle state changes potentially moving in either, or both, directions.

As per CIRP Encyclopedia of Production Engineering [139], a digital twin is a digital representation of a unique active (real device, object, machine, service, or intangible asset) or unique product-service system (a system consisting of a product and a related service) that comprises its selected characteristics, properties, conditions, and behaviours by means of models, information, and data within a single or even across multiple life cycle phases.

The advances in IoT/IIoT technologies accelerate the development of tools and technologies that can support the design of digital models for the IoT digital devices twins integrating 3D modelling tools, computer-aided engineering (CAE) software with IoT/IIoT platforms. The diversity and heterogeneity of IoT/IIoT devices is reflected as well in the data formats used for representing, designing the digital twin. The next-generation IIoT platforms are designed to include capabilities for building digital twins and simulate the interactions between them as well as simulate different what-if scenarios that will support preparedness, impact, and mitigation management.

The capabilities of these platforms include key elements such as data integration, the accuracy of the physics simulation software and the capacity to update physics models in digital twins with real-time data streaming from IoT devices placed in the field.

Thanks to technologies, such as blockchain, swarm logic and AI, digital twins now have these capabilities. In the pursuit of better security, digital twins can trigger and simulate threat scenarios in the digital world, as well as optimise the security strategy to handle such scenarios should they occur in the real world.

In the context of IoT, digital twins are the representation of physical IoT devices that offer information on the state of the physical twin, respond to changes, improve operations, and add value. The digital twin, as a virtual representation of the IoT's physical object or system across its lifecycle, using real-time data to enable understanding, learning, and reasoning is a one-element connecting the IoT and AI.

The digital twin represents the virtual replica of the IoT physical device by acting like the real thing, which helps in detecting possible issues, testing new settings, simulating all kinds of scenarios, analysing different operational and behavioural scenarios and simulating various situations in a virtual or digital environment, while knowing that what is performed with that digital twin could also happen when it is done by the "real" physical "thing".

Digital twins as part of IoT technologies and applications are being expanded to more applications, use cases and industries, as well as combined with more technologies, such as speech capabilities, AR for an immersive experience and AI capabilities, enabling to look inside the digital twin by removing the need to go and check the "real" thing. The digital twins' evolution is shown in Figure 3.9

Digital twins for IoT must possess a minimum of attributes:

- Correctness – give a correct replication of the IoT ecosystem and its devices.
- Completeness – updated vis a vis the functionality in the real-world system.
- Soundness – exhibit only the functionality available in the real-world system.
- Abstractness – free from details specific to implementations.
- Expandability – adapt easily to emerging technologies and applications.
- Scalability – must be able to operate at any scale.
- Parameterisation – accessible for analysis, design, and implementation.

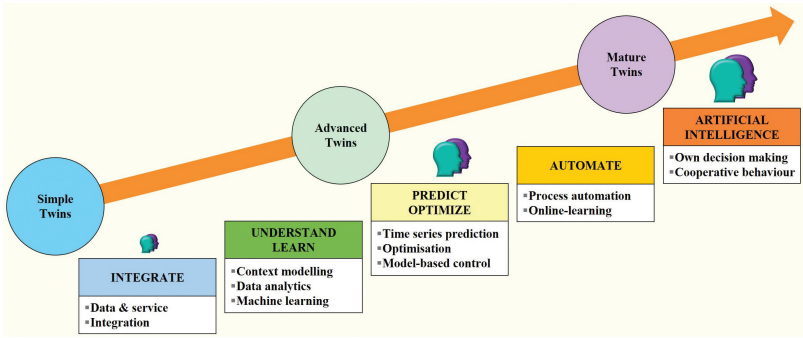

Figure 3.9 Digital Twins evolution. Adapted from IBM.

- Reproducibility – be able to replicate the same result for the same input as the real system.

For IoT, digital twins can expand the interface between man and machine through their virtual representation and advanced technologies on levels, such as AI and speech, which enable people and devices/machines to take actions based on operational data at the edge (provided by IoT devices and edge computing processing).

To fully exploit the potentials of IoT's digital twins several gaps need to be addressed to foster the market uptake and adopt the technology in several application scenarios.

There is a need to put more effort into the validation and quantification of the benefits perceived using DTs against exiting processes and systems. Both the scale and the type of benefits and improvements need to be formalised, and the selection of the type of digital twins should be justifiable. There is an emerging need for an evaluation framework that can categorise the various levels of sophistication of digital twins. These levels should help the industry to use a common language when describing a digital twin and its capabilities. An evaluation framework that is based on 5 distinct levels has been proposed for the built environment [140]:

- Level 1: A digital model linked to the real-world system but lacking intelligence, learning or autonomy.
- Level 2: A digital model with some capacity for feedback and control, often limited to the modelling of small-scale systems.
- Level 3: A digital model able to provide predictive maintenance, analytics, and insights.

- Level 4: A digital model with the capacity to learn efficiently from various sources of data, including the surrounding environment. The model will have the ability to use that learning for autonomous decision making within a given domain.
- Level 5: A digital model with a wider range of capacities and responsibilities; ability to autonomously reason and to act on behalf of users (AI); interconnected incorporation of lower-level twins.

Another aspect that will further boost DT innovation is the consideration of solutions that span across the entire product life cycle. Current frameworks and implementations focus on specific use cases and specific life cycle phases. It is critical to understand the requirements of the DTs at each phase of the life cycle, as well as the required number and interdependencies of digital twins required for example in the production phase in contrast to the operational and maintenance phase.

Other important aspects to be addressed that may accelerate future DTs solutions are (i) standardisation and interoperability aspects such that virtual entities can communicate and interoperate, (ii) the data ownership of Digital Twin data (iii) aspects related to data privacy and personal data protection.

Several standardisation activities have emerged in recent years. The International Organisation for Standardisation (ISO) covers industrial data in TC 184 SC 4 [141]. The standard for a Digital Twin Manufacturing Network is currently under development (ISO/DIS 23247) [142]. ISO/IEC JTC1 provided a technology trend report by its joint advisory group on Emerging Technology and Innovation (JETI). In the report, "Digital Twin" was identified as the number one area needing in-depth analysis [143]. Within the ISO/TC 184 a working group for Digital Twins has been created ISO/TC 184/AHG 2 [144]. The IEEE Standards Association initiated a project IEEE P2806 that aims to define the system architecture of digital representation for physical objects in factory environments [145]. In ITU-T SG 13 study group [146], requirements and capabilities of a digital twin system for smart cities is under study. The recently established Digital Twin Consortium is a program of Object Management Group dedicated to the widespread adoption of digital twin technology and the value it delivers [138]. It aims at contributing to standards by deriving requirements that will be submitted to international standards' development organisations, such as Object Management Group and ISO/IEC.

BuildingSMART International (bSI) is another initiative, committed for creating and developing open digital ways of working for built asset

environment and their standards help asset owners and the entire supply chain work more efficiently and collaboratively through the entire project and asset lifecycle. They recently published a position paper "Enabling an Ecosystem of Digital Twins" [147], where three areas identified to focus further developments. They are closely related to the topic of standardisation: (i) standards for data models, (ii) standards for data management and integration and (iii) data security and privacy.

Projects funded under the H2020 work programme such as IoTwins project [148], which aims to build a reference architecture for developing and deploying distributed and edge-enabled digital twins of production plants and processes will act as facilitators for Digital Twin implementations. More initiatives in this direction are required.

The idea of the implementation of National Digital Twin infrastructures on local and/or national level has also gained attention. A good example is "Virtual Singapore" [149] a research and development programme initiated by the National Research Foundation of Singapore, which is a dynamic three-dimensional (3D) city model and collaborative data platform, including the 3D maps of Singapore. When completed, Virtual Singapore will be the authoritative 3D digital platform intended for use by the public, private, people and research sectors. It will enable users from different sectors to develop sophisticated tools and applications for test-bedding concepts and services, planning and decision-making, and research on technologies to solve emerging and complex challenges for Singapore.

Another example is the National Digital Twin Programme of the Centre for Digital Built Britain [150]. The programme seeks to deliver a smart digital economy for infrastructure and construction, and to transform the UK construction industry's approach to the way we plan, build, maintain and use our social and economic infrastructure for the future. The programme's objectives are to deliver the Information Management Framework, to enable the National Digital Twin and to align industry, academia, and Government on this agenda.

A DT life cycle and supporting tools and functionalities to exploit the approach are presented in Figure 3.10 [35]. The model can be applied to DTs used in IoT applications together with the functions associated. The DTs for IoT will be represented in the design phase by a logical object that is the first software representation of the physical object or thing. In the production phase, the digital twin is released, and the software representation is used to test and experiment with the future physical product/thing. The software aspects of the DT support optimising the physical item and in carrying out

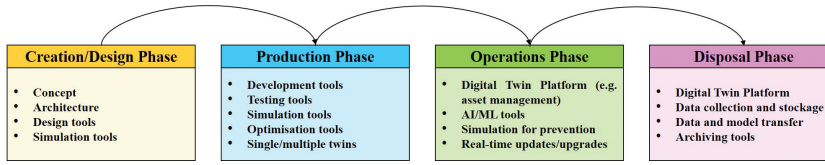

Figure 3.10 Life cycle of a DT – Functions and tools. Adapted from [35].

the validation, testing and optimisation. In the operation phase, the DT runs in the digital domain as a one to one representation of the real thing. In some cases, several physical things can have a DT representation that represents the capabilities of the class of instantiated physical things. In many IoT applications, the production and operation phases are used to manufacturing physical items or products and later use the DT to mirror the properties of the real thing to measure, simulate and optimise the behaviour and performance of products.

Mirroring the characteristics of the real thing into the DT requires that there be a continuous link between these entities that allows for updates/upgrades. Depending on the type of application, the link/connection can be real-time, permanent/intermittent, resilient, etc. The flow of information is predominantly from the physical thing to the DT, with specific situations when the DT sends data and information to the physical thing. In all IoT/IIoT applications, the DT must be continuously synchronised with the production system and the IoT device that it represents. The synchronisation supports the use of the DT to simulate and predict the behaviour of the real thing in a new scenario and use case. Using the bidirectional exchange of information based on real data measurement and events flowing from real to virtual can improve the accuracy of the simulations and predictions.

The DT can be used in the TIoT/TIIoT to support the optimisation of the global loop (master, network, slave domains) by identifying through simulations the parameters of the physical thing to perform in specific environments and conditions.

A representation of an architectural model and general framework for DT is illustrated in Figure 3.11 [35]. The description is following a layering concept that can be mapped to the 3D IoT reference architecture. The first layers are mapped to the physical and network layers in the 3D IoT reference architecture. The upper layers integrate the properties of the DT, into the processing, storage, abstraction, and service IoT architecture layers where different components implement the functions for modelling

of objects, instantiation, self-management, orchestration, entanglement, and other. Collecting and contextualising the data as well as to execute data analysis and information inferring for the DT must be integrated into the IoT data processing and abstraction.

Semantics and ontologies need to be aligned with the ones defined for IoT devices and integrated into the service layer. The DT simulation functions must be implemented as part of the IoT abstraction and service layer components. The use of open APIs that can be programmable at different levels and with different abstraction capabilities support the interaction of IoT applications with different IoT platform functions using structured data.

Applying the digital twins to different sensors, is exemplified by the work on IEEE 1451 smart sensor digital twin federation for cyber-physical systems (CPS) [36]. The IEEE 1516 high-level architecture (HLA) is a standard for the modelling and simulation of distributed, heterogeneous processes. The digital twin developed is a digital simulator or digital replica of a real IEEE 1451 smart sensor. The DT emulates both desired, non-linear behaviours and failure modes to simulate an actual sensor in the field [36].

Several features of the DTs in IIoT applications are the following [37]: connectivity (e.g. the ability to communicate with other entities and DTs), autonomy (e.g. the possibility for the DT to live independently from other entities), homogeneity (e.g. the capability to allow the use of the same DT regardless of the specific production environment), customisation flexibility (e.g. ability to modify the behaviour of a real thing by using the functionalities exposed by its DT), and traceability, (e.g. capability to follow the traces of DT's activity of the corresponding physical item). One of the challenges with DTs in IIoT applications is data interoperability as data moves from physical devices and equipment in the field (e.g. manufacturing floor), to the software used in data IoT twin modelling systems. Different IIoT applications can use several different digital twins, at different levels (e.g. physical component, asset, system, process, etc.). The hierarchy of IoT digital twins can produce different perspectives and generate different types of data and complex relationships between data sets that can result in data interoperability issues.

An architectural model, which incorporates digital twins into edge networks for real-time data analysis and network resource optimisation for smart manufacturing, is presented in Figure 3.12 [38]. The framework contains three layers the user layer, edge layer, and digital twin layer. The user layer consists of client devices in IIoT such as smart machines, vehicles, IoT devices and can be mapped to the physical layer in the 3D IoT reference architecture.

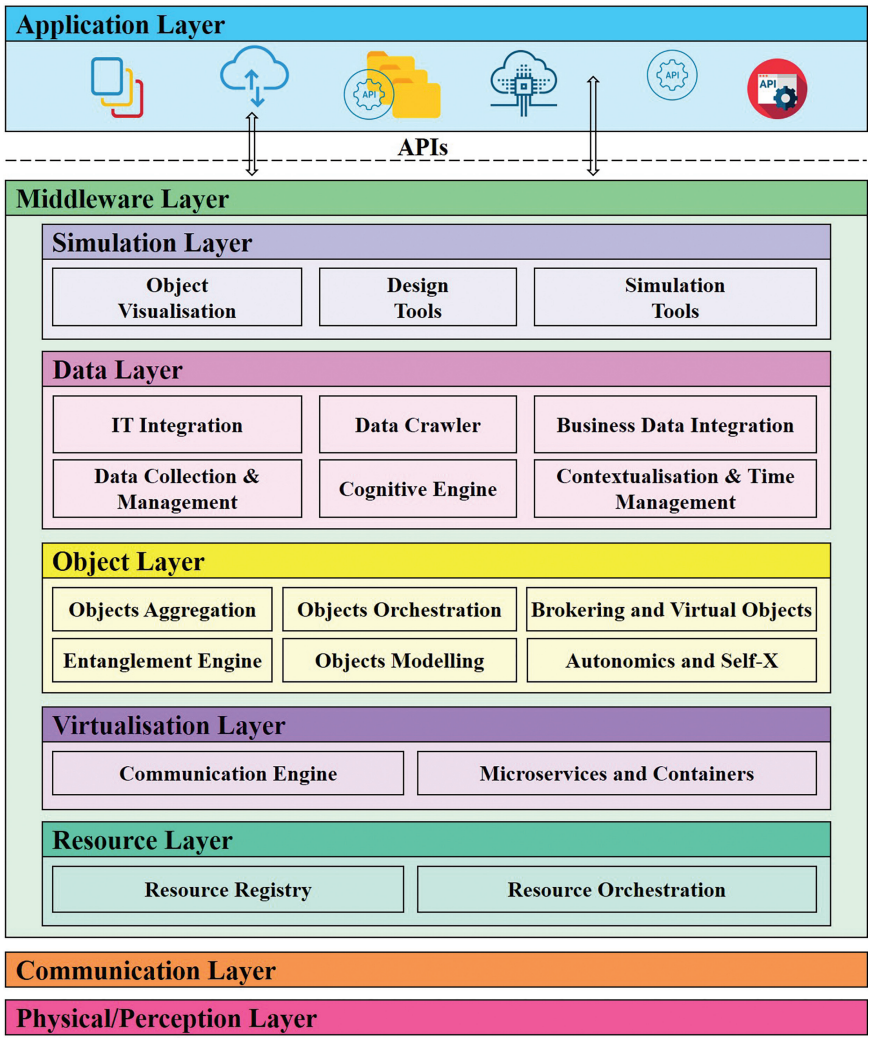

Figure 3.11 DT general framework. Adapted from [35].

The edge layer is composed of base stations that are equipped with MEC servers. The edge layer can be mapped to the network and processing layer in the 3D IoT reference architecture. The base stations are connected with user devices under their coverage via wireless communications. The digital twin layer can be mapped to the abstraction and service layers in the 3D IoT reference architecture.

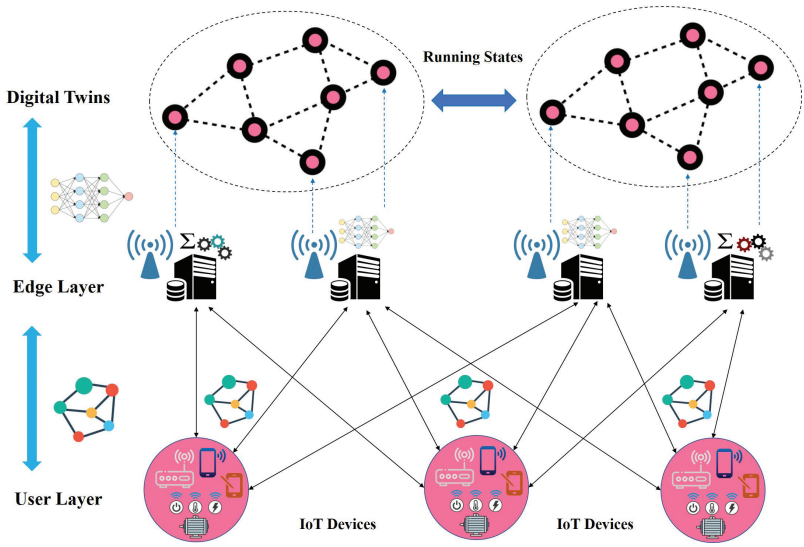

Figure 3.12 Digital Twin Edge Networks for IIoT. Adapted from [38].

To model digital twins, the authors used federated learning to build digital twins from the historical running data of devices. The raw data transmission is avoided, and data privacy is enhanced by federated learning. An optimisation problem was formulated that aimed at reducing the communication cost of federated learning and provided the solution by decomposing it and using DNN for communication resource allocation. Numerical results on the benchmark real-world dataset corroborated that our proposed mechanism could improve the communication efficiency and reduce the overall energy cost [38].

It is evident that standardisation efforts, research, and development funding programmes, as well as reference implementations on local, national, and international level with public-private funding schemes, will be the main drivers towards acceleration and market up-take of DT technology.

3.4 Internet of Things Augmentation

The advances in IoT technologies enabled by heterogeneous integration of functions to sense, collect, act, process, infer, transmit, create notifications of/for, manage and store information allows the interactions between autonomous systems and humans in a continuum of physical, digital, virtual and cyber environments.

IoT augmentation is a blend of methods, techniques and technologies that are applied to improve the sensing, action, or cognitive abilities of IoT devices. The concept is based on the integration of the capabilities offered by augmented reality (AR), virtual reality (VR), digital twins, and AI, to generate and visualise virtual 3D models of the real world that are evolving into smart and interactive environments related to the context of things for physical objects. The concept expands to humans by providing interactive digital extension of human capabilities (e.g. replication, supplementation) by using IoT sensing and actuation technologies, AI, fusion and fission of information, AR, VR, and digital twins to improve human productivity and capabilities.

IoT devices act as bridge between physical assets, digital and cyber infrastructure while the AR/VR brings the digital content and the digital twins by interacting with the physical objects in real-time in the virtual environment.

These new developments allow the technicians accessing real-time IoT data on the shop floor and augmented reality IoT devices aware of the spatial configuration of the environment of the operator can sense what the operator is looking/focusing at to intuitively display only the data needed for the operation to be performed.

IoT augmentation expands its potential in healthcare for virtual monitoring and simulations of the living environment integrating the information from various IoT devices connected in real-time for measuring vital health parameters and information combined with environment and context monitoring.

In retail and the smart, connected supply chain, IoT augmentation bridges the online and offline environments providing the users with the tools to model, simulate and visualize the products and how those products can be customised to fit the user's needs or interact with other products.

From vehicles to mobile devices, the manufacturing processes require putting together hundreds of various components in a precise and predetermined order. Using a 3D design superimposed onto the actual process provides more straightforward access and stepwise guidance. Implementing an IoT-based connected supply chain assures that machines parts are always stocked and ready for use.

IoT augmentation extends to augmented senses, augmented actuation and augmented cognition that applies both to humans and machines/things and enhance the human-machines interactions and collaboration.

As the IoT augmentation develops, the need for standardisation is evident as the new IoT applications encompass various systems and technologies that

deliver real-time information, make inferences, provide analytics and take decisions that have to be interoperable with other systems with which they are interacting and collaborating.

The standardisation needs to address the fail operational functions for IoT augmentation that are important features that are required to be integrated into the system to provide the functional safety and eliminate the sources of error and ensures that the right parts, processes and sequences are always followed.

The next wave of IoT/IIoT applications will include a form or another of IoT augmentation that need to scale, dynamically adapt to the context and seamlessly adopted for multiple functions and in diverse environments, and operation conditions to enhance the human and machines capabilities.

3.5 Edge Computing

The model of centralised cloud computing, including data analysis, and storage for IoT/IIoT is limiting the capabilities of IoT applications, creating silos, challenges regarding interoperability and data spaces. To further develop their capabilities and provide new opportunities for enterprise data management, IoT/IIoT applications (e.g. real-time) and services are moving the developments toward the edge and, as a consequence, most of IoT data generated and processed by enterprises will exist at the edge rather than in the traditional centralised data centre in the cloud. Edge computing provides significant benefits such as reduced latency, data analysis close to the data source, lower bandwidth, and reduced energy consumption in networks.

IoT/IIoT technologies including IoT devices, advanced autonomous IoT systems, IoT intelligent servers providing gateways capabilities, end-point-enabling networks form the foundation of edge computing, accelerating the move of IoT applications to edge in industries where IoT is deployed, such as energy, utilities and manufacturing. Edge computing allows the easy deployment of private and local area networks with all the capabilities associated to the IoT/IIoT systems.

IoT systems deliver disconnected or distributed capabilities into the embedded IoT world and edge computing is part of the technological fabric across the industrial sector that deliver these capabilities empowered with advanced and specialised processing resources, data storage, and analytics. The IoT applications powered by edge computing allow to keep the traffic and processing local, to reduce latency, exploit the capabilities of the edge and enable higher autonomy of the IoT devices at the edge.

Edge computing is redefining the IoT/IIoT, embedded, and mobile processor landscape, accelerating the development of high-performance circuits using AI/ML techniques and embedded security for addressing the edge and deep edge data analysis processing. Edge computing provides mechanisms for distributing data and computing on the edge, which makes IoT applications much more resilient to malicious and no malicious events. Distributed deployment models are expected to address more efficiently connectivity and latency challenges, bandwidth constraints and higher processing power and storage embedded at the edge of the network. In addition, they preserve privacy as raw data is processed locally and only aggregated data is shared in the cloud. Using efficiently the edge computing layer of the IoT architecture move most of the data traffic and processing closest to the end-user applications and devices that generate and consume data. The use of IoT edge, near edge and deep edge capabilities and diverse edge systems, centralised cloud services will enhance the functionalities of cloud technology to provide, manage and update software and services on edge and near edge devices. Centralised cloud services could become hubs in coordinating and federating operation across highly distributed edge devices, and in aggregating and archiving data from the edge or intermediate gateways and servers. The centralised cloud services for intelligent IoT applications will be used as robust and additional scalable ML and sophisticated processing capabilities linked to traditional back-office processing.

The edge computing through different forms (e.g. mobile edge, fog, dew computing, etc.) creates multi-dimensional architecture consisting of a wide range of heterogeneous "things" with different sensing/actuating, connectivity, processing and intelligence capabilities connected to services/applications in a dynamic mesh linked by platforms and distributed services located at the edge/cloud level.

From an architecture perspective, the computing landscape is represented by a tier model as represented in [25] with many alternative implementations of hardware and software at each tier, but all of them are subject to the same set of design constraints. In this model the Tier-1 represents the data centres and the cloud, Tier-2 is organised into small, dispersed data centres called cloudlets representing the essence of edge computing, Tier-3 represents the gateway-base processing devices, and Tier-4 represented by the edge physical devices. The future computing landscape is evolving, including new computing paradigms such as distributed AI in the future IoT and the emergence of new programmable quantum architectures, new compilers that target emerging quantum machines. Future decentralised and

distributed architectures and computing paradigms for IoT/IIoT inspired by neuromorphic computing with hybrid implementations based on architectures for neural DL that could provide optimising computing in such systems leading computing paradigm.

Embedding sensors/actuators, storage, compute and advanced AI capabilities in IoT edge devices and integrate them within mesh architectures, will enable dynamic, intelligent, responsive, peer-to-peer IoT end-devices that exchange information with enterprise IoT platforms and conduct peer-to-peer exchanges with other IoT devices operating across different industrial sectors.

The integration of data analysis technologies in IoT devices is raising edge security challenges that have to be addressed to set and enforce security, privacy and compliance standards for the vastly increased number of devices on the edge, dealing with expanded network types and connections, and various software deployments supporting key features such as low latency, the ability to perform deterministic real-time computing, the support for mission-critical or safety-critical use cases, and the ability to extend computing beyond humans to the extremes of the environment and IoT devices.

IoT devices at the edge need to embed privacy-by-design mechanisms (e.g. encryption of static and dynamic/moving data, anonymisation and credential protection against physical intrusion), together with an increased authentication process governing device participation and data exchange in the cloud, in order to ensure the full authentication and encryption of all traffic between cloud resources and edge IoT devices.

Edge computing is filling the gap between the centralised cloud and the need for a decentralised processing medium for IoT devices and paves the way for IoT connectivity (e.g. 5G and beyond), and AI convergence. Mobile edge computing (MEC), and other edge computing paradigms such as Mobile Cloud Computing (MCC), fog computing, and cloudlets are complementary, possibly competing, options for filling the gap. MEC offers new opportunities for network operators, service, and content providers to deploy versatile and uninterrupted services on IoT applications. MEC supports IoT by providing IoT devices with significant additional computational capabilities through computation offloading.

Edge computing provides low latency connectivity by processing the information closer to its source and to the end user. Edge computing supports the implementation of 5G and AI to intelligently and efficiently manage the network edge by provisioning load balancing, supporting multiple levels of nodes for hierarchical networking, allowing for resource pooling, universal

orchestration and management, multiple access modes providing each edge network node with the resource applications it requires, improving reliability, security, resiliency, supporting virtualisation, mobile IoT applications, providing agility with a horizontal platform and supporting all vertical markets and becoming more scalable by moving computation, networking, or storage capabilities across or through levels of hierarchy.

3.5.1 Edge Computing Architectures

As the IoT technologies have evolved and moved from centralised to decentralised and distributed computing, the architectures needed to perform data analysis at the edge, and the development of the edge cloud computational capabilities as applied to distributed node IoT edge have generated new concepts such as edge cloud, edge computing with specific capabilities (e.g. fog, mobile edge computing, dew computing, etc.) that offer complementary and additional features than the ones provided by cloud computing. The evolution was possible due to the technology progress in the network field (e.g., 1000 times bandwidth increase and significant cost reduction), in the computing processing power (e.g., increased computing power and reduced cost), and in the storage capacity (e.g. 10 000 times capacity increase of a single disk followed by cost reduction).

AI edge processing today is focused on moving the inference part of the AI workflow to the device, keeping data constrained to the device. There are several different reasons why AI processing is moving to the edge device, depending on the application. Semiconductors companies are providing dedicated microcontrollers and microprocessors with hardware dedicated to accelerating DL processing coupled with tools for easy mapping of algorithms in those architectures [41]. Privacy, security, cost, latency, and bandwidth all need to be considered when evaluating cloud versus edge processing [29]. Tractica's report [29] provides a quantitative and qualitative assessment of the market opportunity for AI edge processing across several consumer and enterprise device markets. The device categories include automotive, consumer and enterprise robots, drones, head-mounted displays, mobile phones, PCs/tablets, security cameras, and smart speakers. The report includes segmentation by processor type, power consumption, compute capacity, and training versus inference for each device category, with unit shipment and revenue forecasts for the period from 2017 to 2025. The report predicts that AI edge device shipments will increase from 161.4 million units in 2018 to 2.6 billion units worldwide annually by 2025.

Cloud computing is characterised by a high-processing and computer power, centralised architecture, high latency, centralised data analytics, and AI capabilities, implementing centralised cybersecurity mechanisms and providing very large storage capacity. Edge cloud computing offers a decentralised architecture, medium-low latency, dedicated bandwidth based on the computing needs, AI processing capabilities, cybersecurity mechanisms and networking effect through connectivity with the other IoT edge site nodes. Edge computing is further extending the decentralised architecture and moving towards distributed computing, offering low/ultra-low latency, efficient use of bandwidth, local networking, processing, and storage capabilities. These new architectures provide more efficient in energy and cost solutions.

Edge computing is used to process real-time data storage and computation on the device or data source, rather than sending it to a remote data centre, thus decreasing latency and reducing the bandwidth used by IoT devices. When necessary, the centralised cloud can serve as storage facility for large amounts of data and additional processing. The IoT devices where data analysis and control occur act as nodes. Specialised AI circuits used in IoT devices can process more data on the edge in real-time, which combined with the fact that the devices are closer to the data source, results in faster responses and actions while decreasing latency and saving on bandwidth and equipment costs. Edge computing is addressed in the standards setting communities at ISO/IEC JTC-1 SC-41 from a network centric perspective.

In summary, edge computing provides the functions of a distributed open platform at the network edge, close to the IoT devices and other data sources, integrating the capabilities of collecting information, processing information, networking and exchange of information, storage, analytics, services and applications. Edge computing and IoT technologies are enablers for digitisation of industry offering dynamic connectivity, real-time services, data optimisation, application intelligence, security, privacy protection and delivering IoT edge intelligence services and applications.

The term *edge cloud* is used to describe the decentralisation of the traditional, large cloud data centres with moving cloud storage and compute closer to the edge source while also scaling down the size. Edge locations may connect to each other or to a central cloud for added data inputs and processing or storage capabilities, or isolated in instances of a data breach or service compromise. Edge cloud requires additional remotely administered data centres that are called edge sites, close to the end users. The edge sites are placed at specific locations where increased compute and processing

is needed beyond what can be completed at the edge in conjunction with low-latency, time-sensitive IoT operations.

3.5.2 Deep Edge Computing

In many new IoT applications, the data is collected and processed at deep edge (such as sensor nodes) and send via a communication channel to the gateways and processing units at the edge. Learning and inference will happen on the edge and cloud and the decisions are communicated back to the nodes at the edge. The system architecture including the deep edge depends on the envisioned functionality and deployment options considering that at the core of these devices are cognitive sensor modules that can acquire, understand, and react to data.

Deep edge computing architecture rely on the principles of distributed computing with computing units having limited resources and relying on high efficiency energy consumption and connectivity functions.

IoT devices are becoming more intelligent, equipped with various sensors using new computing paradigms for processing the information and providing the "intelligent deep edge", allowing the IoT applications to become ubiquitous and merge into the environment where various IoT devices can sense its environments and react intelligently to it. Providing AI capabilities to IoT devices significantly enhance their functionality and usefulness, especially when the full power of these networked devices is harnessed – a trend that is often called AI on the edge.

To support connected or collaborative operation, the intelligent sensor modules include at least one communication channel that supports the necessary bandwidths and latencies. Using the channel, the local controller can inform other edge devices or an edge-based service of its current context.

3.5.3 Fog Computing

Fog computing, a term created by Cisco, refers to extending computing capabilities to bring cloud computing capabilities to the edge of the network. Fog enables repeatable structure in the edge computing concept, so enterprises can push compute out of centralised systems or clouds for better and more scalable performance. A Fog computing implementation is a virtualised platform, located between cloud data centres (hosted within the Internet) and end user devices, providing strong support for IoT and complementary to cloud computing platforms.

Fog computing is defined in [1] as an horizontal, system-level architecture that distributes computing, storage, control and networking functions closer to the users along a cloud-to-thing continuum. This horizontal architecture provides support for multiple industry verticals and applications. domains, delivering intelligence and services to users and business. The cloud-to-thing continuum of services assure that services and applications are distributed closer to the IoT devices or things, and anywhere along the continuum between cloud and things. The system level concept is covering the cloud-to-thing continuum over the network edges and across multiple protocol layers without depending on specific communication and protocol layer.

An important component of the Fog computing layered architectures is the Fog service orchestration layer that provides dynamic policy-based life-cycle management of Fog services and the distributed orchestration functionality as the underlying Fog infrastructure and services.

Fog computing has the processing capabilities at the LAN end while data is gathered, processed, and stored within the network, using IoT gateways or fog computing nodes (FCN). Information is transmitted to gateways from various sources in the network and the information is processed in FCN, and the processed data and additional commands are transmitted to the other devices. Using these mechanisms, the fog computing implementation enables a single, processing device to process data received from multiple edge devices and send information where it is needed, with lower latency than centralised cloud computing processing. The Fog computing architecture is scalable and can be integrated with the different cloud computing architectures.

Fog computing has several key differences compared with edge computing as it works with the cloud and has a hierarchical structure. Fog addresses computation, networking, storage, control, and acceleration. Fog computing is an extension of the traditional cloud-based computing model where implementations of the architecture can reside in multiple layers of a network's topology. Fog architectures selectively move compute, storage, communication, control, and decision making closer to the network edge where data is being generated in order solve the limitations in current infrastructure to enable mission-critical, data-dense use cases [1].

Fog computing is described in [2] as a huge number of heterogeneous (wireless and sometimes autonomous) ubiquitous and decentralised devices communicate and potentially cooperate among them and with the network to perform storage and processing tasks without the intervention of third-parties. These tasks can be for supporting basic network functions or new services and

applications that run in a sandboxed environment. Users leasing part of their devices to host these services get incentives for doing so.

Fog computing is based on a decentralised computing infrastructure, where computing resources and application services are distributed in the most logical, efficient place, at any point along the continuum from IoT data source to the cloud. The fog provides high data processing efficiency due to the reduction of amount of data to be transported to the cloud for data processing, analysis, and storage, increasing the security and optimising the data transfer.

In many IoT applications Fog can be used to offload the storage and computations from the IoT devices by using a network edge. The Fog-based IoT network offers a number of design improvements compared with the cloud infrastructure [44, 45] as listed below:

- Better offloading and reduced server strain considering that Fog computing allows the network to offload the data processing to its cloudlets, which reduces data traffic and server strain and allow that more users can be managed more efficiently with edge computing.
- Scalability through parallelism by using Fog computing architecture, decentralisation and adding edge devices to the network, while the cloud computing scales a cloud server by increasing its size in terms of data capacity.

For IoT applications, fog computing offers significant amount of storage at or near the IoT nodes data collection (avoiding primarily to store in large-scale data centres), efficient communication close to the IoT nodes (avoid routing through the backbone network) and optimised management, including network measurement, control and configuration, performed close to the IoT devices.

3.5.4 Mobile Edge Computing

Mobile edge computing relates to computing at the edge of a network and the edge can be seen as a distributed cloud with proximity close to the end user that delivers ultra-low latency, reliability, and scalability.

The aim of Mobile Edge Computing (MEC) is to provide an IT service environment and cloud-computing capabilities at the edge of the mobile network, within the Radio Access Network (RAN) and near mobile subscribers. MEC is a new development in the evolution of mobile base stations and the convergence of IT and telecommunications networking. The main expected

goal of MEC is to reduce latency, to ensure highly efficient service delivery, with the result of providing an improved user experience. This goes together with improved and much more efficient and transparent network operation.

Based on MEC parameters such as characteristics, actors, access technologies, applications, objectives, computation platforms, and key enablers a MEC taxonomy is presented in Figure 3.13 [7]. The implementation of MEC in real applications is support by several key enablers, which contribute to provide context-aware, low latency, high bandwidth services to the mobile subscribers at the RAN close proximity as listed below [7]:

- Cloud and Virtualisation: Virtualisation allows to create variant logical infrastructure in the same physical hardware, with the computing platform at the edge of the network creating different virtual machines using virtualisation technology to provide different services of cloud computing (e.g. Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS), and Infrastructure-as-a-Service (IaaS)).
- High Volume Servers: Mobile Edge Servers are deployed in each mobile base station of the edge network to perform network traffic forwarding and filtering and executing the offloaded task by the edge devices.
- Network Technologies: Multiple small cells are deployed in Mobile Edge Computing environment with Wi-Fi and cellular networking as main networking technologies used to connect the mobile devices with the edge sever.
- Mobile and IoT Devices: Mobile and IoT devices at the edge network compute low intensive task, and hardware related tasks which are non-off loadable to the edge network. Mobile devices perform peer-to-peer computing within edge network through Device-to-Device communication.
- Software Development Kit: Software Development Kit (SDK) with standard Application Programming Interface (API) supported in adapting existing services and foster on expediting the development of new elastic edge applications. These standard APIs can be easily integrated in application development process.

MEC can create opportunities for new use cases and complementary roles for mobile operators as well as for application and content providers by enlarging their business models to better monetise the mobile broadband experience.

MEC enables the creation of new and innovative services over the mobile network for consumers, enterprise customers as well as industries with

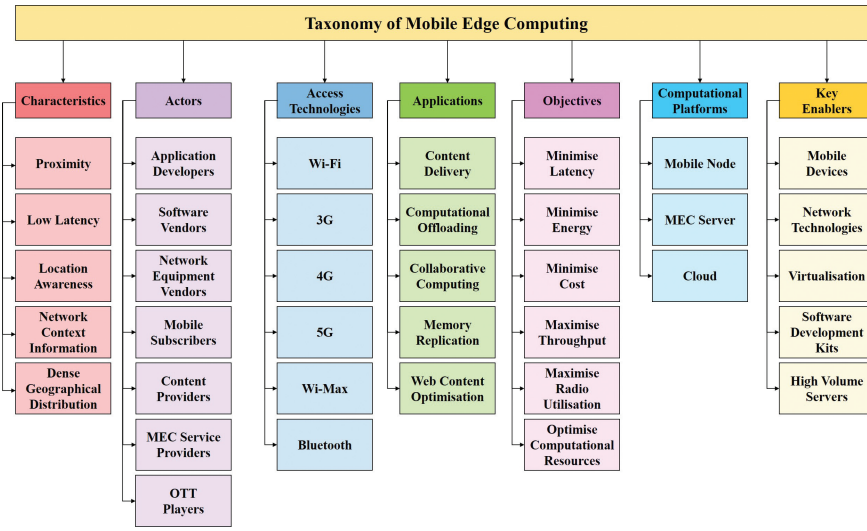

Figure 3.13 Taxonomy of Mobile Edge Computing. Adapted from [7].

mission-critical applications. To this extent, MEC must rely on a standardised, open environment to allow the efficient and seamless integration of applications across multi-vendor Mobile Edge Computing platforms. This was a major driver for the creation of a standardisation effort in support of MEC.

3.5.4.1 The Industry Specification Group (ISG) on Multi-access Edge Computing

Multi-access edge computing technology is currently being standardised in an ETSI Industry Specification Group (ISG) of the same name launched in 2015 [1, 5]. The ISG MEC has already published a set of specifications (including a "Framework and Reference Architecture" with a first release in 2016 [6] and an updated one in 2018 [8]) focusing on management and orchestration (MANO) of MEC applications, and a large range of APIs for application enablement, service deployment and the User Equipment (UE) application. Multi-access edge computing is shifting compute functions from a centralised location to the edge of the network closer to the end user that enables real-time analytics of video surveillance, vehicle-to-vehicle communications, traffic management and other public safety functions.

Multi-access edge computing as it is deployed currently in the 4^{th} generation LTE networks, is connected to the user plane via one of the options

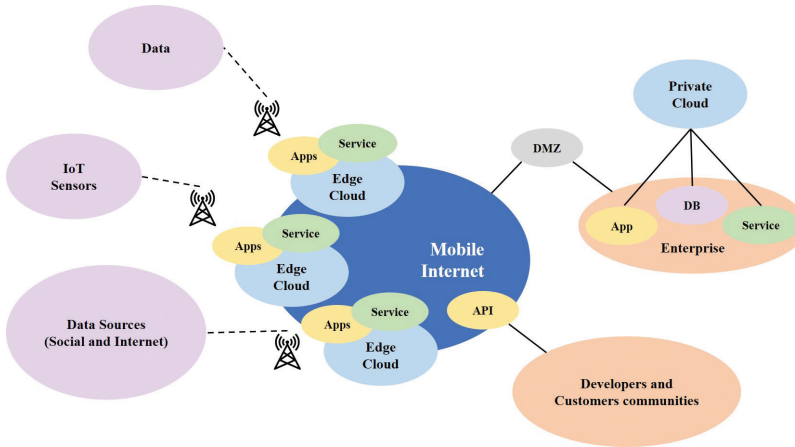

Figure 3.14 Mobile Edge Computing – Improving quality of experience through proximity with end users.

described in [9, 10]. Considering that LTE networks have been already deployed for several years, it was necessary to design the MEC solution as an add-on to a 4G network to offer services in the edge. Consequently, the multi-access edge computing system and its interface specifications, as defined in [8], is to a large extent self-contained, covering everything from management and orchestration down to interactions with the data plane.

As a result, multi-access edge computing servers can be deployed at multiple locations, such as at the LTE macro base station (eNodeB) site, at the 3G Radio Network Controller (RNC) site, at a multi-Radio Access Technology (RAT) cell aggregation site, and at an aggregation point (which may also be at the edge of the core network).

3.5.4.2 A new paradigm for the development of applications

Multi-access edge computing offers to application developers and content providers cloud-computing capabilities and an IT service environment at the edge of the network. Consequently, Multi-access edge computing introduces a standard for supporting an emerging cloud paradigm for software development communities.

Multi-access edge computing in 5G gives service providers and wireless carriers the opportunity to offer distributed sites, including cloud radio access networks, aggregation points, central offices and cell towers as real points for the end users to compute data and store them as part of a virtualised

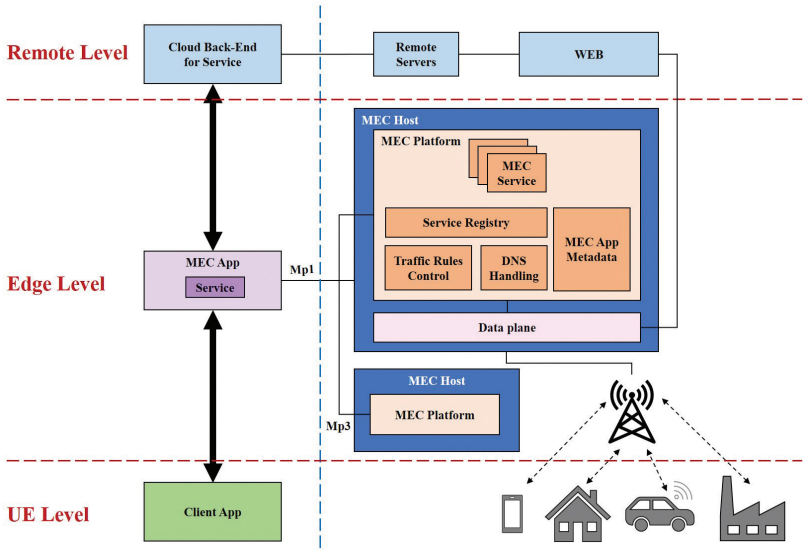

Figure 3.15 The MEC application deployment approach.

and automated cloud network. The role of the edge and multi-access edge computing used for 5G is to reduce network congestion, increase speed, efficiency, scalability and improve application performance by achieving related task processing closer to the user. Edge computing provides more throughput by storing and processing data closer to where it is generated or used.

3.5.4.3 The challenge of network transformation

The deployment of complex, large-scale, and heterogeneous 5G networks is posing a major challenge to the telecom industry. One of the most critical aspects is the issue of management with two main requirements regarding the possible solutions. Firstly, they need to be highly automated and not to require lots of complex (and possibly manual) configuration. Secondly, they need to collect large amounts of relevant data, process them and act on them in an automated fashion with the support of AI and ML.

To this extent, a lot of work has been done in emerging NFV-based architectures and solutions, and there is an opportunity for multi-access edge computing to benefit from the large developments (e.g., the Management and Orchestration (MANO) framework) and deployments done around NFV (the NFV operationalisation) and to extend NFV in order to accommodate the

MEC architecture. To that end, ETSI MEC has defined a "MEC-in-NFV" reference architecture in ETSI Group Specification MEC 003 [6].

3.5.5 Cloudlet

The *Cloudlet* was introduced in [11] as a framework to overcome the overhead of Virtual Machines (VMs) while benefiting of the other features of VMs (i.e. management and reliability issues). The Cloudlet is an edge computing technology that enables new classes of mobile applications that are both compute-intensive and latency-sensitive in an open ecosystem based on cloudlets.

A definition of the cloudlet is given in [14] as a trusted, resource-rich computer or cluster of computers that is well-connected to the Internet and is available for use by nearby mobile devices. For IoT this means the existence of a resource rich computing infrastructure with high-speed Internet connectivity to the cloud that can be used by the IoT mobile devices to augment its capabilities and to enable real-time applications.

Several optimisations in Cloudlet were proposed to reduce the amount of outer data transfer (i.e. the amount of data sent across the network) by processing data within the physical node and provide locality aware global execution.

The Open Edge Computing [23] addresses the Cloudlet based on OpenStack [24]. Open Edge Computing considers that any edge node can offer computational and storage resources to any user in proximity using a standardised mechanism. Edge computing technologies are characterised by openness, as operators open the networks to third parties to deploy applications and services, while their differences enable edge computing technologies to support broader IoT applications with various requirements.

A cloudlet represents a small-scale cloud that provides services like cloud computing with limited resources closer to users. The users are directly connected to other devices in a cloudlet and the latency of the response to user requests is significantly reduced. The cloudlet can ensure better security and privacy. The cloudlet concept forms an initial prototype of fog computing and has four key attributes [22]:

- Maintains only soft state: It is built for microservices and containers and may buffer data originating from a mobile device to a cloud service. Each cloudlet adds close to zero management burden after installation and can be entirely self-managing.

- Powerful, well-connected, safe and secure: It has sufficient compute power (i.e., CPU, RAM, etc.) to offload resource-intensive computations from one or more mobile devices with proper connectivity to the cloud and is not limited by finite battery life. Its integrity as a computing platform is assumed and, in a production, quality implementation this will have to be enforced through some combination of tamper-resistance, surveillance, and run-time attestation.
- Located at the edge of the network: It is logically close to the mobile devices. "Logical proximity" is defined as low end-to-end latency and high bandwidth (e.g., one-hop Wi-Fi).
- Builds on standard cloud technology: It encapsulates offload code from mobile devices in VMs, and thus resembles classic cloud infrastructure such as OpenStack. Each cloudlet has functionality that is specific to its cloudlet role. The cloudlet term is driven primarily by the ISO / IEC JTC-1 SC38 standardisation subcommittee working group.

The cloudlet is enabling resource-intensive and interactive mobile applications such as IoT by providing computing resources to mobile devices with lower latency.

The clouds and cloudlets need a strong isolation between untrusted device-level computations, mechanisms for authentication, access control, and metering, dynamic resource allocation for device-level computations; and the ability to support a wide range of device-level computations, with minimal restrictions on their process structure, programming languages or operating systems. The clouds and cloudlets achieve device and user isolation through virtual machines.

In the case of cloudlets as the mobile and IoT device moves from one physical area to another, its current cloudlet has to hand off the device/user's virtual machine to the new cloudlet.

3.5.6 Dew Computing

Dew computing is defined in [26] as an on-premises computer SW/HW organisation paradigm in the Cloud computing environment where the on-premises computer provides functionality that is independent of cloud services and is also collaborative with cloud services. The Dew computing aims to maximise the potential of on-premises computing and cloud services and use the resources at the lowest level as self-organising systems solving the processing in these environments. Two features describe the nature of dew computing

applications: independence, which indicates that the application is inherently distributed and collaboration that indicates that the application is inherently connected. Collaboration implies that the dew computing application automatically exchange information with cloud services during its operation through synchronisation, correlation, or other kinds of interoperation. Independence represents the ability of the on-premises computing units to provide the requested functionality without cloud services connection meaning that an dew application is not a completely-online application or cloud service [26].

Another definition [28] is considering the Dew computing as a programming model for enabling ubiquitous, pervasive, and convenient ready-to-go, plug-infacility empowered personal network that includes Single-Super-Hybrid-Peer P2P communication link with the aim to access a pool of raw data equipped with meta-data, which can be rapidly created, edited, stored, and deleted with minimal internetwork management effort (i.e. offline mode) on local premises. The Dew computing model is composed of six essential characteristics such as. Rule-based Data Collection, Synchronisation, Scalability, Re-origination, Transparency, and AnyTime Any How Accessibility; three service models such as Software-as-a-Dew Service, Software-as-a-Dew Product, Infrastructure-as-a-Dew; and two identity models (e.g. Open, Closed) [28]. IoT can be one major application for Dew where heterogeneous devices act together to perform a set of tasks and do not need to be connected to the Cloud services all the time creating the so call Dew of Things (i.e. DoT). The Dew computing technical challenges include power management, processor utility, data storage, viability of existing operating system, network model, communication protocols, programming principles, database security, and data efficiency exchange.

The Dew computing paradigm is different from the Fog/Cloud paradigms, as the edge and low-level devices cannot be used in a "conventional" programmable way, but they have to cooperate on the lowest level to solve processing needs, and be able to pass (and consume) information from all hierarchical levels.

Dew computing uses on-premises processing units to provide decentralised, cloud-friendly, and collaborative micro services to end-users. Dew Computing is complementary to Cloud computing, and the Dew computing processing units provide on-premises functionality independent of cloud services and exchange information and collaborate with cloud services.

The Dew computing architecture includes modules for addressing the hybrid and extremely heterogeneous information collectors, information distributors, information processors, information presenters and information consumers, at the edge level directly connected to processing units and IoT devices which are part of the common physical environment, and at the highest level interconnected into the global information processing and distribution system. In the Dew computing paradigm, the individual IoT devices are collecting/generating the raw data and are the components of the computing ecosystem that are aware of the context the data were generated in, therefore dew devices must produce and exchange information.

Dew computing is context-aware, giving the meaning to data being processed with data that is context-free, while information is data with accompanying meta-data and the meta-data places the data in a specific context.

A Dew computing architecture for cyber-physical systems and IoT is proposed in [27]. The dew computing implementation in cyber-physical systems allows autonomous devices and smart systems, that can collaborate and exchange information with the environment, and be independent of external systems or act as processing element in connected complex cyber-physical system of systems.

3.6 Artificial Intelligence IoT

AI has developed over time several (sometimes nature-inspired or human-behaviour mimicking) computational methodologies that address complex real-world problems, in particular when mathematical modelling is not able to provide effective solutions. AI is one of the modalities by which a machine is able to perform logical analysis, acquire knowledge, and adapt to an environment that varies over time or in a given context by making use of abilities that allow a machine (i.e. computer, robot, or intelligent IoT device) to perform functions such as learning, decision making, or other intelligent human behaviours. Applications where IoT technologies converge with AI are growing in depth and functionalities, creating new markets and opportunities that, in turn, require clarifying the future technology requirements and the effective ways to integrate more AI techniques and methods in IoT applications.

The basic concepts behind AI have been around since the 1950's but thanks to modern programming techniques (such as python), the availability of huge quantities and qualities of data, open source tools for neural network

training, powerful computing centres and ever improving embedded processing systems, AI is taking off as a world-changing technology today with a special role for ML in general and DL in particular.

ML is a subset of AI that refers to techniques which enable machines to recognise underlying patterns and learn to make predictions and recommendations by analysing data and experiences, rather than through traditional explicit programming instructions. The system adapts with new data and experiences to improve prediction performance over time. DL is a subset of ML. It is based on an ability to learn data patterns and dependencies by using a hierarchy of multiple layers that mimics the neurons connections of the human brain and make up any deep neural network. DL techniques can work with very large data sets by analysing data, recognizing patterns and making prediction on next data. With DL, a computer can train itself with a large set of data collected for this purpose. If the Training stage, in which the neural network learns to classify different patterns, use datasets labelled in advance the process is referred to as Supervised Learning. In the case of unlabelled datasets, the learning process is called Unsupervised Learning and the neural network tries to cluster the dataset into groups with similar patterns. In both cases the result is an Artificial Neural Network (ANN) that contains all the information necessary to carry out the task. The ANN uses the knowledge acquired in the training to infer data features from new incoming data. This is called Inference stage and can be deployed in embedded devices with memory and processing capabilities orders of magnitude smaller than the servers used to train the ANN itself.

AIoT is enabling and accelerating the developments of IoRT applications to achieve more efficient IoRT operations, improve human-machine interactions and enhance data management and analytics.

The edge computing paradigm is a key element in the evolution of IoT platforms towards IoT decentralised and distributed architectures, intelligence at the edge and optimising the use of resources in IoT applications (i.e. communication, processing, energy consumption at IoT devices, subsystems and systems level, etc.). In many applications, the processing is done at the edge, close to the real-time processes (such as in factories and industrial plants). Many of these processes are time-critical, business-critical, privacy-critical, and part of connectivity bandwidth-intensive, and bandwidth-void applications.

In edge computing, the data generated by different types of IoT devices can be processed at the network edge instead of being transmitted to the centralised cloud infrastructure thus creating bandwidth and energy consumption

concerns. Edge computing can provide services with faster response and greater quality, in comparison with cloud computing. The integration of edge computing with IoT provides efficient and secure services for many end-users, and the development of intelligent edge computing-based architectures can be considered a key element for the future Intelligent IoT infrastructure [42].

That real-time system properties are essential in mobility applications such as autonomous driving applications is obvious, but also new applications within manufacturing, predictive maintenance and robotics are real-time dependent to ensure increased quality, efficiency, reliability, and safety. Moving computing capabilities from the cloud to the edge by implementing AI methods, for instance in IoT devices that operate directly or close to the actual operation and independently of external input from the cloud or other computer centre, will in many cases be the preferred or only possible solution. Enabling technologies like IoT are developing fast, including increasing computing and wireless communication facilities for a range of applications, although interoperability issues are still needed to overcome.

Bringing AI in IoT devices will require, on the one hand, to define the adequate processing unit and, on the other hand, to ensure the easy porting of AI apps on the target platform in an effective way. Regarding the processing unit, either it is a microcontroller or a microprocessor, and in many IoT domains the inference engine usually does not require high computational power. Application domains such as activity monitoring, acoustic event detection, predictive maintenance deal with data (in term of sampling rate and size) that can be analysed by a software base inference engine.

More complex application domains, as for example those based on computer vision, deal with huge amount of data to be analysed in real-time. Inference engines for these domains require dedicated hardware accelerators to be integrated in the processing unit. In both cases tools should be designed and adopted to easy the porting of AI solutions in embedded devices, enabling system designers to concentrate on the building blocks where they have the core IP.

The advantages of bringing AI-enhanced decision-making at the edge (edge-based AI) include the following [39]:

- Edge-based AI is highly responsive and closer to real-time than the typical centralised IoT model deployed to date. Insights are immediately delivered and processed, most likely within the same hardware or devices.

- Edge-based AI ensures greater security. Sending data back and forth with Internet-connected devices exposes data to tampering even without anyone being aware. Processing at the edge minimises this risk, and preserve privacy, with an additional plus: edge-based AI-powered devices can include enhanced security features.
- Edge-based AI is highly flexible. Smart devices support the development of industry-specific or location-specific requirements, from building energy management to medical monitoring to predictive maintenance.
- Edge-based AI does not require highly qualified personnel to operate. Since the devices can be self-contained, AI-based edge devices do not require data scientists or AI experts to maintain. Required insights are either automatically delivered where they are needed, or visible on the spot through highly graphical interfaces or dashboards.
- Edge-based AI provides for superior customer experiences. By enabling responsiveness through location-aware services, or rerouting travel plans in the event of delays, AI helps companies build trust and rapport with their customers.
- Edge-based AI seems to currently provide the required processing power and speed to execute tasks relating to data security or may identify points of altered data or intrusions (data smoothness check, data abnormality checks etc.).
- Edge-based AI reduces data transfers back and forth and thus supports lower bandwidth investments and backhaul costs.
- Edge-based AI supports more efficient and real-time decision making and requires reduced latency as ML triggers real-time actions and decision making. Having data closer to the decision-making point is an added value.

The intelligent edge allows humans to tackle multi-faceted processes by replacing the manual process of sorting and identifying complex data, key insights, and actionable plans. This can help humans gain a competitive edge by having better decision-making, improved return of investment (ROI), operational efficiency, and cost savings. At the same time, there are several challenges facing ML based edge computing such as [40]:

- The cost of deploying and managing an edge.
- The evaluation, deployment, and operation of edge computing solutions.
- Security challenges for processing at the edge and the fear of data breach.
- The type and number of operations performed on the data.

- The size of the data to be processed during the learning phase when the inference engine is generated.

The concept of incorporating AI into edge computing is evolving and more research is needed to get intelligent edge-based solutions to be fully set up, functional and running smoothly in production.

As smart devices require access to data for optimised performances and analysis accuracy, data access must be granted to a technology or solution provider, a new scenarios must be regulated to prevent restriction of data access to only a few cloud service providers able to collect and restrict access to new data. Industry 4.0 and IIoT are examples where there will be an exponential use of AI solutions. The IoT paradigm has evolved towards intelligent IoT applications which exploit knowledge produced by IoT devices at the edge using AI techniques and methods. Knowledge sharing between IoT devices is a challenging issue and the way in which IoT devices effectively produce, cumulate, and share the self-taught knowledge with other IoT devices at the edge in the vicinity to form the distributed intelligence in the IoT ecosystem is a research priority for the future.

Advances in AI and ML combined with the sensing/actuating, processing, connectivity capabilities of the IoT devices is driving transformation across industries and workstreams, across various industrial vertical sectors.

AI and ML processing happened in the past mostly in the cloud. As the enabling technologies are evolving AI computing is increasingly moving onto the IoT devices themselves, reducing dependence upon the cloud. This goes together with the emergence of a new kind of architecture that supports AI at the edge as depicted in Figure 3.16.

The intelligent IoT devices at the edge generate information that can be processed locally or close to the source by AI techniques and methods and exchange the information among the edge computing based IoT platforms setting the stage for AI-enabled IoT applications and devices at the edge.

Edge computing is putting the data processing power to the edge by providing energy efficient computing processing with guaranteed performance for real-time operations to manage the data generated by IoT devices, while preserving end-to-end security and privacy. In this context, edge AI means that AI software algorithms are processed locally on the hardware of IoT devices and the algorithms are using data (sensor data or signals) that are created on the IoT device, the device using edge AI software can process data and take decisions independently without a centralised cloud.

Running the edge AI software on IoT devices, gateways and edge infrastructure is critical in the new IoT/IIoT system architecture. The software

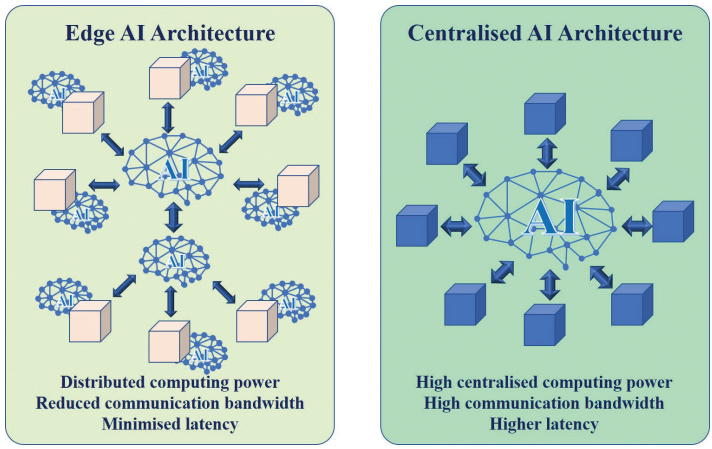

Figure 3.16 Edge AI vs centralised AI architecture (Source: STMicroelectronics).

works in real-time, reports conditions and behaviours on the device based on real-time sensor data. The AI based IoT devices at the edge integrate hardware and software for capturing sensor data, software for training the AI model for different application scenarios and binary software that runs the AI models and algorithms on the IoT devices and gateways. The use of virtualisation is extending the capabilities of edge computing and edge AI.

Edge computing is enabling edge AI as a distributed computing paradigm that brings computation and data storage closer to the location of the edge devices. Edge computing and edge AI are providing the processing of data at the edge allowing to deliver AI-enabled personalised features by deploying different distributed architectures with increased security mechanisms and features for different devices. Edge computing and edge AI create more distributed services at the network and device level implementing connectivity techniques with low-latency when exchanging information across networks and edge devices.

The use of edge computing and edge AI enables the load and processing balancing and increases the applications end to end resiliency when running on distributed systems architectures.

Combining the edge AI and cloud AI many applications in the different industrial sectors will operate in a hybrid manner with a part of AI processing on-device, at the edge and other at the cloud level. The type of AI processing needs determines the optimal federation between edge AI and cloud AI in any given use case depending on the application requirements.

The edge AI processing circuits are focusing on AI workloads that are deployed in edge environments, which include IoT/IIoT edge devices, gateways, edge micro servers and on-premise servers.

IoT continues to evolve and the development of affordable and accessible AI-enabled solutions across industries provide new concepts and integrated solutions combining IoT and AI.

In this context, new reference architectures, design languages, application generators, design automation and respective standardisation are obviously constituents of such engineerable new IoT, AI solutions. Further work is needed to address challenges such as:

- Device-centric AI that runs and trains DL models on edge devices and the use of AI in autonomous manufacturing processes and blockchain networks.
- Development of AI based processor chips for edge devices, PLCs, Distributed Computing Systems (DCSs), and integration of distributed intelligence for wireless and Ethernet-based connectivity on new processor platforms.
- Embedded advanced algorithms. AI based algorithms that can learn with less data and are applied to the edge devices and systems.
- The scale of the IIoT requires that DL networks capabilities are pushed out from the cloud to the edge, and into edge mobile and fixed devices. The edge devices in the industrial sector require intelligent control and coordination, based on DL and AI systems that are collecting data locally, process it and make decisions at the edge in real-time.
- Device-centric AI that runs and trains ML models on edge devices and the use of AI in autonomous manufacturing processes and blockchain networks.
- Security related involving the execution of cryptographic algorithms or other security related firmware at the edge.

The next generation IoT systems require distributed architectures connecting heterogeneous intelligent IoT devices having a context based dynamic behaviour. AI and ML combined with the sensing/actuating, processing, connectivity capabilities of the IoT devices affect the decisions on how to distribute the ML algorithms across device, edge and cloud to create a continuum of knowledge transfer and exchange across various applications domains.

Appling and implementing AI and ML techniques and methods at the edge at the device level requires the semiconductor technology providing the

following capabilities resources low energy consumption, high integration (size, weight, volume), real-time operation, high processing power, integrated connectivity, tools to easy develop new algorithms features, security features, and cost.

The energy efficiency and battery lifetime (e.g. low-power design requirements) for mobile and IoT devices at the edge require implementing AI methods and techniques that take into consideration the duty cycle between the different power modes (active, standby, off) and the energy consumption of the sensors/actuators, CPU and communication interface, while optimising data processing/exchange to the edge/cloud that affect the power consumption of the communication interface in relation to processing the data locally.

The implementation of AI and ML at the edge is highly dependent on the real-time requirements as the IoT systems demand real-time or near real-time responses of 1–2 seconds delay that can be associated with sending the data to the cloud and waiting for the response, while other demand real-time responses of milliseconds or below a millisecond that can be implemented only with edge computing processing.

In many cases loading the communication channels by sending gigabits of data to the cloud, in real-time as in the case of autonomous vehicles applications is not feasible due to the cost, latency and due to the fact that such loads can block all other traffic. Furthermore, the cloud solutions are very costly energy hungry.

One challenge for developing AI and ML for IoT devices at the edge is addressing the multiple platforms and in multiple languages as the AI and ML algorithms are implemented most likely of heterogenous environments with different operating systems, a different toolchain and a different software language. Splitting an AI and ML algorithm and migrating it from one environment to another is a resource-intensive process, which could affect the performance of the AI and ML algorithm, the performance of the IoT application as well as introduce errors and bugs.

Updating and maintaining the AI and ML algorithms on the IoT devices deployed in the field is critical to keep the functionality and the security of the IoT systems. The distribution of configuration and updates to IoT devices need new solutions to orchestrate the operation of semi-autonomous and autonomous IoT nodes to provide higher level functions. To achieve the full functionality expected of an IoT system, research should be done in advanced network reorganisation and dynamic function reassignment.

Considering that the AI and ML algorithms on the IoT devices are dynamic and require to evolve based on learning, all updates,

debug/troubleshoot need to be considered in real-time. The IoT edge devices will experience continuously new scenarios where the AI and ML algorithms will need to be adapted, based on failure or the need to improve their performance. Identifying the issues and improving accuracy, requires parallel simulation and modelling of the raw data considering different scenarios.

AI and ML implementation at the edge for different heterogenous IoT devices requires flexibility, agility and scaling at the device, network, edge processing and application levels. In this context, considering that the edge resources are agile and support flexible architecture that make use of the intelligent connectivity infrastructure capabilities and cloud environments, performance and memory can scale up and down depending on context and analytics requirements which allows data to improve and upgrade the ML algorithms or transfer the processing across this continuum.

The edge processing and transfer of AI and ML at the device level can improve privacy, ethical and security concerns as data or commercially sensitive data is restricted to the local use. Data collected locally can be processed locally and only information filtered by the user can be sent to the cloud. When large amount of sensitive data needs to be processed the edge, this can go together with the capability to transfer only filtered results to the cloud.

Developing IoT/IIoT applications integrating AI end-to-end solutions require addressing the components illustrated in Figure 3.17.

The IoT/IIoT devices and other data generators represent the structured and unstructured data sources that provide the raw input that is fed into a data conditioning step in which they are fused, aggregated, structured, accumulated, and converted to information. The information is processed by different AI algorithms. The components are implemented in different layers of the 3D IoT architecture. The computing capabilities required for performing AI functions are integrated with the other IoT/IIoT functions and provide acceleration mechanisms for processing the AI algorithms at the edge or in the cloud. To support robust AI implementations of IoT/IIoT applications elements like training/learning, metrics, bias assessment, verification, validation, security, safety, policy rules and ethics must be considered when designing complete solutions.

3.6.1 Training/Learning – Federated Learning

The development of AI technologies has to address the continuum between deep edge, edge, cloud and data centres, with the AI technology stack addressing the infrastructure and developer environment to cover the

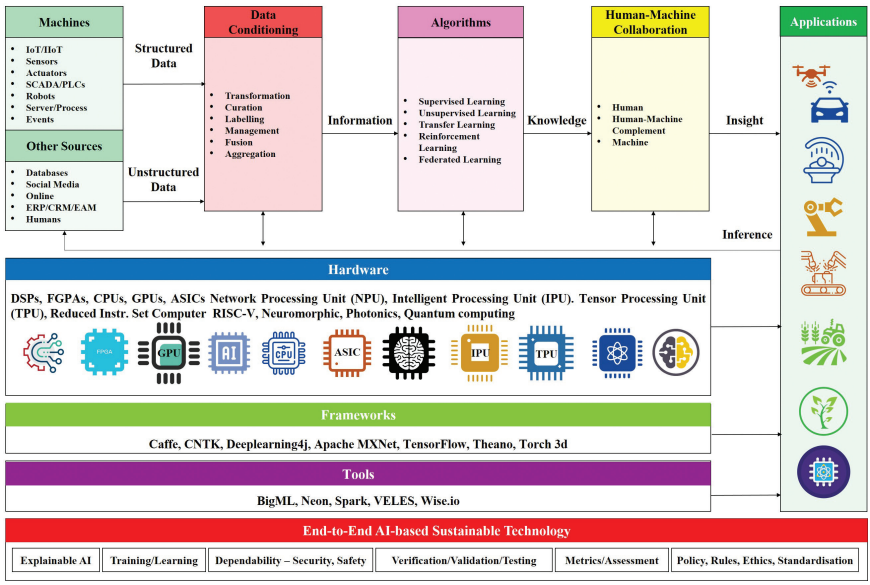

Figure 3.17 Edge AI ecosystem.

hardware, interfaces, platforms, training/learning, applications and services. The intelligent infrastructure at the edge refers to the tools, platforms, and techniques used to run store data, build, and train AI/ ML algorithms, and the algorithms themselves. This infrastructure already exists and is run today by large corporations and AI cloud service providers.

The developer edge environment is a key element for the transition of AI processing to the edge and refers to the tools that assist in developing code to bring out AI capabilities. The developer edge environment must cover all the layers of the AI technology stack offering end-to-end (E2E) development solutions.

Training is a key part of AI technology stack. AI training is defined as the process of creating ML algorithms and involves the use of a learning framework (e.g., TensorFlow) and training datasets provided by static or real-time sources (e.g. databases, IoT/IIoT edge devices, etc.). The data collected is the source of the training data that can be used to train AI-based models for (e.g. ML, DL, etc.) a different use cases, from pattern recognition, object detection, failure detection to consumer intelligence.

AI training systems need to store large volumes of data as the systems refine the algorithms. AI inference systems store only input data that could be useful in future training.

The AI training requires hardware platforms optimised for the type of AI neural networks and algorithms for efficient training used to address how the processing of neural networks is being performed on the platforms, and how application-specific accelerators are designed for specific neural networks for optimising the throughput and energy efficiency.

The emergence of new AI technologies has brought several problems, especially regarding communication efficiency, security threats and privacy violations. To this end, Federated Learning (FL) has received widespread attention. In contrast to centralised training, Federated Learning is a ML setting where the goal is to train a high-quality centralised model while training data remains distributed over a large number of clients each with unreliable and relatively slow network connections that facilitate a distributed learning process [151]. FL allows a ML to be synchronously dispatched to distributed data source locations to be locally trained. The resulting updated ML models are subsequently aggregated at a central location *i.e.,* the trained model parameters are transferred instead of the data, delivering a new updated global model ready for subsequent dispatching cycles. As a result, sensitive data are not exposed to the entity that maintains the global ML model. In certain cases, this approach further yields network resource savings, where training data transfer volume exceeds that of the ML model e.g., video stream data.

ML frameworks such as TensorFlow and PyTorch have already taken steps recently towards privacy with solutions that incorporate federated learning. Instead of gathering data in the cloud from users to train data sets, federated learning trains AI models on mobile devices in large batches, then transfers those learnings back to a global model without the need for data to leave the device. With open source initiatives such as OpenMined [152], AI models can be governed by multiple owners and trained securely on an unseen, distributed dataset.

In the IoT ecosystem, federated learning is proposed to train a globally shared model by exploiting a massive amount of user-generated data samples on IoT devices. Although this is very promising approach the heterogeneities of IoT device as well as the complexity of IoT environments pose great challenges to traditional federated learning. Potential concerns such as man-in-the-middle attacks, model poisoning, bandwidth and processing limitation need to be carefully addressed. Enabling technologies such as 5G and DLTs will play an important role towards this direction.

Cross-silo applications have also been proposed or described in several domains such as finance, health, and smart manufacturing. Cross-silo setting

can be relevant where several companies or organisations share incentive to train a model based on all their data but cannot share their data directly. This could be due to constraints imposed by confidentiality or due to legal constraints [153].

3.7 Distributed Ledger Technologies (DLTs)

A Distributed Ledger is a record of transactions or data that is maintained in a decentralised form across different systems, locations, organisations, or devices. It allows data (e.g. funds) to be effectively sent between parties in the form of peer-to-peer transfers without relying on any centralised authority to broker the transfer. A distributed consensus mechanism allows members of the network (nodes) to establish "trust" and thus maintain a common "distributed trust machine". Adding trust in IoT improves the system capabilities and enhance the IoT platforms capabilities [60, 62, 63].

The blockchain refers to Distributed Ledger Technology (DLT) solution where data from different transactions is linked, hashed, and organised per unit, one block at a time and each block is cryptographically "sealed". The unique seal is the start of the next block of transactions that creates the blockchain structure. Examples of DLTs that are classified as blockchains' applications are Bitcoin, Ethereum, Neo, Stellar, Hyperledger. In this context, the blockchain is the mechanism that allows the implementations (e.g. Bitcoin, Ethereum, Neo, Stellar, Hyperledger, etc.) to work, and the implementations are applications that uses blockchain. The main characteristics of blockchains are the decentralised architecture, "trust less" system properties (e.g. the fact that the system can operate without the need for participants involved to know or trust each other or a third party), the existence of consensus mechanisms, the maintenance of history of transactions and the insurance of immutability.

A blockchain is constituted of a digital DLT that is immutable, noneditable and shared among all participants in a blockchain network. The blockchain is constructed as a data structure constituted of time-stamped and cryptographically linked blocks. In this context, the individual blocks have a cryptographic hash, a list of validated transactions, and a reference to the previous block's hash, Using this mechanism the nodes can verify that a participant owns an asset without the need for a central governing authority.

The blockchain allows for participants to engage in trust less peer-to-peer transactions and the decentralised, trust less transactions are the key innovation of the blockchain. The following element are part of the common

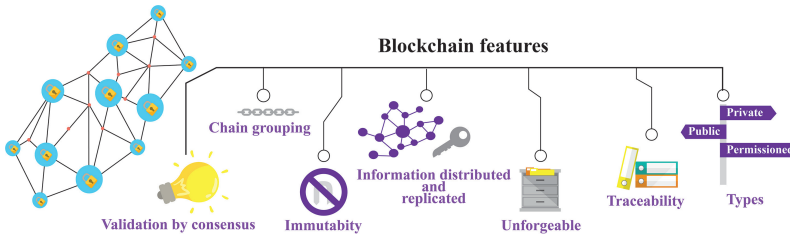

Figure 3.18 Blockchain features.

vocabulary of the blockchains applications and are used to describe the different functions performed by the blockchain.

- Consensus refers to consensus algorithms that represent the mechanism by which all nodes in the network agree on the same version of the truth. Consensus algorithms permit nodes on the system to trust that a specific part of data is valid and that it has been synchronised with all other nodes.
- Cryptographic keys refer to the use of symmetric keys and asymmetric (public-private) key pairs for the use of signing and verifying transactions.
- Decentralised Application (DAPP) refers to decentralised applications that are built on top of a blockchain-based system.
- Ledger represents a shared and distributed history of all transactions and balances.
- Merkle Tree Root (also called binary hash tree) is the result of all leaves hashed together to a single hash.
- Mining/Miners is the process of generating a new legitimate block by applying proof-of-work (e.g. the case of Bitcoin). In specific applications the nodes are dedicated to "mine" new blocks and these nodes are defined as "miners".
- Nodes are represented by any computer or device (e.g. IoT device) connected to a blockchain network.
- Secure Cryptographic Hash Function is defined as a secure cryptographic hash function that preserves one-way, easy computation, and makes impossible to reverse engineer.

The blockchains are classified in the following different categories [67] as presented in the following:

- **Blockchain 1.0** is represented by using blockchain in digital currency (e.g. Bitcoin and other cryptocurrencies payment systems) applications for the decentralisation of money or payment systems.

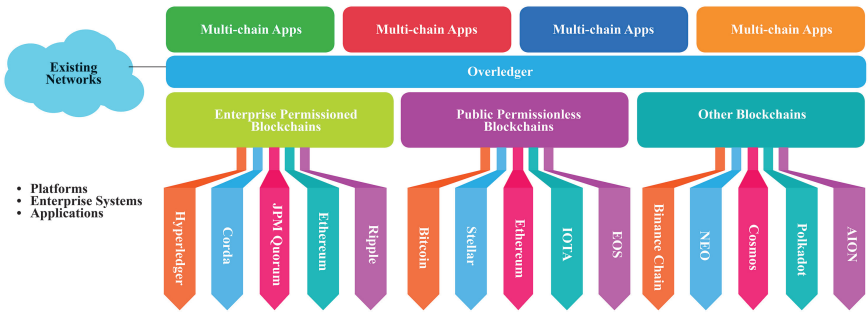

Figure 3.19 Overleger enterprise operating system [69].

- **Blockchain 2.0** is represented by the technology known as contracts, which goes beyond peer-to-peer payment systems and includes the transfers of other property such as stocks, bonds, smart property, and smart contracts.
- **Blockchain 3.0 is** found in the other applications beyond currency and markets, including the use of blockchain in areas like healthcare, governments, and commercial settings.

In the last years, the blockchain developments have addressed the creation of blockchain operating system (OS) that inter-connects blockchains and existing enterprise platforms, applications and networks to blockchain and facilitates the formation of internet scale multi-chain applications known as mApps [69]. Overledger provides interoperability with the full range of DLT technologies including Enterprise Permissioned blockchains like Hyperledger, R3's Corda, JP Morgan's Quorum, permissioned variants of Ethereum and Ripple (XRPL) and Public Permissionless blockchains / DAGs such as Bitcoin, Stellar, Ethereum, IOTA, EOS and the most recent blockchain like Binance Chain. The overledger is positioned as an enterprise operating system that interconnect enterprise platforms and networks providing interoperability to connect to any networks and hyper decentralised applications that run and store data on multiple blockchains is illustrated in Figure 3.19.

Blockchain interoperability is defined as "a composition of distinguishable blockchain systems, each representing a unique distributed data ledger, where atomic transaction execution may span multiple heterogeneous blockchain systems, and where data recorded in one blockchain are reachable, verifiable, and referable by another possibly foreign transaction in a semantically compatible manner" [70].

The emergence of several blockchain platforms brings challenges for IoT applications in particular with respect to the solutions and the interoperability

features provided considering the multitude of IoT platforms and blockchain platforms custom-made for specific purposes, (e.g. public, private, or consortium), that adds overhead to manage workflows. Public and private blockchains scalability is a challenge and new techniques for implicit consensus and data sharing (e.g. cross-blockchain transaction routing and retrieval, and asset referencing/discovery) are needed to provide improvements in transaction throughput and storage. Another interoperability issue is related to the fact that different blockchains have different architectures, use different protocols, service discovery, access control, and use different transaction mechanisms. For the integration of IoT and blockchains interoperability, security, privacy, and scalability remain priority research challenges in the near future for the emergence of efficient IoT blockchain integration.

The usage of ledgers (for record-keeping) has been applied in the past but have been recently enhanced to distributed forms, thus significantly improving their capabilities based on easy to understand architectures.

DLTs are intertwined with IoT platforms and used to provide efficient data management in terms of security, privacy, and safety [66].

There are several benefits from using DLTs in IoT applications, as summarised below:

- Increased security: Avoiding centralised networks, increased autonomy of devices, immutability over injection or penetration vulnerabilities, advanced cryptography, prevention of data augmentation and overall increased trust on the IoT nodes.
- Improved secure data management and reliability: devices communicating with each other directly, consensus on data propagation, reduced junk data processing, reduction of costs associated to centralised architectures setup and maintenance.
- Lower bandwidth: reduction of connectivity bottlenecks, cost efficiency, improved latency for decisions.
- Increased auditability: Through ledger immutability and auditing features.
- New types of contracts: ability to automate decision making once particular conditions are met (smart contracts).

DLTs have still to demonstrate their high value and applicability in modern IoT systems being supported by PKI (public key infrastructure) as well as Physical Unclonable Function (PUF) technologies. Combinations of several of these technologies have strong potentials of increasing even further the safety, security, and privacy capabilities of DLTs. However, this combination may in most cases require support from the hardware devices, especially when providing more processing power is required at the edge.

3.8 Intelligent Connectivity

Next generation IoT will make use of an intelligent connectivity that rely on the consolidation of communication technologies. Whereas the Internet is today the largest world-wide communication network, Intelligent connectivity for IoT still requires further development to allow high demand in bandwidth and quality in signal in support of more important content and data exchange. The requirements of real-time response may be, amongst other, safety and mission critical (like in telemedicine for example) and, consequently, IoT communications solutions are numerous and diverse. An important characteristic of the next generation IoT is the requirements range from high reliability and resilience in the communication network to ultra-low latency and high capacity at the communication channel. IoT intelligent connectivity solutions are also very dependent on the context in which they are applied and whether it is necessary to respond to strict energy efficiency constraints or cover large outdoor areas, deep indoor environments or vehicles moving at high speeds. As they are highly dependent upon both the creation of new technologies and the deployment of new communication networks (e.g. 5G mobile network), many of the next generation IoT capabilities may not be largely available before 2025.

3.8.1 Wireless and Cellular Communication Protocols Used for IoT Applications

The next generation IoT will rely on new communication network technologies. An important development is 5G because of its wider broadband capacity, and other technological enhancements to the 5G network that will allow the new connectivity to be a catalyst for the next-generation IoT services through advanced modulation schemes for wireless access, network slicing capabilities, automated network application lifecycle management, software-defined networking and network function virtualisation. In addition, support will be provided for edge- and cloud-optimised distributed network applications.

Next generation IoT builds on the technologies that have been developed and deployed over time (as summarised in Figure 3.20) providing constant evolution of the available data rates over time.

Similarly, next generation IoT takes advantage of the evolution on the spectrum availability and potential usage as summarised in Figure 3.21.

The next generation IoT must address the convergence between different cells and radiation and develop new management models to control roaming

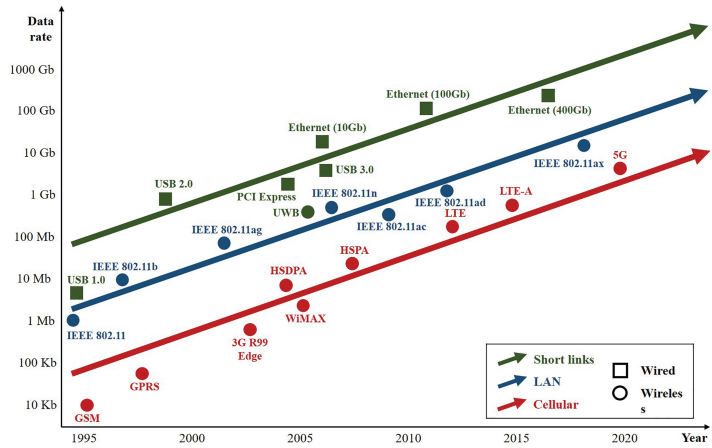

Figure 3.20 Evolution communication technologies (Source: IEEE).

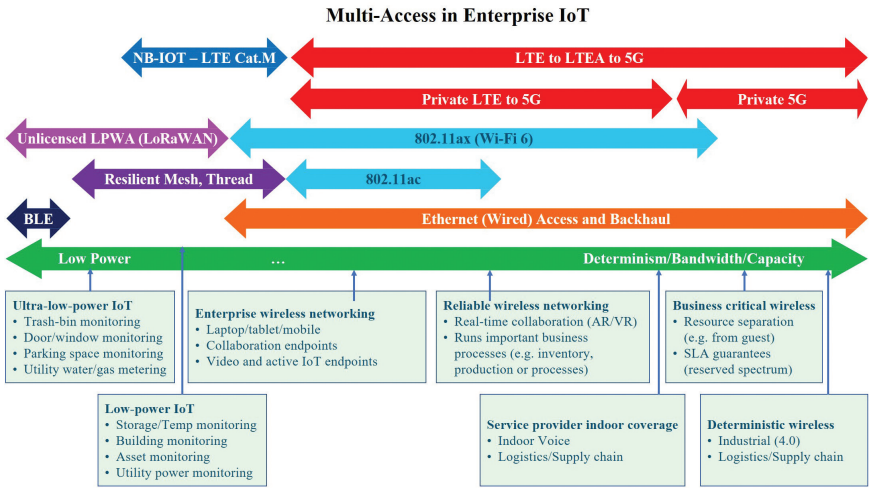

Figure 3.21 Spectrum use by IoT communication protocols. Adapted from GSMA.

while exploiting the coexistence of different cells and radio access technologies. New management protocols to control user assignment regarding cells and technology will have to be deployed in the mobile core network to access network resources more efficiently. Satellite communications must be considered a potential method of radio access, especially in remote areas. With the emergence of safety applications, minimising latency and various protocol translations that will benefit end-to-end latency. It is part of this

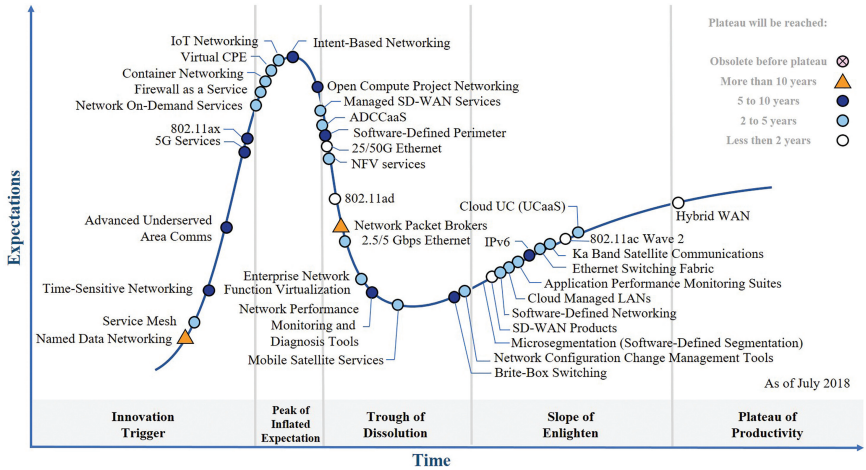

Figure 3.22 Enterprise networking and communications (Source: Gartner Hype Cycle).

latency where tactile technologies as part of the tactile Internet and the capacity to exchange data in real-time where the next generation IoT are involved.

Intelligent connectivity covers the networks and communications technologies as part of the IoT Network Communication Layer covering various communication protocols to provide seamless connectivity for heterogenous IoT devices with different levels of intelligence and connectivity needs.

Intelligent connectivity creates a scalable mobile platform and at the same demands a modular integration of technology (modems for 2G/3G/4G LTE), enabling high-speed data and voice, the implementation of various onboard protocols (i.e. LoRa, Sigfox, On Ramp Wireless, NWave/Weightless SIG, 802.11 Wi-Fi/Wi-Fi Aware, Bluetooth, ZigBee, 6LowPAN, Z-Wave, EnOcean, Thread, wMBus) and the simultaneous use of multiple ISM radio bands (i.e. 169/433/868/902 MHz, 2.4 GHz and 5 GHz).

Connectivity modules are based on integrated circuits, reference designs and feature-rich software stacks created for flexibility and modularity so that they could be implemented in various application domains.

As a result of the progress in data rates, the evolution of the available spectrum and, most importantly, the very large and diverse work in standardisation, the IoT protocols landscape is very large, with varying degrees of development and maturation as shown in Figures 3.22 and 3.23.

Communication technologies used by different IoT devices have various power consumption, bandwidth used, range, data rate, frequency

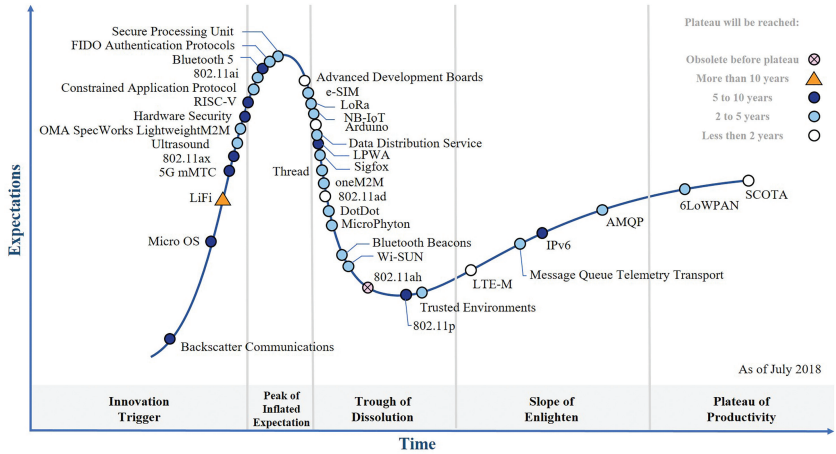

Figure 3.23 IoT standards and protocols (Source: Gartner Hype Cycle).

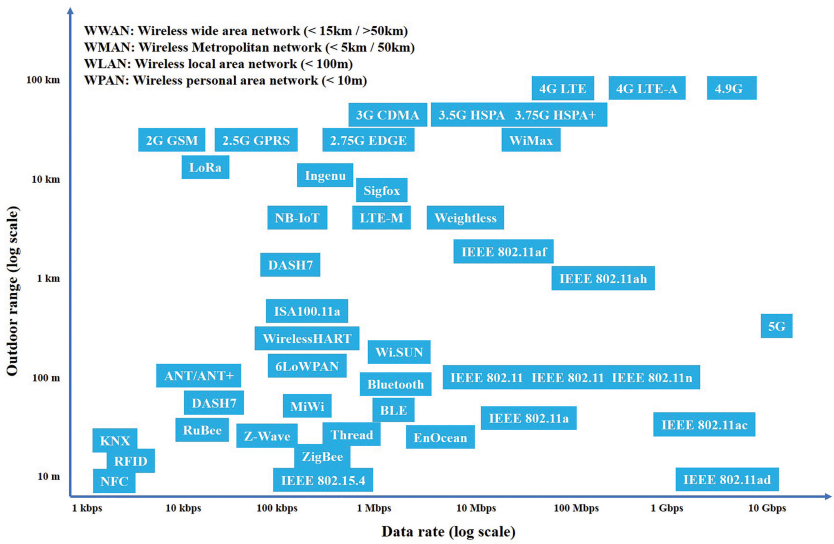

Figure 3.24 IoT wireless technologies landscape.

requirements to ensure seamless communication with other IoT devices and with infrastructure in the operating environments to connect to Anything (any device), to transfer information from/to Anyone (anybody), located Any place (anywhere), at Any time (any context), using the most appropriate physical path from Any path (any network) available between the sender and the

recipient based on performance and/or economic considerations, to provide Any service (any business).

The protocol landscape has evolved and has been structured according to two complementary dimensions. The dimensions in the organisation of the protocol landscape regards the available data rate and the range as shown in Figure 3.24.

The following sub-sections will detail the available protocols with a specific sub-section dedicated to the application protocols.

3.8.2 Wireless Personal Area Network (WPAN)

The development of IoT and mobile devices support various connectivity protocols for different applications and industrial sectors. Different types of wireless area networks are defined to cover the various requirements such as range, data rates, power consumption, and applications.

The wireless personal area network (WPAN) is defined as a personal, short distance area wireless network for interconnecting devices e.g. IoT devices, mobile/cellular phones, tablets, PCs, wireless wearable devices, pagers consumer electronics, PCs, etc.) centered around an individual person's workspace.

The IEEE 802 LAN/MAN Standards Committee develops and maintains networking standards and recommended practices for local, metropolitan, and other area networks, using an open and accredited process, and advocates them on a global basis. The standards covered are used for Ethernet, Bridging and Virtual Bridged LANs Wireless LAN, Wireless PAN, Wireless MAN, Wireless Coexistence, Media Independent Handover Services, and Wireless RAN. IEEE created an individual Working Group that provides the focus for each area.

The WPAN standard protocols for IoT devices are covered under IEEE P802.15 Wireless Personal Area Network (WPAN) Working Group that has 10 major areas of development: IEEE 802.15.1: WPAN / Bluetooth, IEEE 802.15.2: Coexistence, IEEE 802.15.3: High Rate WPAN, IEEE 802.15.4: Low Rate WPAN, IEEE 802.15.5: Mesh Networking, IEEE 802.15.6: Body Area Networks, IEEE 802.15.7: Visible Light Communication, IEEE P802.15.8: Peer Aware Communications, IEEE P802.15.9: Key Management Protocol and IEEE P802.15.10: Layer 2 Routing.

The discussions in this section will focus mostly on the IEEE 802.15.1 and IEEE 802.15.4. IEEE 802.15.1 will be further analysed for the Bluetooth and Bluetooth Low Energy. The IEEE 802.15.4 standard defines the functions

of the Physical and Media Access Control (MAC) layers and is the foundation for different protocol stacks, (e.g. Zigbee, Zigbee RF4CE, Zigbee Pro, WirelessHART, ISA 100.11a, etc.). In the 802.15.4 network there are two types of devices, one that is the full-function device (FFD) implementing all of the functions of the communication stack, which allows it to communicate with any other device in the network and the other that is the reduced-function device (RFD), with very reduce resource and communication capabilities. FFD can relay messages, and it is dubbed as a personal area network (PAN) coordinator, which oversees its network domain by allocating the local addresses and acting as a gateway to other domains or networks. RFDs can only communicate with FFDs and they cannot act as PAN coordinators as their rationale is to be embedded into the "things".

Different network topologies (e.g. star, mesh, or cluster tree topology) formed from clusters of devices separated by suitable distances can be built and every network needs at least a single FFD to act as the PAN coordinator. For example, the star topology is represented by a hub-and-spoke model where all devices communicate through a single central controller, namely, the PAN coordinator. The PAN coordinator (typically main powered) is represented by the hub, and all other devices (battery operated) form spokes that connect only to the hub. Several IoT applications use this type of network configuration (e.g. home automation, personal health monitors, wearables, etc.) where each star network selects a PAN identifier, that is in use by any other network within the radio range, allowing each star network to operate independently of other networks.

3.8.2.1 6LoWPAN

6LoWPAN combines the Internet Protocol (IPv6) and Low-power Wireless Personal Area Networks (LoWPAN) applied to devices by providing encapsulation and header compression mechanisms that allow IPv6 packets to be sent and received over IEEE 802.15.4 based networks. IoT on the Internet layer is the adaptation of the layer's functions to Link layer technologies with restricted frame size.

The base maximum frame size for 802.15.4 is 127 bytes, out of which 25 bytes need to be reserved for the frame header and another 21 bytes for link layer security. IEEE 802.15.4 g increased the maximum frame size to 2047 bytes, that make possible to compress IPv6 packet headers over the Link layer.

IETF defined in RFC6282 IPv6 over low-power wireless personal area networks, and 6LowPAN provides three main functions, IPv6 header compression, IPv6 packet segmentation and reassembly, and layer 2 forwarding. 6LowPAN, allows to compress the IPv6 header into 2 bytes, as most of the information is already encoded into the Link layer header.

6LowPAN uses a mesh topology, operates in the 2.4 GHz frequency band providing a data rate of 250 kbps with a coverage range of 100 m. The security implemented is AES-128 link layer security defined in IEEE 802.15.4 protocol providing link authentication and encryption. Security features defined in RFC 5246 standard are enabled by the transport layer security mechanisms over TCP. The RFC 6347 standard defines the transport layer security mechanisms over UDP.

3.8.2.2 ANT/ANT+

ANT is an ultra-low power wireless protocol designed for low data rate sensor network topologies (e.g. peer-to-peer, star, mesh, broadcast, ANT-FS, shared cluster) in personal area networks and for local area networks using Gaussian frequency-shift keying (GFSK) modulation. The IoT applications include sports, fitness, wellness, home health, homes, and industrial automation applications.

The protocol uses the 2.4GHz frequency and provides ultra-low power, network flexibility and scalability (e.g. self-adaptive and able to do practical mesh). ANT devices may use any RF frequency from 2400MHz to 2524MHz, except for 2457MHz, which is reserved for ANT+ devices. ANT devices may use the public network key, a private network key, or a privately-owned managed network key. The ANT+ network key is reserved only for ANT+ devices. The ANT protocol provides device profiles that are tied to a specific use case. The device profiles are shared among all the ANT+ adopters, enabling any ANT+ adopter to create a specific device for the specific use case (e.g. heart monitor) that will operate interchangeably with one another. ANT+ branded on a device assure the interoperability with other ANT+ branded devices [104, 106].

Each ANT node can operate as a slave or master and can transmit and receive as well as function as a repeater. ANT uses a very short duty-cycle technique and deep-sleep modes to ensure very low power consumption and a single 1-MHz channel for multiple nodes due to a time-division-multiplex technique. Each node transmits in its own time slot, with basic message length of 150 μs, and message rate ranging from 0.5 Hz to 200 Hz with an 8-byte payload per message. ANT ensures management of physical,

data link, network, and transport layers of OSI stack and ANT+ manages session, presentation, and application layers to provide data and devices interoperability.

ANT devices that communicate with each other are part of the same network and the ANT channel (e.g. independent, shared and scan channel) between two devices uses the same frequency, message period, device type and transmission type (i.e. slave or master). The ANT devices are establishing a channel by pairing for communication, process in which an ANT slave device gets the complete unique channel ID of the master, that plans to communicate with, and agrees upon same frequency and message period. Establishing a channel can be permanent, semi-permanent or transitory. The ANT nodes representing a wireless sensor device, include an ANT protocol engine, a Micro Controller Unit (MCU) and can be configured as master and slave nodes to participate in one or more networks [104, 106, 107].

The range provides is approximate 30 m and the data rates are 12.8 Kbit/s – 60kbit/s. ANT supports a 8-byte (64-bit) network key and 128-bit AES encryption for ANT master and slave channels. Authentication and encryption can be further implemented through the application level.

3.8.2.3 Bluetooth and Bluetooth Low Energy WPAN – IEEE 802.15.1

Bluetooth was designed for point-to-point or point-to-multipoint, star (up to seven slave nodes) network configurations and data exchange among mobile devices. Bluetooth is designed as an open wireless protocol used in the unlicensed Industrial, Scientific, and Medical (ISM) 2.4 to 2.483 GHz short-range radio frequency bandwidth. The protocol is utilised for exchanging data over short distances from fixed and mobile devices, creating personal area networks (PANs). Bluetooth uses frequency-hopping spread spectrum, which chops up the data being sent and transmits chunks of it on up to 79 frequencies. The Bluetooth modulation used in the basic mode is GFSK. Bluetooth devices must operate in one of four available modes from mode 1 (insecure mode) to mode 4 -where security procedures are initiated after link setup. Secure Simple Pairing uses Elliptic Curve Diffie Hellman (ECDH) techniques for key exchange and link key generation in mode 4.

Bluetooth Low Energy (BLE) was optimised for power consumption to address small-scale consumer IoT applications. BLE is integrated into several IoT devices such as fitness and medical wearables (e.g. smart-watches, glucose meters, pulse oximeters, etc.), smart home devices (e.g. door locks) and IoT tracking devices (e.g. objects, animals, etc.) whereby data

is conveniently communicated to and visualised on smartphones. Bluetooth Mesh specification aims to enable a scalable deployment of BLE devices (e.g. retail contexts. logistics, etc.). BLE is providing versatile indoor localisation features and IoT beacon networks are used for different IoT service applications (e.g. in-store navigation, personalised promotions, content delivery, etc.). BLE is incompatible/non-interoperable with Bluetooth and to achieve interoperability a dual-mode device needs to be implemented.

BLE uses different set of technical and radio techniques to ensure very low power consumption implementing the data protocol to create low-duty-cycle transmissions or a very short transmission burst between long periods, combined with very low-power sleep modes.

BLE uses a specific frequency-hopping spread-spectrum (FHSS) scheme that employs forty 2 MHz-wide channels to ensure reliability over longer distances. Bluetooth offers data rates of 1, 2, or 3 Mbits/s, while BLE's rate is 1 Mbit/s with a net throughput of 260 kbits/s. BLE has a 0 dBm (1 mW) power output, latency of 6 ms, uses 128-bit AES security and provides a range of 50 meters. The adaptive frequency-hopping technique to avoid interference, a 24-bit cyclic redundancy check (CRC), and a 32-bit message integrity check improves BLE link reliability. In BLE (e.g. version 5.0), the new waveforms and coding techniques are implemented in order to achieve longer ranges, less power consumption and latency, better robustness and support for higher number of subscribers in a single Bluetooth network.

3.8.2.4 UWB

UWB is using wireless connectivity at wide bandwidth for short-range applications. UWB technology transmits information spread over a large bandwidth (more than 25% of the centre frequency or at least 500 MHz) with very low power levels therefore not interfering with other narrower band devices nearby. The receiver translates the pulses into data by listening for a familiar pulse sequence sent by the transmitter. As the data is moving on several channels at once, it can be sent at high speed, up to 1 gigabit per second. UWB technology can penetrate walls. Frequency regulations limit UWB to low power levels to keep interferences below the level of noise produced unintentionally by electronic devices (e.g. TV sets). UWB is limited to short-range applications, enabling wireless connectivity. IEEE 802.15.3 – standard addresses the high-data-rate WPAN designed to provide sufficient quality of service for the real-time distribution of multi-media. Apple launched the three phones with ultra-wideband capabilities in 2019 (e.g. iPhone 11, iPhone 11 Pro, iPhone 11 Pro Max).

3.8.2.5 EnOcean

EnOcean is a wireless technology based on the wireless EnOcean radio standard (ISO/IEC 14543-3-10/11) in sub 1GHz is optimised for use in buildings, with a radio range of 30m indoors and 300m outdoors. The standard covers the OSI (Open Systems Interconnection) layers 1-3 which are the physical, data link and networking layers. EnOcean wireless data packets are relatively small, with the packet being only 14 bytes long and are transmitted at 125 kbit/s. RF energy is only transmitted for the 1's of the binary data, reducing the amount of power required.

The transmission frequencies used for the devices are 902 MHz, 928.35 MHz, 868.3 MHz, and 315 MHz. EnOcean technology uses energy harvesting techniques to harvest the energy from mechanical motion, indoor light, temperature differences. Energy converters are used to transform energy fluctuations into usable electrical, electromagnetic, solar, and thermoelectric energy.

Th authentication method offers field-proven secure and reliable communication in building automation. The unique 32-bit identification number (ID) of the standard EnOcean modules cannot be changed or copied. For additional data security, the security mode protects battery less wireless communication with enhanced security measures to prevent replay or eavesdropping attacks and forging of messages. A specific feature is a maximum 24-bit rolling code (RC) incremented with each telegram which is used to calculate a maximum 32-bit cypher-based message authentication code (CMAC). The CMAC uses the AES 128 encryption algorithm. Another security mechanism is the encryption of data packets by the transmitter. The data is encrypted using the AES algorithm with a 128-bit key.

3.8.2.6 ISA100.11a

ISA100 is developed by ISA as an open-standard wireless networking technology based on IEEE 802.15.4 (MAC and physical layer), TDMA (Time Synchronized Mesh Protocol), utilising DSSS with channel/frequency hopping (with blacklists of noisy channels) and mesh routing. The official description is "Wireless Systems for Industrial Automation: Process Control and Related Applications".

The standard provides reliable and secure wireless operation for non-critical monitoring, alerting, supervisory control, open-loop control, and closed-loop control applications. ISA100.11a specification was approved as IEC 62734.

The standard adds features and functions that give support to comply with ETSI EN 300328 v1.8.1 (e.g. country codes to identify device location and the capacity to attenuate output power levels). The IEC 62734 standard incorporates Annex V, that presents multiple scenarios and approaches to assure ETSI compliance.

ISA100.11a uses mesh and star topologies, operates in the 2.4 GHz frequency band providing a data rate of 250 kbps with a coverage range of 100 m. The security implemented in ISA 100.11a standard is embedded with integrity checks and optional encryption at data link layer of the OSI reference model. Security mechanisms are provided in transport layer and 128 bits keys are used in both transport and data link layers. To join a ISA 100.11a network, a sensor node needs to use a shared global key, a private symmetric key or certificate.

3.8.2.7 NFC

The Near Field Communication (NFC) protocol operates at high frequency band at 13.56 MHz and supports data rate up to 424 kbps. The applicable range is up to 10 cm where communication between active readers and passive tags or two active readers can occur. NFC is a short-range technology and protocol that allows two devices to communicate when they are placed into the touching distance. NFC enables sharing power and data using magnetic field induction at 13.56MHz (HF band), at short range, supporting varying data rates from 106kbps, 212kbps to 424kbps.

A key feature of NFC is that it allows two devices to interconnect. In reader/writer mode, an NFC tag is a passive device that stores data that can be read by an NFC enabled device (e.g. smart poster, for which a technical specification was developed). NFC devices are designed to exchange data in Peer-to-Peer mode.

Bluetooth or Wi-Fi link set up parameters can be shared using NFC and data like digital photos, and virtual business cards can be transferred. The NFC device itself acts as an NFC tag in Card Emulation mode, resembling an external interrogator as a traditional contactless smart card. This facilitates contactless payments and e-ticketing. NFC standards are acknowledged by major standardisation bodies like ISO (e.g. ISO/IEC 18092).

Security is implemented in the NFC using mechanisms such as Digital Signature and Trusted Tag. The Digital Signature (defined in the NFC Forum Signature RTD 2.0) uses asymmetric key exchange and the Digital Signature is a part of the NFC Data Exchange Format (NDEF) message, which includes also a Certificate Chain and a Root Certificate. Each NFC device has a private

and a public key. The Trusted Tag method is fully compliant with NFC Forum Tag Type 4 and works with any NFC Forum compatible devices. The Trusted tag is protected from cloning and embedded with cryptographic code that is generated by every "tap" or click on NFC button. This cryptographic code protects the content of the transmitted information.

3.8.2.8 RuBee

RuBee is a peer to peer communication protocol designed for active or passive tags operating at low frequency (using magnetic induction), suitable in environments containing water and/or metal. IEEE is standardising it as P1902.1 "IEEE Standard for Long Wavelength Wireless Network Protocol". The standard offers a "real-time, tag searchable" protocol using IP and subnet addresses associated with asset taxonomies that operate at speeds of 300 to 9600 Baud. RuBee Visibility Networks are operated by Ethernet-enabled routers. RuBee enables tag networks and applications. RuBee tags and tag data may be seen as a stand-alone via web servers from anyplace in the world. RuBee tags that are properly programmed can be discovered and monitored over the Internet using search engines.

3.8.2.9 DASH7 (D7A)

The DASH7 protocol complies with the ISO/IEC 18000-7 standard. ISO/IEC 18000-7 is an open standard for the license-free 433 MHz ISM band air-interface for wireless communications. Using 433 MHz frequency provides D7A devices with long propagation distance and better penetration in objects (e.g. buildings).

The protocol defines four different device classes such as blinker that only transmits and does not use a receiver, endpoint that can transmit and receive the data and supports wake-up events, sub controller that is a full featured device using wake on scan cycles similar to end points and gateway that connects D7A network with the other networks and is always online and always listens unless it is transmitting.

DASH7 uses AES-CBC for authentication and AES-CCM for authentication and encryption.

3.8.2.10 RFID (WPAN – IEEE 802.15)

RFID standards include ISO 11784/11785, ISO 14223, ISO/IEC 14443, ISO/IEC 15693, ISO/IEC 18000, ISO/IEC 18092, ISO 18185, ISO/IEC 21481, ASTM D7434, ASTM D7435, ASTM D7580, ISO 28560-2.

Passive RFID tags have no power source/battery and the power is provided by the interrogator and operate in short read range (from mm to m). The passive RFID tags consist of an integrated circuit attached to an antenna and use either the magnetic field or the electric field to transmit their signals.

Semi passive tags are RFID devices which use a battery to maintain memory in the tag and sometimes provide power to the processing unit or power the electronics that enable the tag to modulate the reflected signal, or to support sensors. The radio aspects remain passive in that they only react to received signals and use the power from the carrier signal.

Active RFID systems usually operate at 433 MHz, 868MHz (no standardised RFID protocol) 2.45 GHz, or 5.8 GHz and have a read range of up to 100 meters. Low frequency (LF) and very high frequency (VHF) systems are also available. Some active tags now include GPS positioning capability.

Active RFID beacons are used in real-time locating systems (RTLS), where the precise location of an object is tracked. In an RTLS, a beacon emits a signal with its unique identifier at pre-set intervals. The beacon's signal is picked up by at least three interrogator antennas (triangulation or trilateration) positioned around the perimeter of the area where assets are being tracked. RTLS are usually used outside, in a container base or inside large facilities to track parts or personnel. Active tags can be read reliably because they broadcast a signal to the interrogator.

The implementation of security mechanisms in RFID technology is based on confidentiality, integrity, and availability. Confidentiality is the information protection from unauthorised access. Integrity is related to data protection from modification and deletion by unauthorised parties. Availability represents the capability for data access when needed.

An overview of the RFID technologies and the related standards is given in Table 3.3.

One functional area of great relevance to many supply chain applications is the ability to monitor environmental or status parameters using an RFID tag with built in sensor capabilities. Parameters of interest may include temperature, humidity, and shock, as well as security and tamper detection.

3.8.2.11 Thread

Thread is a mesh networking low-power wireless protocol, based on the IPv6, designed to address the interoperability, security, power, and architecture challenges of the IoT. Thread utilises 6LoWPAN that employs the IEEE 802.15.4 wireless protocol with mesh communication. Thread is IP-addressable, with cloud access and AES encryption. Thread 1.2 release

Table 3.3 RFID technologies and related standards

Frequency	LF Low Frequency	HF High Frequency	UHF Ultra-High Frequency	SHF Super High Frequency
Frequency range	30kHz to 300kHz	3MHz to 30MHz	300MHz to 3GHz	3GHz to 30GHz
RFID Frequency Air Interface	30-50kHz 125/134kHz[1] 131/450kHz	6.78MHz[2] 7.4-8.8MHz 13.56MHz 27MHz	433MHz 840-960MHz 2.45GHz	3.1-10,6GHz 5.8GHz 24.125GHz
Standard/Protocol Air Interface	ISO/IEC 18000-2 USID (Ultrasound) ISO 11784/5 ISO 14223 IEEE P1902.1/ RuBee EM4100, Sokymat UNIQUE	ISO/IEC 18000-3 ISO/IEC 15693 ISO/IEC 18092/NFC ISO/IEC 10536	ISO/IEC 18000-7 ISO/IEC 18000-6 Type A, B. C EPC C1G2 ISO/IEC 18000-4 IEEE 802.11 IEEE 802.15 WPAN IEEE 802.15 WPAN Low Rate IEEE 802.15 RFID	ISO/IEC 18000-5 (withdrawn) IEEE 802.15 WPAN UWB
Availability	$>$ 30 years RuBee -2 years	$>$ 10 years	US $>$ 9 years, EU $>$ 7 years, (4-channel plan is from 2008)	$>$ 10 years
Multiple tag reading	Limited	Up to 50 tags/sec	Up to 200 tags/sec	Up to 300 tags/sec
Reading distance	0.001–1 m	0.001–0.5m	US $\sim$ 0–6 m, EU $\sim$ 0–4 m Passive tags $\sim$ 0–100 m, Active tags	$\sim$ 0–1000 m, Active tags
Data transmission rate	Low	Medium	Fast	Very fast
Power Source	Passive – inductive coupling Active – using battery RuBee protocol considers both active and passive tags.	Passive tags, using inductive or capacitive coupling	Active tags with embedded battery Passive tags using capacitive, E-field coupling, Backscattering	Active tags with embedded battery Passive tags using capacitive, E-field coupling

[1,2]According to Annex 9 of the ERC Rec 70-03, inductive RFID Reader systems primarily operate either below 135 kHz or at 6.78 or 13.56 MHz. Therefore, the correlated transponder data return frequencies reside in the following ranges:
LF Range Transponder Frequencies: $f_C = <$ 135 kHz, f_{TRP} = 135 to 148.5 kHz
HF Range Transponder Frequencies: f_C = 6.78 MHz f_{TRP} = 4.78 to 8.78 MHz
f_C = 13.56 MHz f_{TRP} = 11.56 to 15.56 MHz.

includes commercial extensions, the integration of enterprise device lifecycle management and Bluetooth Low Energy Extensions. Enterprise-level security is handled using information technology authentication, authorisation, accounting, and the Thread 1.2 release expands the IP connectivity and end-to-end security for BLE devices. Applications using the protocol can consolidate multiple Thread networks into an extensive virtual Thread network with thousands of nodes, including predictable, stable addressing and management even as devices migrate within that virtual network.

Thread uses mesh network topology, in the 2.4GHz frequency spectrum, providing data rates of 250 kbps with a coverage range of 30 m. The security implemented uses a 128-bit AES encryption system. The encryption cannot be disabled.

Thread utilises a network-wide key that is used at the Media Access Layer (MAC) for encryption. This key is used for standard IEEE 802.15.4 authentication and encryption. IEEE 802.15.4 security protects the Thread network from over-the-air attacks originating from outside the network. Each node in the Thread network exchanges frame counters with its neighbours via an MLE handshake. The frame counters help protect against replay attacks. Thread allows the application to use any internet security protocol for end-to-end communication. The protocol can connect up to 250 devices in a wireless mesh network.

3.8.2.12 WirelessHART

WirelessHart is an open standard wireless networking technology developed by HART Communication Foundation. The protocol utilises a time synchronised, self-organising, and self-healing mesh architecture. WirelessHart currently supports service in the 2.4 GHz ISM Band using IEEE 802.15.4 standard radios. WirelessHART was developed as an interoperable wireless standard specifically for the requirements of Process field device networks. The HART Protocol uses Frequency Shift Keying (FSK) standard to superimpose digital communication signals at a low level on top of the 4-20mA.

WirelessHart uses mesh and star topologies, providing data rates of 250 kbps with a coverage range of 200 m. The security implemented uses a 128-bit AES encryption system. The encryption cannot be disabled.

The security manager in the WirelessHART gateway administers the Network ID, Join key and Session key. A common network key is shared among all devices on a network to facilitate broadcast activity in addition to

individual session keys. A 128-bit join encryption key is used to keep sent and received data private, during the joining process.

3.8.2.13 Z-Wave

Z-Wave is a mesh network low-energy wireless communications protocol used mainly for wireless control of home appliances, lighting control, security systems, swimming pools, garage door openers, thermostats, windows, locks, etc. Z-Wave systems can be controlled via the Internet and locally through devices or using a Z-Wave gateway or central control device serving as hub controller and portal. Z-Wave's interoperability at the application layer assures that Z-Wave devices share information and allows all Z-Wave hardware and software to work together.

Z-Wave uses the unlicensed industrial, scientific, and medical (ISM) band and operates at 868.42 MHz in Europe and 908.42 MHz in the US. Z-Wave provides data rates of 9600 bps and 40 kbps, with output power at 1 mW.

The Z-Wave range between two nodes is up to 100m in an outdoor, unobstructed setting. For in-home applications, the range is 30 m for no obstructions and 15 m with walls in between. Z-Wave Alliance requires the mandatory implementation of Security 2 (S2) framework on all devices receiving certification.

Z-Wave mesh networks become more reliable as more devices are added (e.g. a Z-Wave network with 50 devices is more reliable than a Z-Wave network of 25 devices).

Z-wave provides packet encryption, integrity protection and device authentication services. End-to-end security is provided on application level (communication using command classes). It has in-band network key exchange and AES symmetric block cipher algorithm using 128-bit key length.

3.8.2.14 ZigBee WPAN – IEEE 802.15.4

ZigBee is a short-range, low-power, wireless standard (IEEE 802.15.4), deployed in mesh topology to extend coverage by relaying IoT sensor data over multiple sensor nodes. ZigBee/ZigBee Pro is based on a specification that adds application profile, security, and network layers to IEEE 802.15.4 standard for wireless low-rate personal area networks. It operates in the UHF/microwave bandwidth with battery-powered tags that communicate with each other. Zigbee protocol features include support for multiple network topologies such as point-to-point, point-to-multipoint and mesh networks, low duty cycle that provides long battery life, low latency, direct

sequence spread spectrum (DSSS), 128-bit AES encryption for secure data connections, collision avoidance, retries and acknowledgements.

ZigBee protocols are designed to be used in embedded applications requiring low data rates and low power consumption, enabling devices to form a mesh network of up to 65 000 nodes, covering a very large area target general-purpose, inexpensive, self-organising, mesh networks which are deployed for building/home automation, industrial control, embedded sensing, medical data collection, smoke and intruder warning, domotics, etc. ZigBee networks use low power for communication, and individual devices can run for a few years on the installed battery. ZigBee is a perfect complement to Wi-Fi in many of these applications. ZigBee provides data rates at medium power-efficiency due to mesh configuration. The protocol is designed for physical short-range ($<$ 100m) mesh configurations best-suited for medium-range IoT applications with an even distribution of nodes in proximity.

The Zigbee PRO Specification adds child device management, improved security features, and new network topology options to Zigbee networks. Commissioning devices into networks has also been improved and made more consistent through Base Device Behaviour (BDB). The specification furthermore requires Green Power Basic Proxy functionality in all devices to further support Green Power capabilities and compiles all profile clusters into a single specification.

The Zigbee standard operates on the IEEE 802.15.4 physical radio specification and operates in unlicensed bands including 2.4 GHz, 900 MHz and 868 MHz Zigbee 3.0 supports the increasing scale and complexity of wireless networks, and copes with large local networks of greater than 250 nodes. The data rates provided are 250 kbps (2.4 GHz) 40kbps (915 MHz) 20kbps (868 MHz). Zigbee also handles the dynamic behaviour of these networks (with nodes appearing, disappearing, and re-appearing in the network) and allows orphaned nodes, which result from the loss of a parent, to re-join the network via a different parent.

The self-healing nature of Zigbee Mesh networks also allows nodes to drop out of the network without any disruption to internal routing. Zigbee's supports over-the-air (OTA) upgrade for software updates during device operation and provides enhanced network security using methods such as centralised security by employing a coordinator/trust centre that forms the network and manages the allocation of network and link security keys to joining nodes or distributed security where there is no coordinator/trust

centre. Any Zigbee router node can subsequently provide the network key to joining nodes.

ZigBee is a secure wireless communication protocol, with security architecture built in accordance with IEEE 802.15.4 standard. Security mechanisms include authentication – authorised access to network devices, integrity protection and encryption with key establishment and transportation.

Device authentication is the procedure of confirming a new device that joins the network as authentic. The new device must be able to receive a network key and set proper attributes within a given time frame to be considered authenticated. Device authentication is performed by the Trust Center.

Integrity protection is realised on the frame level using message integrity checks (MIC) to protect the transmitted frames and ensure they are not accessed and manipulated. A 128-bits symmetric-key cryptography is implemented in ZigBee's security architecture.

JupiterMesh is a robust, low-power industrial IoT wireless mesh network for Neighbourhood Area Network (NAN) with flexible data rates that enables neighbourhood and field area communications for utilities and municipalities deploying intelligent grid and smart city solutions. JupiterMesh is supported by ZigBee Alliance and is built on open IETF and IEEE standards uses parts of the IEEE 802.15.4g standard used by the Wi-SUN Alliance, as well as IEEE 802.15.4e and the IETF's Ipv6, 6LoWPAN, UDP, TCP, RPL, and CoAP protocols.

The protocol includes advanced technologies such as IPv6, frequency hopping, multi-band operation, authentication, encryption, and key management to drive industry realisation of interoperable multi-vendor implementations that scale and that are secure and easy to manage. JupiterMesh can operate in the sub-GHz ISM bands, as well as the 2.4GHz band, using FSK, O-QPSK, and OFDM modulation schemes.

3.8.3 Wireless Local Area Network (WLAN)

3.8.3.1 Wi-Fi WLAN – IEEE 802.11

IEEE 802.11 wireless technologies (Wi-Fi) address many of the requirements of IoT and have a number of challenges related to power consumption for end devices, due to the need for client devices to wake up at regular intervals to listen to AP announcements, waste cycle in contention processes, etc. and the frequency bands (e.g. 2.4–5 GHz), which are characterised by short transmission range and high degree of loss due to obstructions.

Wi-Fi provides high-throughput data transfer for both industrial and home environments. Wi-Fi has high energy requirements and is used for IoT applications and services that do not require large networks of battery-operated IoT sensors. For such applications Wi-Fi major limitations are in coverage, scalability, and power consumption. Many IoT devices that are connected to power outlet (e.g. smart home IoT monitoring devices and appliances, digital signages or security cameras, etc.) are using Wi-Fi as connectivity protocol.

Wi-Fi – specification (ISO/IEC 8802-11) is an international standard describing the characteristics of a wireless local area network (WLAN). The name Wi-Fi (short for "Wireless Fidelity") corresponds to the name of the certification given by the Wi-Fi Alliance, formerly WECA (Wireless Ethernet Compatibility Alliance), the group which ensures compatibility between hardware devices that use the 802.11 standards. Wi-Fi networks are networks that comply with the 802.11a-x specifications. The different frequency bands used by the different Wi-Fi implementations is presented in Figure 3.25 [91].

The new Wi-Fi generation is represented by the Wi-Fi 6 protocol that enhances network bandwidth (i.e. <9.6 Gbps) to improve data throughput per user in congested environments.

Wi-Fi 6E brings the technology into 6 GHz and features more contiguous spectrum, wider channels, less interface, gigabits speed, very low latency and high capacity [213]. The increased network capacity and the improved simultaneous communication between access points and multiple endpoints allows Wi-Fi 6 (IEEE 802.11ax) to offer improved performance in crowded areas. Several technologies facilitate these improvements and the most important are [214]; firstly building upon the Orthogonal Frequency Divisional Multiple Access (OFDMA) technology, already available in previous Wi-Fi standards, gives access points the ability to divide channels into many sub-channels, i.e. that the access points can communicate with multiple devices at the same time at a lower data rate; secondly, utilising the Multi-User Multiple-Input-and-Multiple-Output (MU-MIMO) which already are used by Wi-Fi 5 for downlinks.

Wi-Fi 6 utilises MU-MIMO for uplinks and enable access points to simultaneously receive communication from multiple clients. The combination of these technologies and some others (e.g. 1024-QAM) results in more efficient networks and reduces the overall network latency. The power management for connected IoT devices is also improved. Wi-Fi 6 supports a so called "target wake time" feature, which allow the Wi-Fi access points to put IoT devices in sleep mode for a given period [214]. The possibilities to reduce the power consumption of battery driven IoT devices can be very useful.

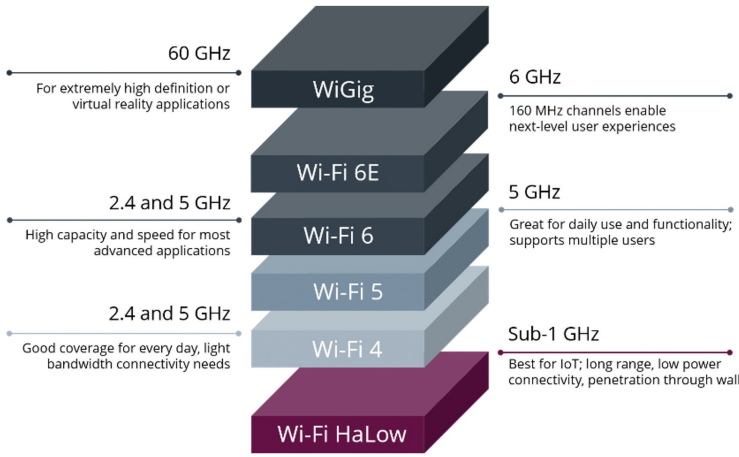

Figure 3.25 Wi-Fi frequency bands [91].

Today, Wi-Fi dominates as the infrastructure from small companies to large enterprises at the expense of the mobile network operators and their cellular network. The technologies are fundamentally different, primarily due to their differences in unlicensed and licenced environments. With the upcoming Wi-Fi 6 and 5G technologies, Wi-Fi and cellular becoming closer to each other, and Wi-Fi 6 is the first generation of WLAN technology that promises to seamlessly interact with the cellular solutions [215]. Wi-Fi 6 access points are already commercialised and are being built into networks [214]. Wi-Fi 6 phones are available together with Wi-Fi 6 adapters and routers. There are still early for private cellular IoT, but it is moving in that direction. Early findings indicate that unlicensed and licensed spectrum will continue to exist as complements [215], but it is a matter of quality of service (QoS) and costs for the users.

Wi-Fi 6 is a private-network technology with similar capabilities to 5G. The majority of IoT devices connect to Wi-Fi or an IoT-specific protocol designed for local ranges. Wi-Fi 6 lacks roaming capability, which has a significant impact on mobile sensor applications. Both 5G and Wi-Fi 6 are based on Orthogonal Frequency Division Multiple Access (OFDMA).

Wi-Fi 6 is less prone to interference than previous Wi-Fi standards, requires less power consumption and has improved spectral efficiency. 5G offers fully private 5G networks based on either licensed or unlicensed, shared spectrums and could offer alternatives to Wi-Fi 6. Specific IoT applications will require the capabilities offered by 5G and Wi-Fi 6 and the factors

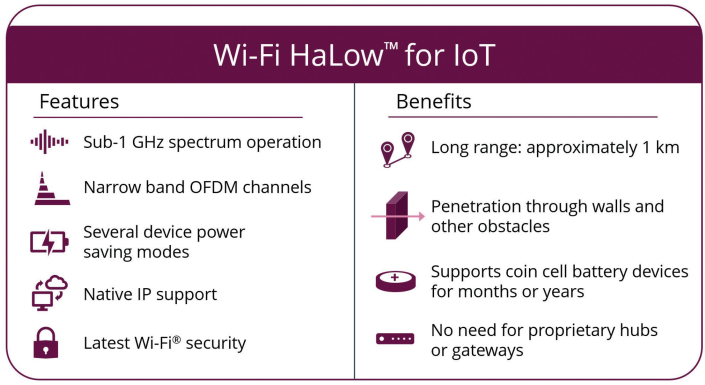

Figure 3.26 Wi-Fi HaLow features [91].

for deciding between the technologies are availability, range, the interplay of mobility and roaming capability, overall costs, etc.

3.8.3.2 Wi-Fi HaLOW

Wi-Fi HaLow wireless technology is based on the IEEE 802.11ah protocol standard. The technology augments Wi-Fi by operating in the spectrum below one gigahertz (GHz) to offer longer range and lower power connectivity. Wi-Fi HaLow meets the requirements for the IoT to enable a variety of use cases in industrial, energy (e.g. smart metering), smart home/building, agricultural, and smart city environments [91]. Some features and benefits are illustrated in Figure 3.26.

Wi-Fi HaLow utilises 900 MHz license-exempt bands to give extended range Wi-Fi networks, compared to Wi-Fi networks working in the 2.4 GHz and 5 GHz bands. The protocol has low energy consumption, allowing for IoT implementations with large groups of stations or sensors that cooperate to share data. The protocol's low power consumption competes with Bluetooth and has the benefit of higher data rates and wider coverage range.

3.8.3.3 White-Fi

White-Fi wireless technology is based on the IEEE 802.11af protocol standard that supports WLAN operation in the he VHF and UHF bands between 54 and 790 MHz TV white space spectrum, designed for ranges up to 1 km. The data rate for IEEE 802.11af per spatial stream reaches 26.7 Mbit/s for 6 and 7 MHz channels and 35.6 Mbit/s for 8 MHz channels. The maximum data rate can reach 426.7 Mbit/s for 6 and 7 MHz channels and 568.9 Mbit/s for

the 8 MHz channels by using four spatial streams and four bonded channels. In order avoid that the system does not create any undue interference with existing television transmissions, the White-Fi can utilise cognitive radio to detect transmissions and move to alternative channels or geographic sensing by using a geographic database and a knowledge of what channels are available to avoid interference with used channels. The White-Fi technology has benefits like propagation characteristics by using frequencies below 1 GHz that allow for greater distances or additional bandwidth. To achieve the required data throughput rates is necessary to aggregate several TV channels to provide the bandwidths that Wi-Fi uses on 2.4 and 5.6 GHz. Unused channels in any geographic area can vary in frequency, and special techniques need to be used for managing the data sharing across the different channels (e.g. as in technologies such as LTE).

3.8.3.4 Li-Fi

Light-based Li-Fi (light fidelity) is a wireless communication technology that uses light to transmit data and position between devices. The technology addresses some of the short comings of radio-based wireless communications and has applications for industrial IoT connectivity by improved solutions for security, scalability, bandwidth, interference, latency. Li-Fi uses the modulation of light intensity to transmit data at high speeds over the visible light (e.g. LED devices), ultraviolet, and infrared spectrums. Li-Fi can transmit at speeds of up to 100 Gbit/s. A unique feature of Li-Fi is that it combines illumination and data communication by using the same device to transmit data and to provide lighting [92]. Figure 3.27 shows the concept of a Li-Fi attocell network.

The illumination in the room is provided by several light fixtures with each light driven by a Li-Fi modem that serves as an optical base station or access point (AP). Each optical base station is connected to the core network by high speed backhaul connection. The light fixtures have an integrated infrared detector to receive signals from the terminals. The illuminating lights are modulated at high rates so the high frequency flickers are much higher than the refresh rate of a computer monitor are not visible to the occupants of the room [92].

Li-Fi's can safely function in areas susceptible to electromagnetic interference (e.g. aircraft cabins, hospitals, military, industrial). Li-Fi and visible light communication (VLC) are used in different IoT applications. Several companies offer uni-directional VLC products, that is not the same as Li-Fi. IEEE 802.15.7 supports high-data-rate visible light communication up to

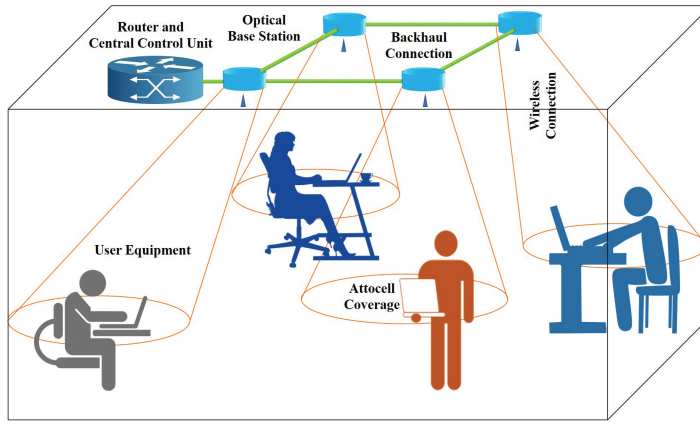

Figure 3.27 Li-Fi concept. Attocell networks applied to indoor wireless networking. Adapted from [92].

96 Mb/s by fast modulation of optical light sources which may be dimmed during their operation. IEEE 802.15.7 provides dimming adaptable mechanisms for flicker-free high-data-rate VLC. VLC technology using Li-Fi was demonstrated and data rates of 8 Gbit/s can be achieved over a single light source (e.g. LED) and complete cellular networks based on Li-Fi can be created.

3.8.3.5 Wi-SUN

Wi-SUN protocol is used for applications like the industrial grade utility, smart cities field area networks, agriculture, and asset monitoring.

Wi-SUN provides the communications profile definitions based on open standards for field area, IoT wireless networks, interoperability testing and certification for peer to peer wireless mesh networks based on IEEE 802.15.4g and IPv6. IEEE 802.15.4g is designed for Smart Utility Networks (SUN) supported by Wi-SUN Alliance. The 802.15.4g variation is aimed at deployments that can have millions of endpoints, deployed across large geographic scales, using minimal infrastructure in a peer-to-peer self-healing mesh. Wi-SUN devices may securely discover and join an existing Wi-SUN network. For this purpose, Wi-SUN relies on EAP and 802.1x security standards. Trickle timers are used to reduce interference and battery consumption while still maintaining responsive connectivity.

Wi-SUN is delivering star and mesh-enabled field area networks (FAN) to provide resilient, secure, cost-effective connectivity with ubiquitous coverage

in a range of topographical environments, from dense urban neighbourhoods to rural areas, with minimal additional infrastructure. Wi-SUN protocol characteristics include a coverage in the range of 1 km, high bandwidth of up to 300 kbps, low latency (e.g. 20 ms), resilient and scalable mesh routing, low power consumption (e.g. less than 2 uA when resting, 8 mA when listening), scalable networks of up to 5000 devices and security using public key certificates (e.g. AES, HMAC, dynamic key refresh, hardened crypto). Wi-SUN protocol is a viable solution as an implementation choice for LPWAN.

The Wi-SUN security is specified by implementation of the x.509 certificate-based, public-key infrastructure to authenticate devices, as well as Advanced Encryption Standard (AES) encryption and message integrity check. Devices protect their digital credentials either by storing them in hardened cryptographic processors that are resistant to physical tampering or by using physically unclonable function (PUF) technology.

3.8.4 Wireless and Cellular Wide Area Networks (WWAN)

3.8.4.1 WiMax IEEE 802.16

IEEE 802.16 technology has been put forward to overcome the drawbacks of WLANs and mobile networks. It provides different QoS scheduling for supporting heterogeneous traffic including legacy voice traffic, VoIP (Voice over IP), voice and video streams and the Internet data traffic. The prominent features of WiMAX include quality of service, high-speed Internet, facility over a long distance, scalability, security, and mobility.

IEEE 802.16 (WiMax) uses PMP network topology, in 2.3 GHz, 3.5 GHz, 5.8 GHz frequency spectrum, providing data rates of 40 Mbit/s for mobile, 1 Gbit/s for fixed networks with a coverage range of 50 km.

Different security solutions are enabled in WiMax networks, like Advanced Encryption Standard (AES) with 128-bit key: Rivest, Shamir and Adelman (RSA) with 1024-bit key and Triple Digital Encryption Standard (3-DES). Both Advanced Encryption Standard (AES) and Triple Digital Encryption Standard (3-DES) are symmetric encryption algorithms using a block-cipher method. Rivest, Shamir and Adelman (RSA) is an asymmetrical algorithm.

The air interface in IEEE 802.16 networks is secured by authentication procedures, secure key exchange and encapsulation. With encapsulating data

from authorised users, the base station limits the access of unauthorised users. Besides, it supports the Privacy Key Management (PKM) protocol for secure two-layer-key distribution and exchange and real-time confirmation of subscribers' identification, which ensures secure wireless data transport.

3.8.4.2 2G (GSM)

IoT applications rely on data transfer over *cellular technologies* such as 2G (GSM, D-AMPS, PDC), 2.5G (GPRS), 2.75G (EDGE), 3G (UMTS/WCDMA, HSPA, HSUPA , EvDO), 4G (i.e. LTE, LTE-A), 5G. M2M (*Machine-to-Machine*) connectivity is referred within the cellular context or MTC (*Machine-type Communication*) within 3GPP (3rd Generation Partnership Project). The approximate range is 35km for GSM and 200km max for HSPA. The data rate for typical download is in the range of 35–170 kps (GPRS), 120–384 kbps (EDGE), 384Kbps–2 Mbps (UMTS), 600 kbps–10 Mbps (HSPA), 3–10 Mbps (LTE).

3.8.4.3 3G (GSM/CDMA)

3G and 4G technologies such as 3GPP LTE are enabling technologies that offer wide area cover- age, QoS support, mobility and roaming support, scalability, billing, high level of security, the simplicity of management as well as connectivity of sensors through a standardised API [127] LTE-A (Long Term Evolution – Advanced) and Mobile WiMAX Release 2 (*Wireless MAN – Advanced* or *IEEE 802.*16 m) enabling higher speeds, more scalability, and low costs.

3.8.4.4 4G (LTE)

3GPP specified technologies such as eMTC (enhanced Machine-Type Communication), NB-IoT, and EC-GSM-IoT. The eMTC brings some LTE enhancements for MTC such as a new Power Save Mode (PSM). Release 14 brings new eMTC feature enhancements such as support for positioning and multicast, mobility for inter- frequency measurements, and higher data rates [131]. It brings enhancements such as lower costs, reduced data rate/bandwidth, and some other protocol optimisations.

Release-14 delivers new enhancements for the NB-IoT technology such as support for multicast, power consumption and latency reduction, mobility, and service continuity enhancements, etc. EC-GSM-IoT delivered EGPRS

enhancements, which in combination with PSM makes GSM/EDGE systems IoT ready.

This technology brings improvements such as extended coverage, support for massive number of devices: at least 50 000 per cell, improved security compared to GSM/EDGE, etc. The eMTC, NB-IoT, and EC-GSM-IoT has are described in the report on progress on 3GPP IoT [130]. QoS and network congestion are very challenging issues due to a huge number of deployed nodes (devices) [128]. In 4G LTE enhanced security was added such as unique identifiers (ID) for end-mobile device (UE), secure signalling between the UE and MME (Mobile Management Entity) and security for interworking between 3GPP networks and trusted non-3GPP users (e.g. using EAP-AKA) the UMTS Authentication and Key Agreement protocol.

3.8.4.5 5G

Cellular next generation 5G with high-speed mobility support and ultra-low latency is positioned to be the future of autonomous vehicles and augmented reality. 5G is also expected to enable real-time video surveillance for public safety, real-time mobile delivery of medical data sets for connected health, and several time-sensitive industrial automation applications in the future.

5G networks are expected to support the new IoT technologies, enabling IoT/IIoT device producers to develop and deploy new IoT devices and systems across multiple industries and provide IoT/IIoT applications globally.

The cell radius in 2 G systems is 35 km, in 3G systems 5 km, in 4G systems 100 m, and in 5G about 25 m to reuse the available RF spectrum more efficiently and to achieve higher data densities.

To implement the ultra-low latency and very high bit-rate applications, the connectivity technologies require more sizable contiguous blocks of the spectrum than those available in frequency bands that have been previously used. As the IoT/IIoT and the underlying connectivity technologies aim for worldwide coverage, there is a need for harmonised worldwide bands to facilitate global roaming and the benefits of economies of scale.

Technology advancements enabling 5G deployments can be divided into radio deployments in new bands in the sub-6GHz range, deployments in millimetre wave frequency bands and deployments in existing LTE bands. The first two categories combined can provide a forecasted population coverage of 55 percent in 2025, while the third category, where the networks can be upgraded to support 5G services in existing LTE bands by utilizing spectrum

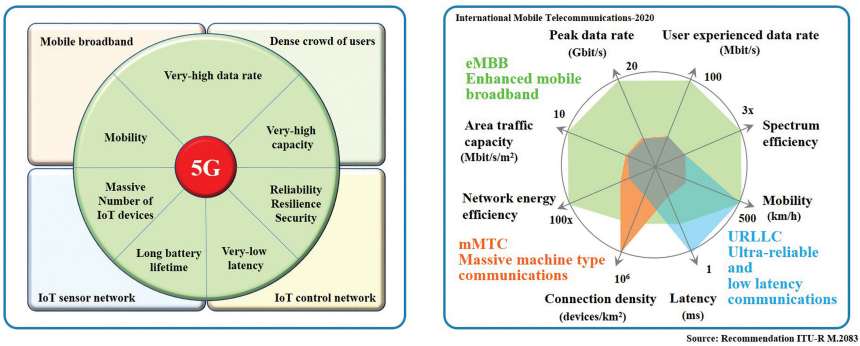

Figure 3.28 5G key capabilities of IMT-2020 defined by ITU.

sharing can provide a 10 percentage population coverage is achievable, creating a potential of up to 65 percent coverage in 2025 [82].

The ITU-R M.2083-0 define the capabilities of 5G to use cases such as mobile broadband, massive-machine communication, and mission-critical communication that include the full deployment of IoT solutions. 3GPP Releases 15 and 16 addresses the set of 5G standard specifications starting from LTE-Advanced Pro specifications. The functional specifications include mMTC (massive Machine Type Communications) requirements, specifications for eMBB (enhanced Mobile Broadband) and URLLC (Ultra-Reliable and Low Latency Communications), etc. A survey of the 5G cellular network architecture and key emerging technologies like interference management, spectrum sharing with cognitive radio, cloud computing, SDN, etc. is described in [119]. An overview of unique characteristics and characteristics of IoT and 5G technologies are given in [129].

The 5G technologies spectrum is distributed within three key frequency ranges to deliver the required coverage and support different use cases across various applications. The three ranges are: below-2 GHz, 2-8 GHz and above 24 GHz as presented in Figure 3.29.

Below 2 GHz low frequencies – low bands, support widespread coverage across urban, suburban, and rural areas and accelerate the IoT services for massive machine-type communications. Low-frequency bands extend the 5G mobile broadband to more extensive areas and deeper indoor environments. URLLC and mMTC type services benefit from enhancing coverage at the low frequency-bands.

The 2-8 GHz medium frequencies – mid bands offer a mixture of coverage and capacity benefits that includes spectrum within the 3.3-3.8 GHz range

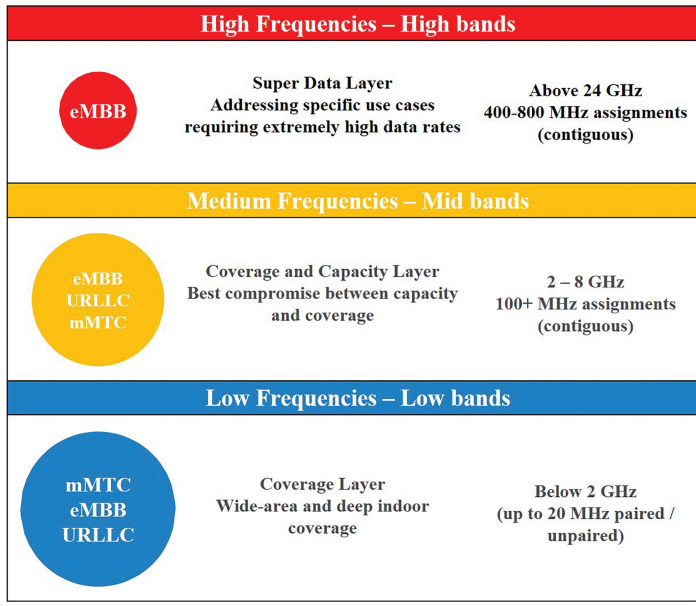

Figure 3.29 Ranges of the frequency spectrum for 5G technologies.

expected to form the base for many 5G services. The mid bands include the 2.3 GHz and 2.6 GHz frequencies. The unpaired (TDD) bands at 3300-4200, 4400-5000, 2500-2690 and 2300-2400 MHz deliver the best compromise between wide-area coverage and high capacity.

Above 24 GHz high frequencies – high bands are needed to ensure the ultra-high broadband speeds. The 26 GHz and/or 28 GHz bands have the most international support in this range. High-frequency bands are essential for providing additional capacity and delivering the extremely high data rates required by some 5G eMBB applications at specific locations ("hotspots"). The 400-800 MHz of contiguous spectrum per network operator is recommended from higher frequencies to achieve good return on investment and meet service requirements.

The frequency bands 24.25–27.5 GHz (global), 37–43.5 GHz (global), 45.5–47 GHz (regional/multi-country), 47.2–48.2 GHz (regional/multi-country) and 66–71 GHz (global) were identified for International Mobile Telecommunications (IMT) for the deployment of 5G networks by the World Radiocommunication Conference 2019 (WRC-19), that took place in Egypt, 28 October to 22 November 2019. WRC-19 took measures to ensure

Group 30 (GHz)	Group 40 (GHz)	Group 50 (GHz)	Group 70/80 (GHz)
24.25-27.5 31.8-33.4	37-40.5 40.5-42.5 42.5-43.5	45.5-47 47-47.2 47.2-58.2 50.4-52.6	66-71 71-76 81-86

Figure 3.30 Frequency bands identified by the WRC-19 for the deployment of 5G networks.

appropriate protection of the Earth Exploration Satellite Services, including meteorological and other passive services in adjacent bands.

The rage for 5G depends on the frequency bands used. The low band 5G has a range of tens of kilometres (similar range to 4G), mid-band 5G has several kilometres range, and high band 5G has hundreds of meters up to 1.5 km range. The data rates for the different bands are 30–250 Mbps for low-band 5G (600–700 MHz), 100–900 Mbps for mid-band 5G (2.5–3.7 GHz) and downloading speeds of 1–3 Gbps for high-band 5G (25–39 GHz and higher frequencies up to 80 GHz).

5G aims to integrate different portions of the unlicensed spectrum, to work in concurrence with licensed bands or independently. 5G NR-U (unlicensed) will support operations on 5 GHz spectrum used by Wi-Fi and new frequencies in the 6 GHz spectrum (e.g. future support includes the 3.5-4.2 GHz, 6-7 GHz, 37.37.6 GHz (US) and 57-70 GHz bands). 5G NR-U supports wideband carriers, flexible numerologies, beamforming, and dynamic TDD, where the uplink-downlink allocation could change to adapt to traffic conditions. The Licensed Assisted Access NR-U (LAA NR-U) aggregates carriers of both licensed and unlicensed spectrum, using NR and LTE carriers in the licensed band as anchors combined with NR-U carriers in the unlicensed spectrum to increase network capacity and provide broadband services at a lower cost. Stand-alone NR-U enables stand-alone operation in the unlicensed spectrum and can be deployed by different stakeholders. The use scenarios include local private networks, for specific uses industrial IoT applications or mobile enterprise broadband. Other potential applications are the delivery of broadband connectivity to public locations such as in stadiums or shopping malls. For private networks, 5G NR-U can deliver improved coverage capacity, mobility, increased reliability and precise timing due to the integration with time sensitive networking (TSN) that provides deterministic services over IEEE standard 802.3 Ethernet wired networks, guaranteeing low-latency packet transport, low packet delay variation and low packet loss.

The encryption is enhanced compared with 4G and the level of 5G security is not defined by the number of specified security mechanisms. A new approach is necessary for addressing 5G security to provide the security baseline of trustworthy, cost-efficient, and manageable 5G networks for various IoT/IIoT applications.

The Next Generation Mobile Network (NGMN) [121] introduced the concept of 5G network slicing and defined as an end-to-end (E2E) logical network/cloud running on a common underlying (physical or virtual) infrastructure, mutually isolated, with independent control and management that can be created on demand. The NGMN slice capabilities consists of several layers [120], the 5G Service Instance Layer (5GSIL) that represents different services which are to be supported, the 5G Network Slice Instance (5GNSI) that provides network characteristics which are required by a 5GSI, and the 5G Resource Layer (5GRL) that consists of physical resources (e.g. assets for computation, storage or transport including radio access) and logical resources (e.g. partition of a physical resource or grouping of multiple physical resources dedicated to a Network Function (NF) or shared between a set of NFs).

A network slice may be expressed by of cross domain components from separate domains in the same or different administrations, or components applicable to the access network, transport network, core network, and edge networks. Network slices are manageable, and programmable, self-contained, mutually isolated, to support multi-service and multitenancy.

Standardisation organisations have defined the network slicing considering various perspectives. The 3GPP [123] described network slicing as a "technology that enables the operator to create networks, customised to provide optimised solutions for different market scenarios which demand diverse requirements (e.g., in terms of functionality, performance, and isolation)" [124]. The ITU-T designated the network slicing as the a concept of network softwarisation that facilitates a Logical Isolated Network Partitions (LINP) composed of multiple virtual resources, isolated and equipped with a programmable control and data plane [123].

The network slicing overview and programmability are illustrated in Figure 3.31 and Figure 3.32. Network slicing concept can support multiple logical and self-contained networks on top of a shared physical infrastructure platform [122]. Network slicing present various operational challenges [125] such as configuration capability to allow users to adjust and modify the network functions as well as underlying resources within the network slice instance provided for them, monitoring capability for following the traffic

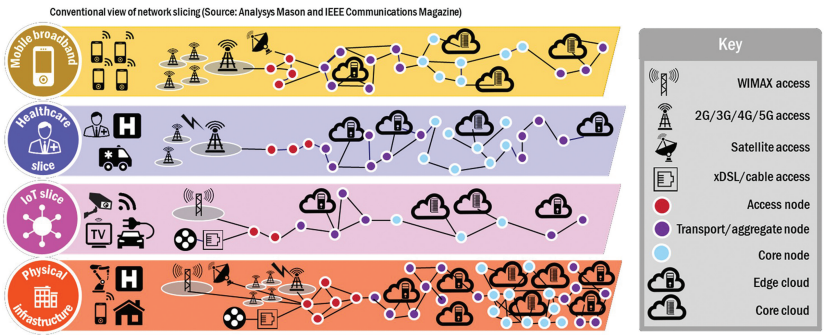

Figure 3.31 Network slicing overview. (Source: Analysis Mason and IEEE Communications Magazine).

characteristics and performance (e.g., data rate, packet drop, and latency), end user's geographical distribution, etc., in addition to per session/user/slice instance-based monitoring, etc. and control capability to enable the customers to utilise application programming interfaces (APIs) provided by the operator to control network service. Considering different applications where the network slicing concept is used, a slice encompasses a combination of relevant network resources, functions, and assets required to fulfil a specific business case or service, including operations support system/business support system (OSS/BSS), and DevOps processes.

Internal and external slices are used. The internal slices are partitions used for internal services of the provider, retaining full control and management of them and the external slices are partitions hosting customer services, appearing to the customer as dedicated networks/clouds/datacentres.In the initial phase of network slicing, operators are likely to launch a handful of slice types (e.g., eMBB, URLLC, and mMTC) with multiple tenants per slice. Over time, the number of slice types should increase and become more service specific (e.g., video gaming, smart meter connectivity, autonomous vehicles, specific IoT/IIoT use cases, etc.). For mobile network operators (MNOs) to capture value beyond basic connectivity, they must change their business models to address industry vertical-specific ecosystems. This will require a fundamental change in the way operators manage monetisation to address the multitude of use cases that network slicing will offer, including fixed wireless access, augmented reality/virtual reality (AR/VR) broadcasting, and industrial IoT [125].

Network programmability and slicing offer flexibility and control and the implementation steps are illustrated in Figure 3.32 [74]

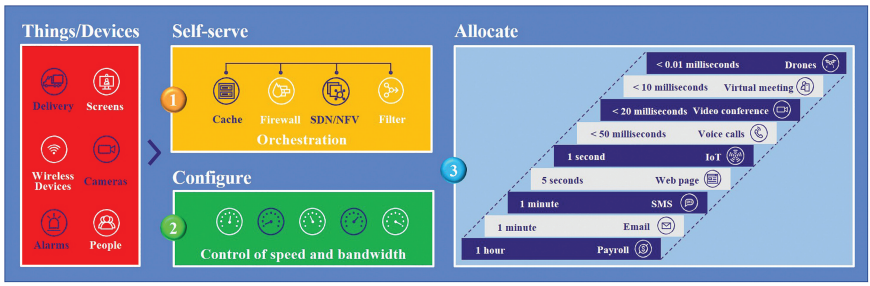

Figure 3.32 Network programmability and slicing. Adapted from Telestra [74].

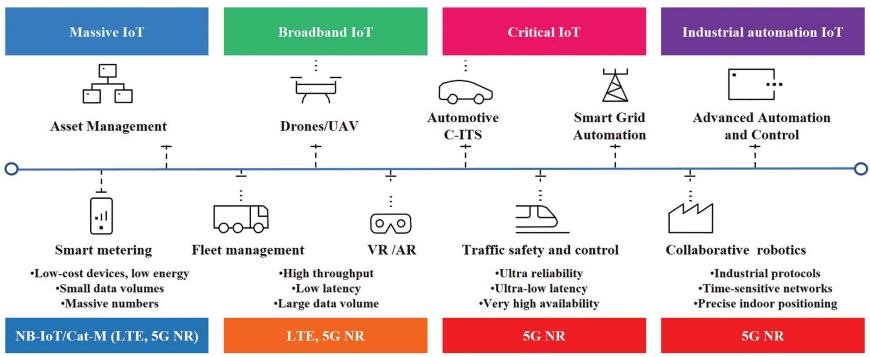

Figure 3.33 Cellular IoT use case segments. Adapted from [79].

Cellular IoT use cases have differing connectivity requirements and can be divided into four segments as presented in Figure 3.33. Verticals using Massive IoT include utilities with smart metering, healthcare in the form of medical wearables and transport with tracking sensors. NB-IoT and Cat-M will be able to fully co-exist in spectrum bands with 5G NR. Broadband IoT includes wide-area use cases that require higher throughput, lower latency, and larger data volumes (e.g. peak data rates in the multi-Gbps range and radio interface latency as low as 10ms) such as smart watches, drones/UAV, etc.

Critical IoT includes both wide-area and local-area use cases that have requirements for extremely low latency and ultra-high reliability such as such as interactive transport systems in the automotive industry, smart grids with real-time control and distribution of renewable energy in the utilities industry, and real-time control of manufacturing robots in the manufacturing industry. Industrial automation IoT includes very specific use cases, with the most demanding requirements coming from the manufacturing and industrial sites.

Table 3.4 IoT connections (billions) [82]

IoT	2019	2025	CAGR
Wide-area IoT	1,6	5,5	23%
Cellular IoT	1,5	5,2	23%
Short-range IoT	9,1	19,1	13%
Total	**10,7**	**24,6**	**15%**

Note: The figures for cellular IoT are also included in the figures for wide-area IoT.

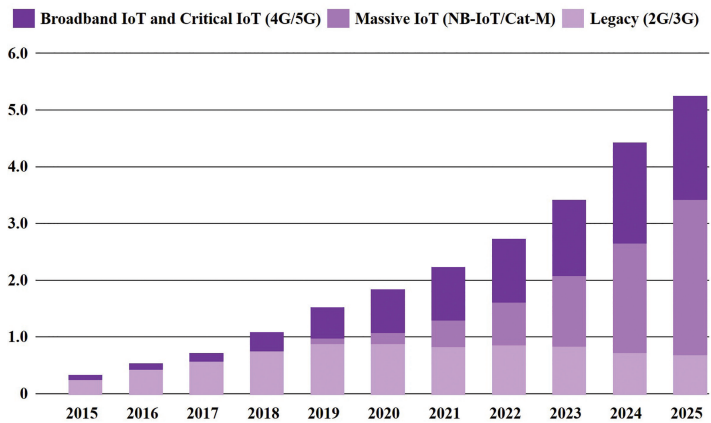

Figure 3.34 Cellular IoT connections by segment and technology (billion) [82].

Time-sensitive networks, industrial protocols running over ethernet, and very precise positioning will be needed [79].

It is projected that more than five billion active cellular IoT connections will be available by 2025 which are split between LTE-M and NB-IoT as shown in Table 3.4. The cellular IoT connections by segment and technology are presented in Figure 3.34.

The massive IoT technologies NB-IoT and Cat-M1 are rolled out at a slightly slower pace in 2020 than previously forecast due to the impact of the pandemic crisis. 2G and 3G connectivity enable the majority of IoT applications. massive IoT connections reached around 100 million connections at the end of 2019.

NB-IoT and Cat-M technologies complement each other, and it is projected to account for 52 percent of all cellular IoT connections [82]. Cat-M includes both Cat-M1 and Cat-M2. Only Cat-M1 is being supported today. Commercial devices for massive IoT include different types of meters, sensors, trackers, and wearables.

The deployment of 5G New Radio (NR) will increase data rates and the broadband IoT that includes wide-area use cases that utilise higher throughput, lower latency, and larger data volumes. Critical IoT that is used for time-critical communications in wide- and local-area use cases requires guaranteed data delivery with specified latency targets, will be introduced in 5G networks with the advanced time-critical communication capabilities of 5G NR in 2021. The use cases for critical IoT include cloud-based AR/VR, robotics, IoRT, autonomous vehicles, edge computing, advanced cloud gaming, and real-time coordination and control of machines and processes.

The 5G devices are initially supporting mobile broadband capabilities, and performance is expected to evolve towards time-critical communication capabilities where needed, via software upgrades on devices and networks [82].

3.8.4.6 NB-IoT (LTE Cat NB1 and LTE Cat NB2)

The 3GPP has started the standardisation of a set of low cost and low complexity devices targeting machine type-communications (MTC) in Release 13. The standardisation addresses the IoT market from the enhanced machine type communications (eMTC), the narrow band IoT (NB-IoT) and the EC-GSM-IoT. eMTC is an evolution of the work developed in Release 12 that can reach up to 1 Mbps in the uplink and downlink and operates in LTE bands with a 1.08 MHz bandwidth.

In Release 14 3GPP introduced five new FDD frequency bands for NB-IoT: 11 (central frequencies – UL 1437.9 MHz, DL 1485.9 MHz), 21 (central frequencies – UL 1455.4 MHz, DL 1503.4 MHz), 25 (central frequencies – UL 1882.5 MHz, DL 3962.5 MHz), 31 (central frequencies – UL 455 MHz, DL 465 MHz), and 70 (central frequencies – UL 1702.5 MHz, DL 2007.5 MHz).

The NB-IoT, CAT-NB1, use the existing 4G/LTE network [216]. NB-IoT and LTE coexist, and re-use of the LTE physical layer and higher protocol layers benefits the technology implementation. NB-IoT has been designed for extended range, and the uplink capacity can be improved in bad coverage areas. NB-IoT devices support three different operation modes [216, 217]:

- Stand-alone operation (in other spectrum/non-LTE) – utilizing one or more GSM carriers (bandwidth of 200 kHz replacements) utilizing for example the spectrum currently being used by GERAN systems as a replacement of one or more GSM carriers.

- Guard band operation (in the guard band of an LTE carrier) – utilizing the unused resource blocks within an LTE carriers' guard-band (frequency bands to prevent interference).
- In-band operation (within an LTE carrier) – utilizing resource blocks within a normal LTE carrier.

NB-IoT and LTE-M defined in 3GPP Release 13, has two user network equipment categories: Cat-NB1 for NB-IoT networks and Cat-M1 for LTE-M networks [217].

The coverage enhancement modes introduced as part of LTE-M can also be optionally supported by ordinary LTE user equipment categories. Ten years battery lifetime and low-cost devices are available and support a high number of low throughputs in IoT for different applications.

NB-IoT is designed to exist in independently licensed bands, in unused 200 kHz bands that have previously been used for GSM or CDMA, or on LTE base stations that can allocate a resource block to NB-IoT operations or in their guard bands (where regulations allow it).

The LTE Cat NB1 provides data rates of 66 kbps (multi-tone), or 16.9 Kbit/s (single tone), while LTE Cat NB2 data rates of 159kbps.EC-GSM-IoT

EC-GSM-IoT is an evolution of EGPRS towards IoT, operates in the frequency band of 850-900 MHz (GSM bands), uses a star topology, provides a data rate of 70 kbps (GSMK), 240 kbps (8PSK) with a coverage approximate range of 15km.

The EC-GSM-IoT has improved security, compared to the existing GSM/GPRS networks – offers integrity protection, mutual authentication and implements stronger ciphering algorithms. The EC-GSM-IoT technology implementation is based on software upgrades of the existing GSM networks.

3.8.4.7 LTE-M

LTE-M is a LPWAN cellular radio technology standard developed by 3GPP to enable a wide range of cellular devices and services IoT applications.

LTE-MTC Cat 0 uses the technology frequency bands (700 MHz, 800 MHz, 900 MHz, 1700 MHz, 1800 MHz, 1900 MHz, 2300 MHz, 2400 MHz, 2500 MHz, 2700 MHz). The technology uses a star topology, provides a data rate of 1 Mbps with a coverage range of 10km that is variable and depends on frequency bands, propagation conditions etc. LTE-MTC system and security management is enhanced compared to LTE, as numbers of devices in LTE MTC network are very large. The request defined in 3GPP TS 22.368 is

"LTE MTC optimisations shall not degrade security compared to non-MTC communications".

LTE-M eMTC (Cat M1, Cat M2) uses the technology frequency bands (400 MHz 450 MHz, 600 MHz, 700 MHz, 800 MHz, 900 MHz, 1400 MHz, 1500 MHz, 1700 MHz, 1800 MHz, 1900 MHz, 2100 MHz, 2300 MHz, 2400 MHz, 2500 MHz, 2600 MHz, 2700 MHz). The technology uses a star topology, provides a data rate of 1 Mbps for LTE-M Cat M1 and $\approx$4 Mbps DL/ $\approx$7 Mbps UL for LTE-M Cat M2 with a coverage range of 10 km that is variable and depends on frequency bands, propagation conditions etc. LTE-M technology offers SIM-based security features requiring device authentication to connect to the network. Security system and management is more complex in LTE-M (eMTC) than LTE due to massive connectivity that is supported in LTE-M (eMTC) networks.

Two innovations support LTE-M improve battery life: LTE eDRX (extended discontinuous reception) and LTE PSM (power saving mode). LTE-M delivering data rates of around 150-200kbps (1.4MHz bandwidth), with the capacity to support higher bandwidth workloads, addresses mobility credentials and provides a 50 ms/100 ms latency response.

The advantage of LTE-M over NB-IoT is its comparatively higher data rate, mobility, and voice over the network (VoLTE). LTE-M benefit from reduced complexity with devices implemented with a very simple frontend and antenna configuration.

LTE-M theoretically can be extremely power-efficient but because eDRX and PSM are only recently being deployed, their power efficiency is still evaluated.

An overview of the cellular IoT LPWAN technologies defined by 3GPP (e.g. NB-IoT, LTE-M, EC-GSM-IoT) their features and properties is presented in Table 3.5.

A comparison of the IoT personal, local, wide area networks is presented in Table 3.6.

3.8.5 Low Power Wide Area Networks (LPWAN)

3.8.5.1 LoRaWAN

LoRaWAN defines a communication protocol and network architecture for IoT low-power wide area networks (LPWANs) and is designed to address the requirements for low power consumption (i.e., long battery life), long range, and high low data rate ($<$ 2 kbit/s) while maintaining low operating and deployment costs.

Table 3.5 Overview of the cellular IoT LPWAN technologies defined by 3GPP

		LTE-M				NB-IoT		
		LC-LTE/MTCe		eMTC				
	LTE Cat 1	LTE Cat 0	LTE Cat M1	LTE Cat M2	non-BL	LTE Cat NB1	LTE Cat NB2	EC-GSM-IoT
3GPP Release	8	12	13	14	14	13	14	13
Downlink Peak Rate	10 Mbit/s	1 Mbit/s	1 Mbit/s	∼4 Mbit/s	∼4 Mbit/s	26 kbit/s	127 kbit/s	474 kbit/s (EDGE) 2 Mbit/s (EGPRS2B)
Uplink Peak Rate	5 Mbit/s	1 Mbit/s	1 Mbit/s	∼7 Mbit/s	∼7 Mbit/s	66 kbit/s (multi-tone) 16.9 kbit/s (single tone)	159 kbit/s	474 kbit/s (EDGE) 2 Mbit/s (EGPRS2B)
Latency	50-100ms	Not deployed	10-15ms			1.6-10s		700 ms-2s
Antennas	2	1	1	1	1	1	1	1–2
Duplex Mode	Full Duplex	Full or Half Duplex	Full or Half Duplex	Full or Half Duplex	Full or Half Duplex	Half Duplex	Half Duplex	Half Duplex
Device Receive Bandwidth	1.4–20 MHz	1.4–20 MHz	1.4 MHz	5 MHz	5 MHz	180 kHz	180 kHz	200 kHz
Receiver Chains	2 (MIMO)	1 (SISO)	1 (SISO)	1 (SISO)	1 (SISO)	1 (SISO)	1 (SISO)	1–2
Device Transmit Power (dBm)	23	23	20/23	20/23	20/23	20/23	14/20/23	23/33

Table 3.6 Comparison of different IoT protocols

	Personal and Local Area		Wide Area				
	BLE	ZigBee	LoRaWAN	MIOTY	Sigfox	LTE-M	NB-IoT
Network type	Mesh	Mesh	LPWA	TS-UNB	UNB	LPWA	UNB
Standard	IEEE 802.15.4	IEEE 802.15.4	Proprietary	ETSI TS 103357	Proprietary	3GPP Rel. 12-14	3GPP Rel. 13-14
Bandwidth	2 MHz	600 kHz-5 MHz	125 kHz	200 kHz	100 kHz	1.4MHz	200kHz
Cell capacity	n/a	n/a	40,000	>1,000,000	1,000,000	1,000,000	200,000
Max nodes	1,000 +	250	n/a	n/a	n/a	n/a	n/a
Max range	30 m	10-100 m	5-15 km	5-15 km	10-30 km	11 km	15 km
Throughput	2 Mbps	250 Kbps	50 Kbps	512 bps	600 bps	1 Mbps	200 Kbps
Latency	< 3 ms	15 ms	1-100 ms		>20s	10-15 ms	< 10ms
Payload	350 bytes	102 bytes	243 bytes	10 – 192 bytes	12/8 bytes (UL/DL)	–	1.6 Kbytes
Alliance	Bluetooth SIG	Zigbee Alliance	Private	MIOTY Alliance	Private	GSMA / 3GPP	GSMA / 3GPP

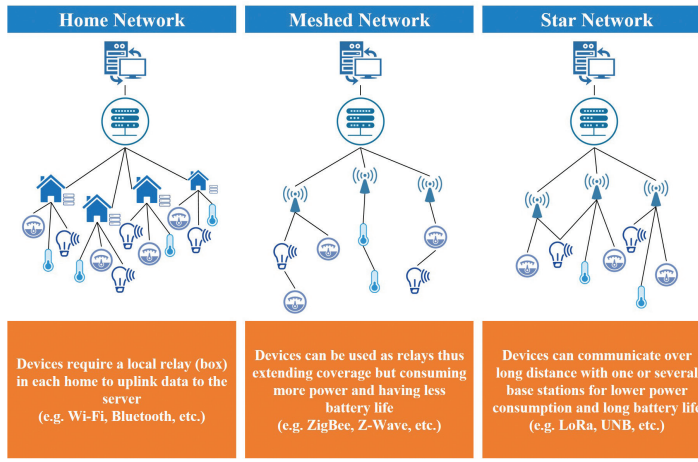

Figure 3.35 LoRa network topologies [111].

The LoRa physical layer uses chirp spread spectrum modulation a spread spectrum technique where the signal is modulated by chirp pulses (frequency varying sinusoidal pulses) hence improving resilience and robustness against interference, Doppler effect and multipath, characterised by low power usage and increased communication range allowing a single base station to cover hundreds of square kilometres. LoRa and LoRaWAN, enable long battery life for devices in the field and covers the IoT communication needs between local wireless such as Bluetooth, Wi-Fi and cellular-based wireless.

LoRaWAN features data rates of 27 kbps (50 kbps when using FSK instead of LoRa), and a single gateway can collect data from thousands of nodes deployed at 5-15 km distance. LoRaWAN networks are organised in a star of stars topology, in which gateway nodes relay messages between end devices and a central network server. End devices send data to gateways over a single wireless hop and gateways are connected to the network server through a non-LoRaWAN network (e.g. IP over Cellular or Ethernet). Communication is bi-directional, even uplink communication from end devices to the network server is preferred. LoRa network topologies are illustrated in Figure 3.35

LoRa has several key features that have made it a strong choice for organisations that are looking to achieve efficiency and cost savings with IoT. However, as use cases for IoT deployments using LoRa become more advanced, and as IoT requirements scale, a new way of approaching the problem of how to best benefit from LoRa is required. That is where Super-B comes in. Super-B is an advanced protocol that is built on top of the

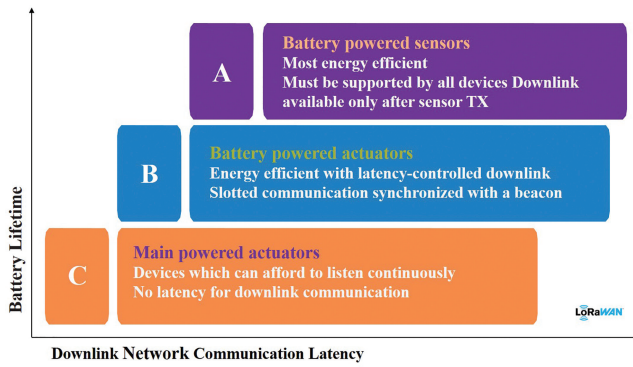

Figure 3.36 LoRaWAN classes of communications between sensors and gateways [73].

LoRaWAN protocol while utilizing the same network infrastructure as LoRa. By selecting IoT solutions using the Super-B protocol, many issues can be alleviated so that organisations can easily [73]:

- Densify IoT networks to meet their organisation's growing requirements
- Vastly improve data delivery and QoS
- Better secure IoT networks through firmware over the air (FOTA) updates
- Deliver strong return on investment (ROI) by reducing the need to add hardware to scale networks

The LoRaWAN "classes" are illustrated in Figure 3.36. LoRaWAN has three "classes" of communications between sensors and gateways [73]:

- **Class A** – The gateway is passive and in listen-only mode. Sensors utilise listen before talking and send messages whenever there is a message to send. Because of its random nature, Class A is prone to significant noise and interference, and thus packet loss, which can be exacerbated by the all-too-common practice of sending a message a few times to increase the likelihood of delivery.
- **Class B** – The network establishes a session to communicate but it's set up and torn down, forcing the sensor to search for the gateway each time it connects, creating inefficiencies. Has limited scheduled receive slots, however, lacks the efficiency needed to scale. Class B sensors frequently cause interference with Class A sensors.
- **Class C** – Has a bidirectional link that is never torn down, so is "always on" to send and receive messages.

Super-B protocol rides atop the LoRa protocol, utilizing standard LoRa but extending its structure to allow for the scheduling of messages from gateway to devices (and vice versa) while maintaining extremely low power. Super-B capitalises on the best parts of LoRa A and LoRa B to deliver scalable, secure IoT.

The LoRa technology is defined by three main parameters: spreading factor (SF), bandwidth (BW), and carrier frequency. The transfer rate varies by using orthogonal spreading factors which is a compromise between distance and emission power [108]. A pseudorandom channel hopping method is natively used in LoRaWAN to distribute transmissions over the pool of available channels, thereby reducing the collision probability.

Several companies offer LoRa-based IoT solutions for different applications. Semtech [109] has released LoRa Edge to simplify IoT deployments for indoor and outdoor asset management applications. The LoRa Edge geolocation platform integrates a LoRa transceiver with Wi-Fi and Global Navigation Satellite System (GNSS) scanning technologies. The hardware is connected to the geolocation and device management services that operate on the LoRa cloud platform. The optimisation of the capacity of the LoRaWAN network, and the possibility to perform traffic slicing for guaranteeing specific requirements in a service basis, remain as open issues. Other issues to be addressed are new channel hopping methods to meet traffic requirements when there are latency, jitter or reliability constraints (i.e. downlink ACKs for all packets), that cannot be adapted according to the noise level of each channel.

The nature of ALOHA-based access is not optimal to serve deterministic traffic and using a hybrid Time Division Multiple Access (TDMA) access on top of LoRaWAN could provide new use cases and adds more flexibility.

The location of LoRa end devices for IoT applications is necessary for different IoT use cases and GPS-based solutions too expensive and requires extra computing resources and energy consumption. TDOA-based (Time Difference Of Arrival) triangulation techniques for LoRaWAN could provide solutions for dense gateway deployments. Future developments could include the integration of cognitive radio into the LoRaWAN standard. Considering the random-based access in unlicensed bands of LoRaWAN, the performance achieved in isolated networks is cross-examined in scenarios with co-existing gateways and limited number of available channels. It is key to devise coordination mechanisms between gateways from the same or different operators to limit interference and collisions and provide co-existence mechanisms

that encompass coordination and reconfiguration protocols for gateways and end-devices. [100, 110].

It should be noted that 1) the reach and autonomy of the IoT devices based LoRa connectivity will depend on the applications requirements. For instance, the reach depends not only on the transmitter performances, but also on the distance (between the end-devices and the closest gateway) and antenna gains. The battery life depends on the reading duty cycle and the number of bytes transmitted in the payload. The reach and the autonomy will be reduced when these parameters have high numbers. 2) LoRaWAN allows developing IoT private networks at cost effective for applications (e.g. in agriculture and Smart city) requiring long reach (around 10 kms) and low data rates.

3.8.5.2 Weightless – N, W, P

Weightless is a LPWAN standard offered by the Weightless Special Interest Group (SIG) in three different protocols: Weightless-N, Weightless-W, and Weightless-P. Each of these are designed to support different end-user cases and modalities.

The Weightless technology can operate in any frequency band and is currently defined for operation in license-exempt sub-GHz frequency bands (e.g. 138MHz, 433MHz, 470MHz, 780MHz, 868MHz, 915MHz, 923MHz). The approximate range is 2km (P), 5km (W, N).

Weightless-N system supports an ultra-narrowband connectivity and has many similarities with Sigfox. It is sometimes considered like a LoRa-based version of Sigfox. Wireless-N system comprises networks of partners instead of a completely end to end enclosed approach. The modulation used is BPSK and is usually intended for uplinks related to sensor data.

Weightless-N was designed to expand the range of Weightless-W and reduce the power consumption (e.g. battery lifetime up to 10 years) at the expense of data rate decrease (from up to 1 Mbps in Weightless-W to 100 kbps in Weightless-N). Weightless-N is based on the Ultra Narrow Band (UNB) technology and operates in the UHF 800-900 MHz band; it provides only uplink communication.

Weightless-W was developed as a bidirectional (uplink/downlink) solution to operate in TV whitespaces (470-790 MHz). It is based on narrowband FDMA channels with Time Division Duplex between uplink and downlink; data rate ranges from 1 kbps to 1 Mbps and battery lifetime is 3 to 5 years.

The characteristics of Weightless-P are bidirectional, fully acknowledged communication for reliability, optimised for a large number of

low-complexity end devices with asynchronous uplink-dominated communication with short payload sizes (typically $<$ 48 bytes) and standard data rates from 0.625kbps to 100kbps.

The Weightless-P systems include end devices (ED) (e.g. the leaf node in the network, low-complexity, low-cost, usually low duty cycle), base stations (BS) (e.g. the central node in each cell, with which all EDs communicate via a star topology) and base station network (BSN) (e.g interconnects all BS of a single network to manage the radio resource allocation and scheduling across the network, and handle authentication, roaming and scheduling).

Weightless-P is proposed as a high-performance two-way communication solution that can operate over 169, 433, 470, 780, 868, 915 and 923 MHz bands. Cost of the terminals and power consumption are higher than in Weightless-N, with a battery lifetime of 3 to 8 years.

In Weightless standard AES-128/256 encryption and authentication of both the terminal and the network guarantees integrity whilst temporary device identifiers offer anonymity for maximum security and privacy. OTA security key negotiation or replacement is possible whilst a future-proof cipher negotiation scheme with a minimum key length of 128 bits protects long term investment in the network integrity.

3.8.5.3 Sigfox

Working in a similar approach to Weightless-N, Sigfox is a low cost, low power, and very reliable means for connecting sensors and related devices. It is of great use in the IoT domain and is expected to remain. The Sigfox protocol focuses on attributes such as extended autonomy using very low energy, simple setup and maintenance with no configuration, quick deployment, low cost, low bandwidth, small message structure (up to 12bytes) and the possibility to be combined with Wi-Fi, BLE or cellular (e.g. GPRS, 3G, 4G, etc.).

Sigfox employs the differential binary phase-shift keying (DBPSK) and the Gaussian frequency shift keying (GFSK) that enables communication using the Industrial, Scientific and Medical ISM radio band which uses 868 MHz in Europe and 902 MHz in the US. It utilises a signal that is extremely narrowband (100 Hz bandwidth), called “Ultra Narrowband” and requires little energy, being termed “Low-power Wide-area network (LPWAN)”. It is based on Random Frequency and Time Division Multiple Access (RFTDMA) and achieves a data rate around 100 bps in the uplink, with a maximum packet payload of 12 Bytes, and a number of packets per device that cannot exceed

14 packets/day. It consumes only 50 microwatts and can deliver a typical stand-by time 20 years with a 2.5Ah battery.

The network is based on one-hop star topology and requires a mobile operator to carry the generated traffic. The signal can also be used to easily cover large areas and to reach underground objects. The range is 30-50km (rural environments), 3-10km (urban environments) and the data rates varies from 10 to 1000bps.

Security is implemented in the Sigfox devices during the manufacturing process, when each Sigfox Ready device is provisioned with a symmetrical authentication key. Additional security is supported by radio technology. The Sigfox technology encryption is designed for use with short Sigfox messages. End-to-end encryption solutions are applicable to the Sigfox networks and applications.

3.8.5.4 Random-phase multiple access (RPMA) – Ingenu

Random-phase multiple access (RPMA) is a bidirectional IoT communications technology developed by Ingenu that operates in the 2.4 GHz ISM band. Ingenu claim that RPMA coverage can extend up to 70 kilometres, throughput of 19 kbps/MHz and an uplink link budget of 180 dB and 185 dB for downlink in the US. Citing security concerns, RPMA does not support IP connectivity. Instead, Ingenu choose to provide a REST API instead for data access [116]. The Ingenu proprietary LPWAN technology in the 2.4 GHz band, is based on Random Phase Multiple Access (RPMA) to provide M2M industry solutions and private networks.

The Ingenu technology provides high data rate up to 624 kbps in the uplink, and 156 kbps in the downlink. The energy consumption is higher, and the range is shorter (a range around 5-6 km) due to the high spectrum band used.

Security in RPMA wireless technology is built on 128-bit AES. It offers security features such as: mutual authentication, message integrity and replay protection, message confidentiality, device anonymity, authentic firmware upgrades and secure multicasts.

3.8.5.5 Neul

Neul technology operates in the sub-1GHz band and leverages very small slices of the TV White Space spectrum to deliver high scalability, high coverage, low power, and low-cost wireless networks.

Neul systems are based on the Iceni chip, which communicates using the white space radio to access the high-quality UHF spectrum, available due to the analogue to digital TV transition.

Data rates can be anything from a few bits per second up to 100kbps over the same single link; and devices can consume 20 to 30mA and that offer a10 to 15 years lifetime in the field.

The frequencies used by Neul are 900MHz (ISM), 458MHz (UK), 470-790MHz (White Space) providing a range of 10km.

The wireless communications links between the gateway (base station) and the network nodes are encrypted.

3.8.5.6 Wavenis

Wavenis is a wireless technology for ultra-low power and long-range Wireless Sensor Networks (WSNs) and promoted by Wavenis Open Standard Alliance.

Wavenis uses tree and star network technologies, in the 433MHz, 868MHz, 915MHz frequency spectrum providing low data rates of 9.6kbps (433 & 868MHz), 19.2kbps (915MHz) and a coverage range of approximate 1 km.

Wavenis technology is supported by 128-bit AES encryption.

3.8.5.7 WAVIoT (NB-Fi – Narrowband Fidelity)

NB-Fi (Narrowband Fidelity) is a narrow band protocol which communicates on the sub 1GHz ISM sub bands. DBPSK is used as the modulation scheme in the physical layer. WAVIoT gateways can provide -154 dBm of receiver sensitivity, and cover over 1 million nodes. On WAVIoT-developed devices, short data bursts use 50mA of current, and in idle mode – a few ţA are used. Devices have a lifetime of up to 20 years, and a 176 dBm link budget.

WAVIoT uses a star network technology, in the 315 MHz, 433 MHz, 470 MHz, 868 MHz, 915 MHz frequency spectrum providing data rates of 10–100 bps and a coverage range of approximate 50 km.

All WAVIoT data is encrypted bidirectionally from the device to the server using an XTEA 256-bit key block cipher.

3.8.5.8 MiWi

MiWi is a wireless technology for low-power, cost-constrained networks, such as industrial monitoring and control, home and building automation, remote control, wireless sensors, lighting control, HVAC systems and automated meter reading.

MiWi uses mesh and star network technologies, in the 2.4GHz, 700MHz/800MHz/900MHz frequency spectrum providing low data rates of 20kbps and a coverage range of approximate 300m.

The MiWi protocol follows the MAC security definition specified in IEEE 802.15.4 and is based on 128-bit AES model. MiWi security mechanisms are classified as three modes: AES-CTR mode that encrypts MiWi protocol payload, AES-CBC-MAC mode that ensures the integrity of the MiWi protocol packet and AES-CCM mode that combines the previous two security modes to ensure both the integrity of the frame and encrypt the MiWi protocol payload.

3.8.5.9 MIOTY™ (TS-UNB)

MIOTY™ (TS-UNB) is a LPWAN, ETSI-standardised telegram-splitting ultra-narrowband (TS-UNB) technology (TS 103-357) that supports MYTHINGS. MYTHINGS is a wireless connectivity platform designed for large-scale industrial and commercial IoT networks.

Telegram-splitting is a standardised LPWAN technology in the licensed-free sub-GHz radio spectrum that feature a data rate of 512 bit/s and divides at the physical layer, an ultra-narrowband telegram into multiple equal-sised sub-packets, each of which is randomly sent at a different time and carrier frequency. Each sub-packet has a much smaller size than the original telegram, and the on-air time is reduced to 16 milliseconds. As the airtime of the sub-packets is much shorter than that of existing LPWANs, the chance of collisions with another message is very low. An algorithm in the base station permanently scans the spectrum for MIOTY sub-packets and reassembles them into a complete message. The technology has high-redundancy as up to 50% of the sub-packets can be lost without reducing the information content.

MIOTY enables energy-efficient, robust, and reliable transmission of sensor data over distances of up to several kilometres. The telegram-splitting mechanism is illustrated in Figure 3.37 [99, 103]. The mechanism allows the implementation of scalable networks for very high-density solutions. A MIOTY network is private and can have over a million devices that can transmit up to 1,5 million data packets a day to a single gateway, with no loss of information in environments with physical obstructions and poor propagation properties.

The protocol fragments data packets into numerous subpackets or telegrams and distributes them over time and frequency. MIOTY is designed to support up to 15 km range in flat terrain, up to 65,000 messages per hour, 407 bits/s, have enhanced interference-resilience features for use in shared

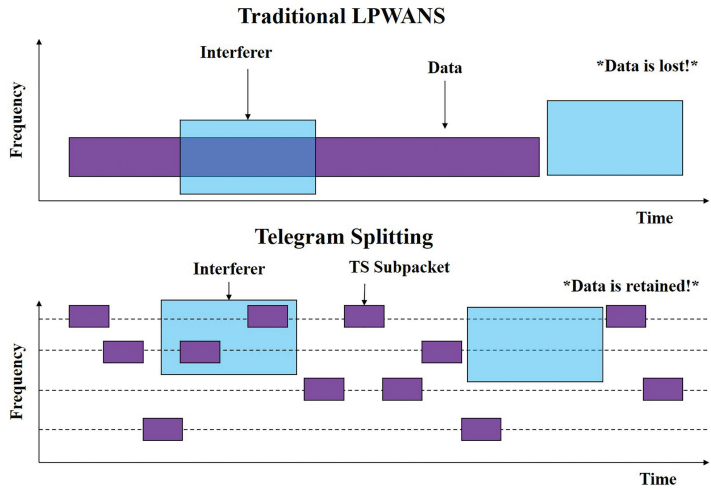

Figure 3.37 Telegram-splitting transmission. Adapted from [99, 103].

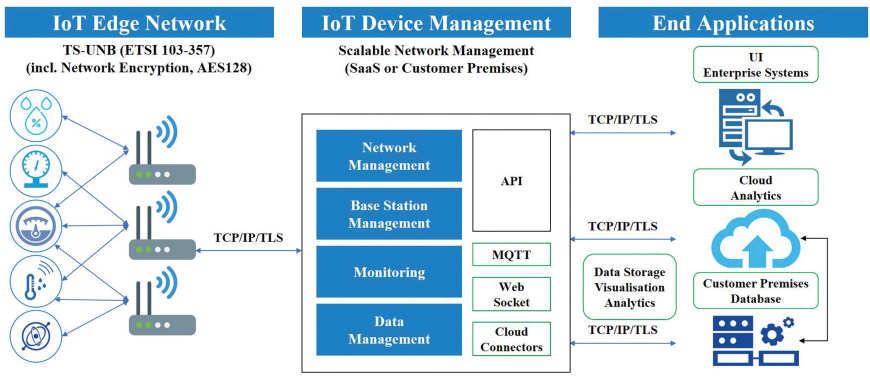

Figure 3.38 MYTHINGS network. Adapted from [99].

spectrum segments, (e.g. 868 MHz, 915 MHz ISM bands) via the time- and frequency-distribution approach, and support mobility up to approximately 80 km/h [104]. MIOTY does not support IP connectivity

An example of implementation of an IoT network based on MIOTY is illustrated in Figure 3.38 [99]. The MYTHINGS base station IoT gateway is leveraging the MIOTY wireless stack including the built-in software, the platform for device on and off-boarding, cloud/backend integration, data monitoring, network troubleshooting and indoor localisation. The MYTHINGS IoT gateway has the capacity to handle millions of messages a day from thousands

of endpoints and provides a web-based user interface for management, cloud integration and MQTT interface for large-scale data collection and business analytics.

3.8.6 Satellite

During the last years, satellite service providers (e.g. Argos, Iridium, ORBCOMM, Inmarsat's Broadband Global Area Network, etc.), have integrated the IoT services into their portfolio and commercial satellite IoT initiatives (e.g. HIBER, DIAMOND, KEPLER), have started using nanosatellites for satellite IoT communication. ORBCOMM operates a satellite network dedicated to IoT providing two-way data communications in remote areas of the world via a network of low-earth orbit (LEO) satellites and ground stations. The OG2 satellites provide network redundancy, minimal line of site issues for complete global coverage, while VHF frequency furthers signal propagation.

These small satellite systems can cover oceans, rural areas, polar regions, and the integration of the nanosatellites networks with IoT networks provide new ways of monitoring the climate change, pollution, and global/regional disasters.

The miniaturisation of satellite technology for LEO (Low Earth Orbit) allows to use these satellites to serve as a backhaul for narrow-band IoT communication. The IoT applications using the satellite communication channels require data rates in the range of up to hundreds of kbps or tens of Mbps, low power consumption for battery life of the IoT nodes up to several years, compact antenna design, and coarse to even omni-directional antenna pointing.

These requirements narrow down the considerable frequency bands for Earth-to-satellite to VHF (0.03–0.3 GHz), UHF (0.3–1 GHz), L band (1–2 GHz), and S band (2–4 GHz). Frequencies, from VHF/UHF to Ka bands, are usable, with several nanosatellite using the UHF band, due to its omni-directional pattern, robustness, and low power consumption.

The IoT backbone requirements ($\sim$Mbps rate and $\sim$1000 km range) need to consider the S band, or X the X band due to the trade-off between bandwidth, directionality ($\approx 10^{\circ}$), antenna (e.g., patch antenna on S/C), and transceiver size and power consumption [113].

The satellite communications offer opportunities by adding capacity on GEO (geostationary) satellites in C-, Ku- and Ka-band for direct or backhaul

connectivity to deploying new LEO (low earth orbit) or HEO (highly elliptical orbit) constellations, optimised for the IoT solutions and applications.

The satellite industry responding to the IoT connectivity require solutions with low cost/low power direct to satellite connectivity and various combinations of terrestrial (cellular and LPWAN) IoT access networks and satellite backhaul.

Comparable to the cellular or Wi-Fi backhaul service, the IoT gateway backhaul over satellite emerges as a new SATCOM application segment.

The IoT market is developing around ultra-low-cost terrestrial radio transmission standards for IoT such as LoRa, Sigfox, LTE-M or NB-IoT targeting low cost per radio transmitter with dedicated gateways to concentrate larger numbers of IoT devices in range of operation. For the satellite industry connecting these gateways is leading to a new satellite application segment [114].

3.8.7 IoT Application Protocols

The IoT application protocols refer to OSI model (ISO/IEC 7498) layer 5, 6 and 7, layers that are responsible for managing communication sessions, data formatting (e.g. date translation, character encoding, data compression, encryption/decryption, presentation) and high-level APIs, resource sharing, remote file access, etc.

These layers are based on HTTP that is not suitable for resource constrained environments (e.g. verbose, requires large parsing overhead, etc.).

The requirements for many IoT applications have accelerate the developments of alternative protocols that address the resource constraint environments.

An extensive overview of the existing IoT application layer protocols is given in the following sub-sections. An illustration of the deployment of the IoT application protocols in the IoT applications is given in Figure 3.39.

3.8.7.1 The Advanced Message Queuing Protocol (AMQP)

The AMQP is largely used in the financial sector applications and applied to other types of applications. The protocol standardised by Organisation for the Advancement of Structured Information Standards (OASIS) and ISO assumes a reliable underlying transport protocol, such as TCP and is a binary message-oriented that provides message delivery guarantees for reliability, including at least once, at most once, and exactly once. This feature is very

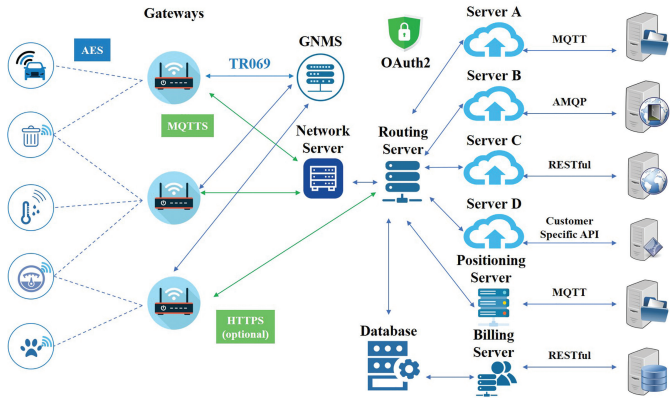

Figure 3.39 IoT applications implementation diagram. IoT application protocols deployment example.

important in the context of financial transactions (e.g., for executing credit or debit transactions). The protocol offers flow control through a token-based mechanism, to ensure that a receiving endpoint is not overburdened with more messages than it is capable of handling. AMQP. Different open-source implementations of the AMQP protocol are available.

AMQP supports both point-to-point communication and multipoint publish/subscribe interactions, defines a type system for encoding message data as well as annotating this data with additional context or metadata and the protocol can operate in simple peer-to-peer mode as well as in hierarchical architectures with intermediary nodes, e.g., messaging brokers or bridges.

The AMQP is an interoperable and cross platform messaging standard. In AMQP the messages along with a header are transmitted by the client to a broker or exchange and there is a single queue to which the message is transmitted by a producer. From the broker, the messages can be transmitted on to one or many queues. The AMQP header contains information about each byte of the message and the routing information. The broker is responsible to read headers and receive, route, and deliver messages to the client applications. The communication in AMQP protocol remains one to one between two nodes.

3.8.7.2 Constrained Application Protocol (CoAP)

CoAP is a RESTful protocol that supports the create, read, update, and delete (CRUD) verbs and in addition provides built-in support for the publish/subscribe paradigm via the new observe verb. CoAP provides a

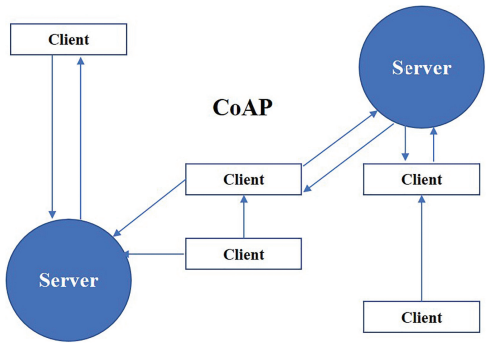

Figure 3.40 The flow of interactions for the CoAP.

mechanism where messages may be acknowledged for reliability and provides a bulk transfer mode.

Using CoAP, the client node can command another node by sending an CoAP packet that is interpreted by the CoAP server (the server may or may not acknowledge the request), which extracts the payload, and decides the action depending on its logic. The flow of interactions for the CoAP is presented in Figure 3.40.

CoAP is standardised as RFC 7252 by the IETF Constrained RESTful Environments (CORE) workgroup as a lightweight alternative to HTTP, targeted for constrained nodes in low-power and lossy networks (LLNs).

CoAP reduces the TCP overhead of seven messages required to fetch a resource by using UDP as a transport in lieu of TCP. CoAP uses short headers to reduce message sizes. IETF is further working to define mechanisms for dynamic resource discovery in CoAP via a directory service.

3.8.7.3 Distributed Data Service Real-Time Publish and Subscribe (DDS RTPS)

DDS RTPS is a data-centric application protocol using UDP as the underlying transport and standardised by Object Management Group (OMG).

The protocol supports the publish/subscribe paradigm, and which organises data into "topics" that listeners can subscribe to and receive asynchronous updates when the associated data changes. DDS RTPS supports different QoS policies for data distribution and provides mechanisms where listeners can automatically discover speakers associated with specific topics.

IP multicast or a centralised broker/server may be used to that effect. Several speakers may be associated with a single topic and priorities can be defined for different speakers, creating a redundancy mechanism for the

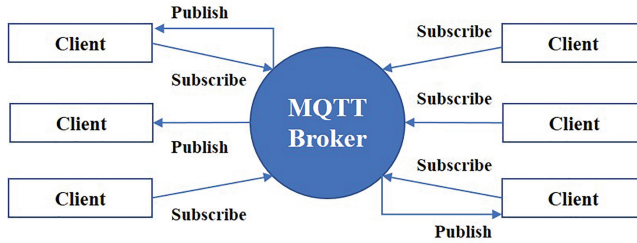

Figure 3.41 The flow of interactions for the MQTT.

architecture in case a speaker fails or loses communication with its listeners. The QoS policies include reliability, data persistence, delivery deadlines, and data freshness.

3.8.7.4 IEEE 1888

IEEE 1888, standardised by IEEE Standards Association is an application protocol for environmental monitoring, smart energy, and facility management applications, supporting reading and writing of time-series data using the Extensible Markup Language (XML) and the simple object access protocol (SOAP). The data is identified using Universal Resource Identifiers (URIs).

3.8.7.5 Message Queue Telemetry Transport (MQTT)

MQTT protocol is a binary protocol, using TCP as transport layer. The protocol was designed by IBM for enterprise telemetry and is a lightweight publish/subscribe messaging protocol standardised by OASIS. The protocol is message oriented, where messages are published to an address, referred to as a topic.

MQTT is a publish-subscribe protocol that facilitates one-to-many communication mediated by brokers with clients that can publish messages to a broker and/or subscribe to a broker to receive certain messages.

Messages are organised by topics, which essentially are "labels" that act as a system for dispatching messages to subscribers. The flow of interactions for the MQTT is presented in Figure 3.41.

Clients subscribe to one or more topics and receive updates from a client that is publishing messages for this topic. In MQTT, topics are hierarchical (like URLs), and subscriptions may use wildcards. MQTT uses a client-server architecture where clients connect to a central server (called the broker).

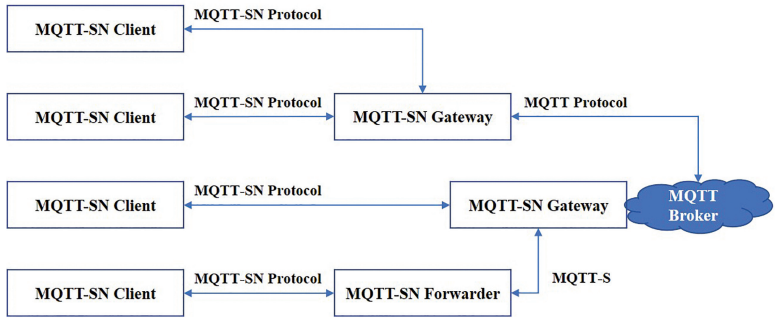

Figure 3.42 The architecture of MQTT-SN Protocol.

The protocol targets end devices where "a small code footprint" is required or where network bandwidth is limited (e.g. constrained IoT devices).

3.8.7.6 MQTT-Sensor Network (MQTT-SN)

The MQTT-SN is an open, lightweight publish/subscription protocol designed for constrained devices i.e. wireless sensor network (WSN). The protocol is based on MQTT and adapted to the specific requirements of wireless communication environments (e.g. short message length, low bandwidth, high link failures, etc.) and devices (e.g. low-cost, very low power consumption, long-battery life, limited processing and storage resources, etc.). Several changes have been introduced in MQTT-SN compared with MQTT, like the topic names are replaced by topic IDs, which reduce the overheads of transmission, topics do not need registration as they are preregistered, messages are split to send only the necessary information. The clients connect to the broker through a gateway device, which resides within the sensor network and connects to the broker and for reducing power consumption there is an offline procedure for clients who are in a sleep state so the messages can be buffered and later read by clients when they wake up.

Three components of MQTT-SN architecture are used such as MQTT-SN clients, MQTT-SN gateways (GW), and MQTT-SN forwarders as illustrated in Figure 3.42. MQTT-SN clients connect themselves to a MQTT server via a MQTT-SN GW using the MQTT-SN protocol, the MQTT-SN GW may or may not be integrated with a MQTT server. When utilising a stand-alone GW, the MQTT protocol is used between the MQTT server and the MQTT-SN GW having the function to translate the messages between MQTT and MQTT-SN.

MQTT-SN clients can access a GW via a forwarder in case the GW is not directly attached to their network. The forwarder encapsulates the MQTT-SN frames it receives on the wireless side and forwards them unchanged to the GW and releases the frames it receives from the gateway and sends them unchanged to the clients.

3.8.7.7 OMA LightweightM2M (LWM2M)

The LWM2M is a device management protocol, designed to be able to extend to meet the requirements of applications by transferring service / application data. The protocol uses a simple object-based resource model with resource operations of creation/retrieval/update/deletion/configuration of attribute. The protocol provides data format support for TLV, Json, Plain Text, Opaque and transport layer support for UDP/IP and SMS. The LWM2M has M2M/IoT functionalities such as LWM2M server, access control, device, connectivity, firmware update, location, connectivity statistics and uses DTLS for security.

3.8.7.8 RESTful HTTP (REST)

The concept of Representational State Transfer (REST) uses the HTTP methods such as GET, POST, PUT, and DELETE to provide a resource oriented messaging system where the interactions can be performed simply by using the synchronous request/response HTTP commands for receiving, modifying, and sending data. REST is based on TCP/IP, uses a request/response architecture with a relative complex implementation at client side, having a larger header compared to other IoT Protocols (e.g. higher bandwidth requirements) and applying SSL/TLS for security.

3.8.7.9 Secure Message Queue Telemetry Transport (SMQTT)

The SMQTT protocol is an encryption based light weight messaging protocol based on MQTT. Compared to an MQTT session, SMQTT session has four levels: setup, encryption, publish and decryption. The protocol uses a similar MQTT broker-based architecture with the difference that both the subscriber and publisher need to register with the broker using a secret master key. The data is encrypted before being published by the publisher and then is decrypted at the subscriber end and different encryption algorithms can be used by developers.

3.8.7.10 Session Initiation Protocol (SIP)

SIP is a text-based protocol that can use different underlying transports, TCP, UDP, or SCTP and is standardised by IETF as RFC 3261. The protocol handles session establishment for voice, video, and instant messaging applications on IP networks. It also manages presence (like XMPP). The invitation messages are used to create sessions carry session descriptions that enable edge devices to agree on a set of compatible media types. The protocol uses elements called proxy servers to route requests to the user's current location, authenticate and authorise users for services, implement call-routing policies, and provide features. A registration function is defined by the protocol to enable users to update their current locations for use by proxy servers.

3.8.7.11 Streaming Text Orientated Messaging Protocol (STOMP)

The STOMP is a text-based protocol for message-oriented middleware based on TCP and uses HTTP like commands. The protocol was designed to provide interoperability among platforms, languages, and brokers. The data is communicated between a client and broker in multi-line frames containing command, header, and content. The commands used can be CONNECT, DISCONNECT, ACK, NACK, SUBSCRIBE, UNSUBSCRIBE, SEND, BEGIN, COMMIT or ABORT.

3.8.7.12 Very Simple Control Protocol (VSCP)

VSCP is an open source standard protocol for M2M, IoT that enables low-cost devices to be networked together with computers and/or to operate as autonomous system, whatever the communication channels are used. The protocol utilises an event-based architecture, provides mechanisms for device discovery, identification, configuration and has support for secure device firmware updates. VSCP is an application level protocol that uses CAN, RS-232, Ethernet, TCP/IP, MQTT, 6LowPan or other protocols as it's transport mechanism and work over cable and over the air.

3.8.7.13 Extensible Messaging and Presence Protocol (XMPP)

XMPP is a message-centric protocol based on the Extensible Markup Language (XML) that use TCP as underlying transport with an option to run XMPP over HTTP. The protocol was designed for instant messaging, contact

list, and presence information maintenance and extended to several other applications, including network management, video, voice-over IP, file sharing, social networks, and online gaming, etc. XMPP is used for many IoT smart grid applications. The XMPP Standards Foundation (XSF) actively develops open extensions to the protocol.

An overview of selected IoT application protocols is presented in Table 3.7.

3.9 IoT Trustworthiness

The trustworthiness of IoT technologies and applications is critical to the acceptance and adoption of the technology, and many ethical issues must be addressed along the way to developing these technologies. The IoT applications are advancing and interact, cooperate, and collaborate with humans and animals. From the point of view of system design, the trustworthiness of IoT technologies are directly connected to the concept of dependability. Assuring dependability is ensuring the basis for trust in IoT technologies. The concept integrates the elements of availability, reliability, safety, security resilience, privacy and it embraces "privacy and security by design" as a model for an implementable IoT application.

The trustworthiness of IoT technologies and applications needs to consider trust semantics, metrics, models, IoT platforms, trusted IoT network computing, operating systems, software, and applications, while addressing the trust in mobile, wireless communications and risk and reputation management. Figure 3.43 illustrates the many channels for IoT applications security breaches. In this context, IoT applications need to embed mechanisms to continuously monitor security and stay ahead of the threats posed by interactions with other IoT applications and environments.

Trust is based on the ability to maintain the security of the IoT system and the ability to protect application/customer information, as well as being able to respond to unintended security or privacy breaches. In the IoT, it is important to drive security, privacy, data protection and trust across the whole IoT ecosystem.

3.9.1 Trust and Privacy in IoT Through Distributed Identity Management

IoT upcoming technologies rapidly thrust digitalisation into many application domains, such as smart cities, supply chain, industrial control, and healthcare

Table 3.7 Overview IoT application protocols

IoT Protocol	Functions	Transport	Format	Applications	SDO
AMQP	Message orientation, queuing, and pub/sub Data transfer with delivery guarantees (at least once, at most once, exactly once)	TCP	Binary	Financial services	OASIS
CoAP	REST resource manipulation via CRUD Resource tagging with attributes Resource discovery through RD	UDP	Binary	Low power and lossy networks	IETF
DDS (RTPS)	Pub/sub messaging with well-defined data types Data discovery Elaborate QoS	UDP	Binary	Real-time distributed systems (military, industrial, etc.)	OMG
IEEE 1888	Read/write data into URI Handling time-series data	SOAP/ HTTP	XML	Energy and facility management	IEEE
MQTT	Lightweight pub/sub messaging Message queuing for future subscribers	TCP	Binary	Enterprise telemetry	OASIS
SIP	Manage presence Session establishment Data transfer (voice, video, text)	TCP, UDP, SCTP	XML	IP telephony	IETF
XMPP	Manage presence Session establishment Data transfer (text or binary)	TCP HTTP	XML	Instant messaging	IETF XSF

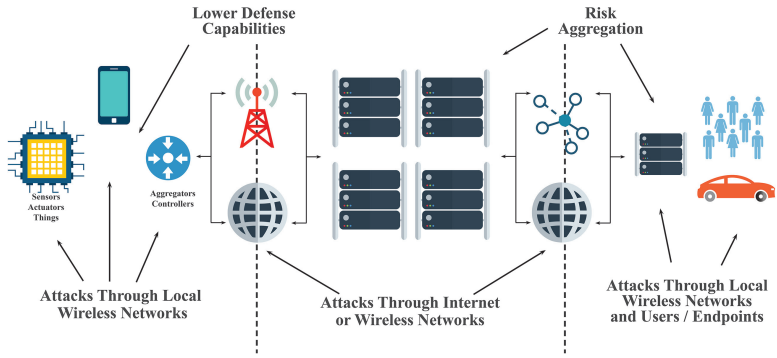

Figure 3.43 IoT security breaches through different channels.

systems. One of the most critical issues and challenges in IoT is to secure communication between IoT devices and internet.

The proliferation of connected IoT devices has led to an increased attack surface that leads to new cybersecurity risks that can compromise device security and data security. Smart healthcare medical systems, machine-to-machine communications (M2M), intelligent transportation systems (ITS), Industry 4.0 have heightened needs for trustworthy components with guaranteed authenticity, integrity, and confidentiality.

To address such security and privacy challenges, different authentication and privacy-preserving solutions are emerging to empower trusted IoT communications and enable anonymous mechanisms to protect sensitive and private data [191–193].

IoT is a service-providing infrastructure, where devices (things) can operate autonomously and need to be uniquely identifiable. Albeit cryptography is applied to protect communicated data when third parties are involved, this prerequisite mutual authentication between the communicating entities.

Maintaining a permanent unique identifier is a privacy risk, as the device can potentially be cloned, and IoT infrastructure maybe misused or wrong information can be provided to hide alerts or to get secret user information.

To guarantee security and privacy in IoT authentication protocols it is fundamentally required to guarantee mutual authentication, identity secrecy, device anonymity, non-traceability, forward security, availability [194].

To deal with these requirements, authentication protocols for IoT environments involve two or three phases and use various cryptographic mechanisms, including physical unclonable functions (PUF), digital signature, private-key cryptography, and public-key cryptography.

Generally, mutual authentication consists of two stages: the enrolment stage, where every node should be identified within the system, and the authentication stage, where a number of handshake messages are exchanged between the end node device and the server, which result is a session key to be used for upcoming communication [195, 196], and with provision of hardware crypto-cores [203].

By using three phases, registration, login and authentication phase, and password change phase, Chung et al. [197] can provide anonymity, hop-by-hop authentication, and non-traceability. Multiple trusted authority with role separation [205] have been proposed to authenticate communication in secure vehicular ad-hoc networks (VANETs) and through using identity-based aggregate signatures [206]. Recently, focusing on IoT setting, proposals involve Near Field Communications (NFC) secure element (SE)-based mutual authentication and attestation for IoT access with a user device, such as a smartphone [204].

Improved public-key-based mechanisms have been investigated, with most existing authentication protocols to be based on elliptic curve cryptography (ECC) to deal with the capabilities of constrained-resources devices compared to RSA algorithm.

Example proposals feature mutual authentication, non-traceability, and session key agreement [198]. Considering M2M communications in IIoT environment, lightweight authentication schemes are needed. Proposals involve a device equipped with a Secure Element (SE), which is authenticated by a network element equipped with a Trusted Platform Module (TPM) in two phases, registration and authentication [199].

Isolated execution environments through a proxy are a prominent solution in interconnected Cyber-Physical systems (CPSs), such as vehicles, given the penetration of the IoT paradigm in vehicles have raised the collection of a huge amount of data [207].

In IIoT, privacy-preserving biometric-based provable secure authentication protocol with ECC has also been proposed, in which a user authenticates himself to a gateway to agree on a session key for all future communications to be made secure and then accesses sensory data of a node [199].

In IIoT with a fog layer Attribute Credential-based Public Key Cryptography (AC-PKC) scheme [201] meets authentication and privacy-preserving access control requirements through employing a two level verification scheme which requires a fog node to generate a signature for the command it issues for an actuator.

The actuator on receiving this, must perform the two-level verification to authenticate that the command was indeed issued by a trusted fog node.

As IoT deployments mature, standardisation efforts have recognised and developed OAuth 2.0 authorisation framework [211] through using cryptographic tokens and authorisation and resource servers. The authorisation consent by the resource owner is provided after the owner is authenticated by an authorisation server; however, the authentication procedure is not part of OAuth 2.0. Authorisation is provided for different levels of access, such as read and write/modify, which are termed scopes, and for a specific time interval.

3.9.2 Decentralised Identification in IoT

The term decentralisation is recently often used to describe a security approach where data are spread over many network nodes in an approach to reduce chances of one point being vulnerable to introducing network threats or risks and anyhow limit or modify the network's or devices intended purpose.

Decentralisation includes aspects such as control level, access and ownership spread around the whole network nodes/actors that comprise it. Recently, this decentralisation refers to a shifting from centralised to distributed modes of network configurations. In mind of IoT devices identification in a network, decentralisation can play a predominant role as through this many issues of IoT ecosystems being brokered in a server/client approach. Up to now, this configuration seemed adequate and 'enough' but as the IoT devices' number and complexity increase, this is not considered as suitable anymore.

In the framework of IoT devices and solutions, core components of the IoT infrastructure and networks were cloud servers raising single point of failures largely in the IoT networks. In mind of critical applications (health, automotive, production etc.), decentralised approaches have a lot to offer regarding shared approaches.

Blockchain and other technologies recently support this decentralisation including concepts of trusting credentials in the network and validation of key modifications, self-sovereignty in identities involving self-trust on control of own identity and zero-knowledge proofs that actually proof identities and data without revealing any secret information. These are being discussed later in this chapter.

Always thinking of the IoT domain, the adoption of decentralised approaches for devices' identification is expected to strongly reduce

installation costs as well as costs for large (centralised) data centres or related infrastructures. This is also expected to support distribution of both storage and computational requirements while support single point of failures risks.

IoT devices to be used and managed when deployed in different applications require a unique identification. From simple device description to more sophisticated appliances as USB serial naming, security keys, cryptography keys need device identification.

Today device identification is done using the device's network address (IP, LPWAN, etc.) , hardware identifier as in the case of RFID, simple hardcoding it into the firmware, separately flashing info into the FLASH memory, or random generating by the first run of the IoT device..

Some microcontrollers such as STM32 provided a unique 96-bit ID encoded identifier that is generated during the manufacturing process. This unique Identifier consists of 3 parts: X and Y coordinates on the wafer expressed in BCD format, the lot number and wafer number. All these solutions are not 100% secure in all situations where it is important to prove the real identity of a device.

Several researches and trials are proposing to shift from a centralised server-client paradigm connecting to cloud via internet to architectures where industrial processes exchange information directly in a Peer-to-Peer (P2P) fashion (i.e., machine-to-machine (M2M) connections), hence promoting the decentralisation of computations across participating entities. Until today in security frameworks, digital identity is supported on centralised, commonly third-party, information systems which raise single point of failure concerns for an infrastructure.

Additionally, centralised identity management has privacy challenges with the possibility of digital identity exposure if the central authority is compromised. Furthermore, IoT devices with a hardware identifier can present privacy problems due to the identifier facilitating tracking and correlation attacks.

With the advent of blockchain, Identity Management (IdM) systems are switching from traditional web-centric approach or identity federation approaches, towards the self-sovereign identity (SSI) paradigm. Blockchain technology allows transferring digital assets (such as IoT data) in a decentralised fashion using the ledger, i.e., distributed ledger technology (DLT) [202], without intermediary central third-parties, while enabling public verifiability as well as provenance of the digital transactions and data. Meanwhile, cryptographic mechanisms (such as asymmetric encryption

algorithms, digital signature, and hash functions) guarantee the integrity of data blocks in the blockchains.

Therefore, a blockchain ensures non-repudiation of transactions, while each transaction in the blockchain is traceable to every user with the attached historic timestamp.

The ledger is immutable, meaning that past transactions cannot be modified by any entity registering transactions in the blockchain, and is shared and synchronised across all participating nodes. This way, the blockchain guarantees that the ledger cannot be tampered with, and that all the data held by the blockchain is trustworthy.

In IoT environment, millions of constrained smart entities with scarce capabilities to enforce proper security mechanisms, strive to cope with cyber-attacks that may leak their communicated data, and ultimately, sensitive, and private information of their owners/users.

Besides, in IoT, user privacy controls are difficult to apply, as the smart objects usually act on behalf of the user without user control and consent, undermining the adoption of the minimal personal disclosure principle. Blockchain brings a fully decentralised root of trust avoiding central authorities, thus creating trust across initially non-trusted or even unknown users and things [202, 208].

Blockchain-based identity schemes encrypt a user's identity, hash it and add its attestations to the blockchain ledger. These attestations are later used to prove the user's identity. Three important schemes and concepts determine the broader landscape of DLT-based IdMs:

- Decentralised Identity: This identity solution is like the conventional identity management solutions where credentials from a trusted service are used. The key modification is the storage of validated attestations on a distributed ledger for later validation by a third party.
- Self-sovereign identity: A user or device is the entity who owns and controls his/her identity without heavily relying on central authorities. It provides a framework to enable exchange of information and propagation of trust between peers.
- Zero knowledge proofs: Zero-knowledge protocols provide that an entity can prove the knowledge of its secret associated to public data without revealing any information about the secret. Essentially, one party (the prover) can prove to another party (the verifier) that they know a value x, without conveying any information apart from the fact that they know the value x. It is used in blockchain to perform authentication without giving the secret to other party.

A distributed identifier (DID) is a new type of identifier without a central issuing and controlling agency that creates and controls the identifier. Instead, DIDs are created and managed by the identity owner, an approach known as self-sovereign identity [210].

DID is acting as a permanent identifier which never changes and is resolvable. It acts as a persistent verifiable identifier through cryptography and can be used to encrypt communication channels for safe and secure messaging. A typical DID contains one or more public keys that can be used to authenticate the DID.

One or more services that can be used to for interaction via protocols supported by those services. Finally contains metadata such as digital signatures, timestamps, and other cryptographic proofs. An entity can have any different number of DIDs for many different purposes.

Current strengths and challenges of applying DLT to identity management together with the evaluation of three proposals (i.e., uPort, ShoCard, and Sovrin) are analysed in [202].

Currently, there are multiple competing DID technologies, in which the Decentralised Identity Foundation [209] is promoting interoperability among implementations. Today DIDs are an elegant solution in IoT to solve identity challenges, such as generated data insecurity, fraudulent identities, or third-party controlled IDs. Since DID are globally resolvable several IoT services can be enabled in a secure manner.

Hyperledger Indy is the only developed ledger to be permissioned. Even though anyone can read the blockchain's contents, only allowed entities can write to it. Since it is permissioned, it does not need proof-of-work to eliminate spam, making the transaction delay in the order of a few seconds (duration of network consensus).

To deploy DIDs and verifiable credentials in IoT, the IoT device should have:

- Sufficient performance for cryptographic operations,
- Enough energy to perform the required operations,
- Non-volatile storage space to store the code and cryptographic keys,
- Sufficient entropy source to generate random cryptographic keys.

Storing the keys on the device can present an unacceptable security risk of key leakage unless the device utilises a secure element, e.g., a trusted platform module (TPM), or embrace a proxy solution to act as a guardian for the keys.

In the scope of Self-Sovereign Identity model as shown in Figure 3.44, the blockchain acts as distributed and reliable identity verifier, providing

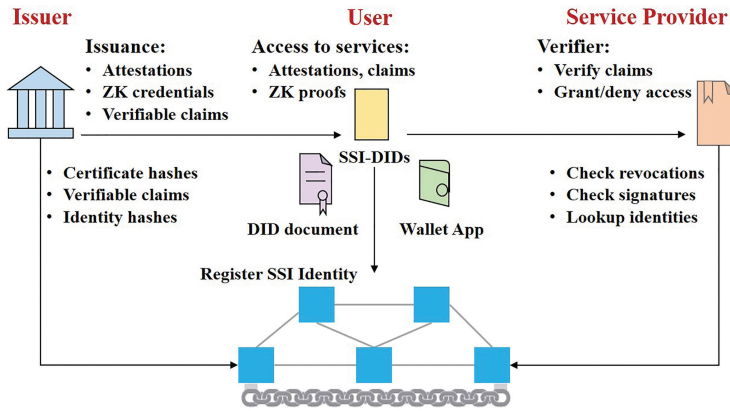

Figure 3.44 Blockchain-based self-sovereign identity management.

provenance and verifiability of identities. Thus, the ledger provides a cryptographic root of trust, which facilitates identity management without external authorities.

The major technical challenge that the SSI technology currently faces is its infancy and lack of widespread usage. It must be noted that privacy enhancing by using decentralised identifiers for the IoT devices is necessary for protecting the privacy of the users and owners of the devices, but similar care has to be taken with all other elements of the system to truly protect the privacy of users.

Additionally, in the direction of W3C community group [212], standard mechanisms must exist to define transparent interfaces and common data models that can be used in the exchange of information among parties to ensure that different entities (agents, devices, etc.) can interact seamlessly.

IoT in conjunction with blockchain, can assist to track, process and exchange transactions among connected devices. However, the development of new privacy-preserving approaches should manage the cost of cryptographic operations, to foster the adoption of blockchain technologies.

Today, lightweight scalable blockchain solutions can address the challenges of traditional security and privacy methods in IoT environment, centralisation, lack of privacy or safety threats. Nevertheless, careful infrastructure tuning should consider control of periodic integrity and authentication verification in the blockchain, to effectively prevent malicious nodes from intruding, to resist DDoS attacks and to prevent tampering with the device firmware.

3.10 Discussion

The rise of the IoT has brought countless applications and numerous opportunities, from the personal sensing applications, all-time connected sensors, smartphone, and wearable computers, to cloud and IoT and intelligent devices for robotic applications. Next wave of Internet of Intelligent Things is yet to come and has qualitatively augmented capacities to collect, manage and process data towards the personal-centric applications that are in high demand today and in the years to come.

Next-generation digital technologies fuelled by varieties of IoT, Industrial IoT, Tactile IoT, Internet of Robotic Things, Intelligent Internet of Things, Artificial Intelligence of Things, Internet of Things Senses, are developing to enable personalisation of services and interactions between things, humans and environments by unifying the physical, digital, virtual and cyberspaces into a continuum of experiences.

By bringing more intelligence to the IoT, a crucial characteristic will be provided to the future systems, which is the capacity for an awareness of the surrounding information through the integration of multiple pieces of information. We are yet to see the seamless integration of the Internet, intelligent things, and AI.

Acknowledgements

The IoT European Research Cluster – European Research Cluster on the Internet of Things (IERC) maintains its Strategic Research and Innovation Agenda (SRIA), considering its experiences and the results from the on-going exchange among European and international experts.

The present document builds on 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017 and 2018 Strategic Research and Innovation Agendas.

The IoT European Research Cluster SRIA is part of a continuous IoT community dialogue supported by the EC DG Connect – Communications Networks, Content and Technology, E4 – Internet of Things Unit for the European and international IoT stakeholders. The result is a lively document that is updated every year with expert feedback from on-going and future projects financed by the EC. Many colleagues have assisted over the last few years with their views on the IoT Strategic Research and Innovation agenda document. Their contributions are gratefully acknowledged.

Contributing Projects and Initiatives

SmartAgriFood, EAR-IT, ALMANAC, CITYPULSE, COSMOS, CLOUT, RERUM, SMARTIE, SMART-ACTION, SOCIOTAL, VITAL, BIG IoT, VICINITY, INTER-IoT, symbIoTe, TAGITSMART, bIoTope, AGILE, Be-IoT, UNIFY-IoT, ARMOUR, FIESTA, ACTIVAGE, AUTOPILOT, CREATE-IoT, IoF2020, MONICA, SYNCHRONICITY, U4IoT, BRAIN-IoT, ENACT, IoTCrawler, SecureIoT, SOFIE, CHARIOT, SEMIoTICS, SerIoT.

References

[1] OpenFog Reference Architecture for Fog Computing, OPFRA001. 020817, 2017.

[2] L. M. Vaquero and L. Rodero-Merino, "Finding your way in the fog: Towards a comprehensive definition of fog computing," SIGCOMM Comput. Commun. Rev., vol. 44, no. 5, pp. 27–32, Oct. 2014. Online at: http://doi.acm.org/10.1145/2677046.2677052

[3] R. van der Meulen, "Edge computing promises near real-time insights and facilitates localized actions", Web article, 2018. Online at: https://www.gartner.com/smarterwithgartner/what-edge-computing-means-for-infrastructure-and-operations-leaders/

[4] J. Morrish, M. Hatton and M. Arnott, "Global IoT Forecast Insight Report 2020", Transforma Insights, 2020. Online at: https://transformainsights.com/research/reports/global-iot-forecast-insight-report-2020

[5] ETSI White Paper, "Mobile Edge Computing A key technology towards 5G", #11, (09/2015), online at: https://www.etsi.org/images/files/ETSIWhitePapers/etsi_wp11_mec_a_key_technology_towards_5g.pdf

[6] ETSI, "Mobile Edge Computing (MEC); Framework and Reference Architecture", GS MEC 003 V1.1.1 (03/2016).

[7] Ahmed, E. Ahmed, "A Survey on Mobile Edge Computing", IEEE, Int'l Conf. on Intelligent System and Control ISCO 2016. Online at: https://www.researchgate.net/publication/285765997

[8] ETSI, "Multi-access Edge Computing (MEC); Framework and Reference Architecture", GS MEC 003, V2.1.1 (01/2019).

[9] ETSI White Paper, "MEC in 5G networks". #28, (06/2018).

[10] ETSI White Paper, "Network Transformation; (Orchestration, Network and Service Management Framework)", #32, (10/2019). Online

at: https://www.etsi.org/deliver/etsi_gs/MEC/001_099/003/02.01.01_60/gs_MEC003v020101p.pdf

[11] S. Ibrahim, H. Jin, B. Cheng, H. Cao, S. Wu, and L. Qi, "CLOUDLET: towards mapreduce implementation on virtual machines," in Proceedings of the 18th ACM International Symposium on High Performance Distributed Computing, HPDC 2009, Garching, Germany, June 11-13, 2009, 2009, pp. 65–66. Online at: https://dl.acm.org/doi/10.1145/1551609.1551624

[12] UN Environment Programme, "Emissions Gap Report 2019". Online at: https://wedocs.unep.org/bitstream/handle/20.500.11822/30797/EGR2019.pdf?sequence=1&isAllowed=y

[13] Sangwon Suh, et. al., United Nations Environment Programme, "Green Technology Choices: The Environmental and Resource Implications of Low-Carbon Technologies", 2017. Online at: https://www.resourcepanel.org/reports/green-technology-choices

[14] M. Satyanarayanan, P. Bahl, R. Caceres and N. Davies, "The Case for VM-Based Cloudlets in Mobile Computing," in IEEE Pervasive Computing, vol. 8, no. 4, pp. 14–23, Oct.–Dec. 2009. Online at: https://doi.org/10.1109/MPRV.2009.82

[15] European Commission, " The European Green Deal", COM(2019) 640 final, 2019. Online at: https://ec.europa.eu/info/sites/info/files/european-green-deal-communication_en.pdf

[16] United Nations, "Transforming our World: The 2030 Agenda for Sustainable Development" 2015. Online at: https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf

[17] European Commission, "Circular Economy Action Plan – For a cleaner and more competitive Europe", 2020. Online at: https://ec.europa.eu/jrc/communities/en/community/city-science-initiative/document/circular-economy-action-plan-cleaner-and-more-competitive0

[18] R. Arshad, S. Zahoor, M. A. Shah, A. Wahid and H. Yu, "Green IoT: An Investigation on Energy Saving Practices for 2020 and Beyond," in IEEE Access, vol. 5, pp. 15667–15681, 2017. Online at: https://doi.org/10.1109/ACCESS.2017.2686092

[19] N. Kaur and S. K. Sood, "An Energy-Efficient Architecture for the Internet of Things (IoT)," in IEEE Systems Journal, vol. 11, no. 2, pp. 796–805, June 2017, doi: 10.1109/JSYST.2015.2469676

[20] T. Qiu, A. Zhao, R. Ma, V. Chang, F. Liu, and Z. Fu. A task-efficient sink node based on embedded multi-core SoC for Internet of Things.

Future Generation Computer Systems, 82, pp. 656–666, 2018. https://doi.org/10.1016/j.future.2016.12.024

[21] S. Murugesan, "Harnessing Green IT: Principles and Practices," in IT Professional, vol. 10, no. 1, pp. 24–33, Jan.-Feb. 2008, doi: 10.1109/MITP.2008.10

[22] Elijah – Cloudlet-based Edge Computing. Online at: http://elijah.cs.cmu.edu/

[23] Open Edge Computing Initiative. Online at: https://www.openedgecomputing.org/

[24] OpenStack. Online at: https://www.openstack.org/

[25] M. Satyanarayanan, W. Gao and B, M Lucia, "The Computing Landscape of the 21st Century", HotMobile '19: Proceedings of the 20th International Workshop on Mobile Computing Systems and Applications, February 2019, pp. 45–50, Online at: https://dl.acm.org/doi/10.1145/3301293.3302357

[26] Y. Wang, "Definition and categorization of dew computing," Open Journal of Cloud Computing (OJCC), vol. 3, no. 1, pp. 1–7, 2016.

[27] M. Gushev, "Dew Computing Architecture for Cyber-Physical Systems and IoT", Internet of Things, Volume 11, September 2020. Online at: https://doi.org/10.1016/j.iot.2020.100186

[28] P. P. Ray, "An Introduction to Dew Computing: Definition, Concept and Implications," in IEEE Access, vol. 6, pp. 723–737, 2018. Online at: https://doi.org/10.1109/ACCESS.2017.2775042

[29] Tractica Report, "Artificial Intelligence for Edge Devices". Online at: https://www.tractica.com/research/artificial-intelligence-for-edge-devices/

[30] D. Schatsky, N. Kumar and S. Bumb, "Intelligent IoT – Bringing the power of AI to the Internet of Things", December 2017. Online at: https://www2.deloitte.com/insights/us/en/focus/signals-for-strategists/intelligent-iot-internet-of-things-artificial-intelligence.html#endnote-sup-1

[31] J. Fu, Y. Liu, H. Chao, B. K. Bhargava and Z. Zhang, "Secure Data Storage and Searching for Industrial IoT by Integrating Fog Computing and Cloud Computing," in IEEE Transactions on Industrial Informatics, vol. 14, no. 10, pp. 4519–4528, Oct. 2018. Online at: https://doi.org/10.1109/TII.2018.2793350

[32] M. Jbair, B. Ahmad, M. H. Ahmad and R. Harrison, "Industrial cyber physical systems: A survey for control-engineering tools," 2018 IEEE Industrial Cyber-Physical Systems (ICPS), St. Petersburg, 2018,

pp. 270–276. Online at: https://doi.org/10.1109/ICPHYS.2018.8387671

[33] M. Frey et al., "Security for the Industrial IoT: The Case for Information-Centric Networking," 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), Limerick, Ireland, 2019, pp. 424–429. Online at: https://doi.org/10.1109/WF-IoT.2019.8767183

[34] A. H. Sodhro, S. Pirbhulal and V. H. C. de Albuquerque, "Artificial Intelligence-Driven Mechanism for Edge Computing-Based Industrial Applications," in IEEE Transactions on Industrial Informatics, vol. 15, no. 7, pp. 4235–4243, July 2019. Online at: https://doi.org/10.1109/TII.2019.2902878

[35] R. Minerva, G. M. Lee and N. Crespi, "Digital Twin in the IoT Context: A Survey on Technical Features, Scenarios, and Architectural Models," in Proceedings of the IEEE. Online at: https://doi.org/10.1109/JPROC.2020.2998530

[36] E. Y. Song, M. Burns, A. Pandey and T. Roth, "IEEE 1451 Smart Sensor Digital Twin Federation for IoT/CPS Research," 2019 IEEE Sensors Applications Symposium (SAS), Sophia Antipolis, France, 2019, pp. 1–6. Online at: https://doi.org/10.1109/SAS.2019.8706111

[37] T. Catarci, D. Firmani, F. Leotta, F. Mandreoli, M. Mecella and F. Sapio, "A Conceptual Architecture and Model for Smart Manufacturing Relying on Service-Based Digital Twins," 2019 IEEE International Conference on Web Services (ICWS), Milan, Italy, 2019, pp. 229–236. Online at: https://doi.org/10.1109/ICWS.2019.00047

[38] Y. Lu, X. Huang, K. Zhang, S. Maharjan and Y. Zhang, "Communication-Efficient Federated Learning for Digital Twin Edge Networks in Industrial IoT," in IEEE Transactions on Industrial Informatics. Online at: https://doi.org/10.1109/TII.2020.3010798

[39] Z. Tsai, " The Emerging Role of AI in Edge Computing", November 2018. Online at: https://www.rtinsights.com/the-emerging-role-of-ai-in-edge-computing/

[40] M. Dsouza, "Intelligent Edge Analytics: 7 ways machine learning is driving edge computing adoption in 2018", August 2018 Online at: https://hub.packtpub.com/intelligent-edge-analytics-7-ways-machine-learning-is-driving-edge-computing-adoption-in-2018/

[41] STMicroelectronics, "Neural Networks on STM32", online at: https://www.st.com/content/st_com/en/about/innovation---technology/artificial-intelligence.html

[42] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang and W. Zhao, "A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications," in IEEE Internet of Things Journal, vol. 4, no. 5, pp. 1125–1142, Oct. 2017. Online at: https://doi.org/10.1109/JIOT.2017.2683200

[43] J. Ni, K. Zhang, X. Lin, X.S. Shen, "Securing fog computing for internet of things applications: challenges and solutions", IEEE Commun. Surv. Tutor. 20 (1) (2018) pp. 601–628. Online at: https://doi.org/10.1109/COMST.2017.2762345.

[44] A.C. Baktir, A. Ozgovde, C. Ersoy, "How can edge computing benefit from software-defined networking: a survey, use cases, and future directions", IEEE Commun. Surv. Tutor. 19 (4) (2017) pp. 2359–2391. Online at: https://doi.org/10.1109/COMST.2017.2717482

[45] M. Marjanović, A. Antonić and I. P. Žarko, "Edge Computing Architecture for Mobile Crowdsensing," in IEEE Access, vol. 6, pp. 10662–10674, 2018. Online at: https://doi.org/10.1109/ACCESS.2018.2799707

[46] A List of All Human Senses. Online at: https://www.scribd.com/document/251594575/A-List-of-All-Human-Senses

[47] J. Sachs et al., "Adaptive 5G Low-Latency Communication for Tactile Internet Services," in Proceedings of the IEEE, vol. 107, no. 2, pp. 325–349, Feb. 2019. https://doi.org/10.1109/JPROC.2018.2864587

[48] D. Van Den Berg et al., "Challenges in Haptic Communications Over the Tactile Internet," in IEEE Access, vol. 5, pp. 23502–23518, 2017. Online at: https://doi.org/10.1109/ACCESS.2017.2764181

[49] A. Aijaz, M. Dohler, A. H. Aghvami, V. Friderikos and M. Frodigh, "Realizing the Tactile Internet: Haptic Communications over Next Generation 5G Cellular Networks," in IEEE Wireless Communications, vol. 24, no. 2, pp. 82–89, April 2017. Online at: https://doi.org/10.1109/MWC.2016.1500157RP

[50] K. Antonakoglou, X. Xu, E. Steinbach, T. Mahmoodi and M. Dohler, "Toward Haptic Communications Over the 5G Tactile Internet," in IEEE Communications Surveys & Tutorials, vol. 20, no. 4, pp. 3034–3059, Fourthquarter 2018. Online at: https://doi.org/10.1109/COMST.2018.2851452

[51] S. K. Sharma, I. Woungang, A. Anpalagan and S. Chatzinotas, "Toward Tactile Internet in Beyond 5G Era: Recent Advances, Current

Issues, and Future Directions," in IEEE Access, vol. 8, pp. 56948–56991, 2020. Online at: https://doi.org/10.1109/ACCESS.2020.2980369

[52] S. Knab, and R.Rohrbeck. "Why intended business model innovation fails to deliver: insights from a longitudinal study in the German smart energy market." Proceedings of the R&D Management Conference, Stuttgart, Germany, June 3–6, 2014.

[53] O. Vermesan and J. Bacquet (Eds.). Next Generation Internet of Things. Distributed Intelligence at the Edge and Human Machine-to-Machine Cooperation, ISBN: 978-87-7022-008-8, River Publishers, Gistrup, 2018. Online at: https://european-iot-pilots.eu/wp-content/uploads/2018/11/Next_Generation_Internet_of_Things_Distributed_Intelligence_at_the_Edge_IERC_2018_Cluster_eBook_978-87-7022-007-1_P_Web.pdf

[54] O. Vermesan and J. Bacquet (Eds.). Cognitive Hyperconnected Digital Transformation. Internet of Things Intelligence Evolution, ISBN: 978-87-93609-11-2, River Publishers, Gistrup, 2017. Online at: https://www.riverpublishers.com/research_details.php?book_id=456

[55] A. Gluhak, O. Vermesan, R. Bahr, F. Clari, T. Macchia, M. T. Delgado, A. Hoeer, F. Boesenberg, M. Senigalliesi and V. Barchetti, "Report on IoT platform activities", 2016, online at http://www.internet-of-things-research.eu/pdf/D03_01_WP03_H2020_UNIFY-IoT_Final.pdf.

[56] IoT Analytics GmbH, IoT Platforms Company Landscape 2020. Online at: https://iot-analytics.com/product/iot-platforms-landscape-database-2020/

[57] M. Heller, "How to choose a cloud IoT platform". Online at: https://sg.channelasia.tech/article/print/679602/how-choose-cloud-iot-platform/

[58] O. Vermesan and P. Friess (Eds.). Digitising the Industry Internet of Things Connecting the Physical, Digital and Virtual Worlds, ISBN: 978-87-93379-81-7, River Publishers, Gistrup, 2016. Online at: https://www.riverpublishers.com/research_details.php?book_id=396

[59] O. Vermesan and P. Friess (Eds.). Building the Hyperconnected Society – IoT Research and Innovation Value Chains, Ecosystems and Markets, ISBN: 978-87-93237-99-5, River Publishers, Gistrup, 2015. Online at: https://www.riverpublishers.com/research_details.php?book_id=307

[60] S. Nakamoto. Bitcoin: A Peer-to-Peer Electronic Cash System. Online at: https://bitcoin.org/bitcoin.pdf

[61] Opportunities and Use Cases for Distributed Ledgers in IoT, GSMA 2018, online at: https://www.gsma.com/iot/wp-content/uploads/2018/09/Opportunities-and-Use-Cases-for-Distributed-Ledgers-in-IoT-f.pdf

[62] O. Vermesan, et. al., "The Next Generation Internet of Things – Hyperconnectivity and Embedded Intelligence at the Edge", in O. Vermesan and J. Bacquet (Eds.). Next Generation Internet of Things – Distributed Intelligence at the Edge and Human Machine-to-Machine Cooperation, ISBN: 978-87-7022-008-8, River Publishers, Gistrup, 2018, pp. 19–91.

[63] O. Vermesan, et. al., "Internet of robotic things: converging sensing/actuating, hypoconnectivity, artificial intelligence and IoT Platforms", in O. Vermesan and J. Bacquet (Eds.). Cognitive Hyperconnected Digital Transformation Internet of Things Intelligence Evolution, ISBN: 978-87-93609-10-5, River Publishers, Gistrup, 2017, pp. 97–155.

[64] What is Blockchain? Blockchainhub Berlin, updated July 2019. Online at: https://blockchainhub.net/blockchain-intro/

[65] K. Loupos, B. Caglayan, A. Papageorgiou, B. Starynkevitch, F. Vedrine, C. Skoufis, S. Christofi, B. Karakostas, A. Mygiakis, G. Theofilis, A. Chiappetta, H. Avgoustidis, George Boulougouris – Cognition Enabled IoT Platform for Industrial IoT Safety, Security and Privacy – The CHARIOT Project, IEEE International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), 11–13 September 2019, Limassol, Cyprus, DOI: 10.1109/CAMAD.2019.8858488.

[66] Papageorgiou, T. Krousarlis, K. Loupos, A. Mygiakis, DPKI: A Blockchain-Based Decentralized Public Key Infrastructure System, Global IoT Summit 2020, 3rd Workshop on Internet of Things Security and Privacy (WISP), 3–5 June 2020, Dublin.

[67] M. Swan, Blockchain: Blueprint for a New Economy, O'Reilly Media, Inc., 2015.

[68] K. Christidis, M. Devetsikiotis, Blockchains and smart contracts for the internet of things. IEEE Access 4, pp. 2292–2303, 2016.

[69] What is a blockchain operating system and what are the benefits? Introducing Overledger from Quant Network. Online at: https://medium.com/@CryptoSeq/what-is-a-blockchain-operating-system-and-what-are-the-benefits-c561d8275de6

[70] D. Yaga, P. Mell, N. Roby, and K. Scarfone. Blockchain Technology Overview. Technical Report. NISTIR. 2018. Online at: https://doi.org/02

[71] Outlier Ventures Research, Blockchain-Enabled Convergence – Understanding The Web 3.0 Economy, online at https://gallery.mailchimp.com/65ae955d98e06dbd6fc737bf7/files/Blockchain_Enabled_Convergence.01.pdf

[72] What is a blockchain? https://www2.deloitte.com/content/dam/Deloitte/ch/Documents/innovation/ch-en-innovation-deloitte-what-is-blockchain-2016.pdf

[73] LongView Whitepaper, Super-B for IoT: Improving QoS in LoRa Networks, 2019. Online at: https://www.longviewiot.com/resources/assets/white-paper-super-b-protocol-pdf/

[74] M. Wright, 5G – A brave new future. Telestra. Presentation at Mobile World Congress, Barcelona, 2018.

[75] S. Lin, H.F. Cheng, W. Li, Z. Huang, P. Hui, C. Peylo, Ubii: physical world interaction through augmented reality, IEEE Trans. Mob. Comput. 16 (2017) 872–885.

[76] X. Sun, N. Ansari, EdgeIoT: mobile edge computing for the Internet of Things, IEEE Commun. Mag. 54 (2016) 22–29.

[77] Edge Computing Market, MarketsandMarkets Report 2017, online at: https://www.marketsandmarkets.com/Market-Reports/edge-computing-market-133384090.html

[78] Ericsson Mobility Report, June 2018, online at: https://www.ericsson.com/assets/local/mobility-report/documents/2018/ericsson-mobility-report-june-2018.pdf

[79] Ericsson Mobility Report, June 2019, online at: https://www.ericsson.com/49d1d9/assets/local/mobility-report/documents/2019/ericsson-mobility-report-june-2019.pdf

[80] Ericsson Mobility Report, November 2019, online at: https://www.ericsson.com/4acd7e/assets/local/mobility-report/documents/2019/emr-november-2019.pdf?_ga=2.258613962.1969153473.1594892941--1856930902.1594892941&_gac=1.212672800.1594892967.EAIaIQobChMIkb6xoL_R6gIVS9OyCh3EWA8tEAAYASAAEgJNYPD_BwE

[81] Ericsson ConsumerLab, “10 Hot Consumer Trends 2030 – Internet of the senses”, December 2019, Online at:

[82] Ericsson Mobility Report, June 2020, online at: https://www.ericsson.com/49da93/assets/local/mobility-report/documents/2020/june2020-ericsson-mobility-report.pdf

[83] Cloud IoT Edge, online at: https://cloud.google.com/iot-edge/
[84] Hyperledger, online at: https://www.hyperledger.org/
[85] Enterprise Ethereum Alliance (EEA), online at: https://entethalliance.org/
[86] Hyperledger Fabric, online at: https://www.ibm.com/blockchain/hyperledger/fabric-support
[87] G. Herr, J. Lyon, and S. Gillen, "Industrial intelligence: Cognitive analytics in action," presentation at EMEA Users Conference, Berlin, 2016.
[88] Research and Markets, online at: https://www.researchandmarkets.com/
[89] State of the IoT 2018: Number of IoT devices now at 7B – Market accelerating, IOT ANALYTICS, online at: https://iot-analytics.com/state-of-the-iot-update-q1-q2--2018-number-of-iot-devices-now-7b/
[90] IDTechEx, Comparison of Low Power Wide Area Networks (LPWAN) for IoT 2018–2019, online at https://www.idtechex.com/research/articles/comparison-of-low-power-wide-area-networks-lpwan-for-iot-2018--2019-00014777.asp
[91] Wi-Fi HaLow, online at: https://www.wi-fi.org/discover-wi-fi/wi-fi-halow
[92] H. Haas, "LiFi is a paradigm-shifting 5G technology", Reviews in Physics 3 (2018), pp. 26–31.
[93] Wi-SUN FAN Overview, IETF LPWAN working group, online at: https://tools.ietf.org/id/draft-heile-lpwan-wisun-overview-00.html
[94] SEMTECH, online at: https://www.semtech.com/
[95] T. Ryberg, BERG INSIGHT, NB-IoT networks are here, now it's time to make business, online at: https://www.iot-now.com/2018/07/04/85156-nb-iot-networks-now-time-make-business/
[96] NB-IoT, CAT-M, SIGFOX and LoRa Battle for Dominance Drives Global LPWA Network Connections to Pass 1 Billion By 2023, ABIresearch, 2018, online at: https://www.abiresearch.com/press/nb-iot-cat-m-sigfox-and-lora-battle-dominance-drives-global-lpwa-network-connections-pass-1-billion-2023/
[97] Wi-SUN FAN Overview, online at: https://tools.ietf.org/id/draft-heile-lpwan-wisun-overview-00.html
[98] Opportunities and Use Cases for Distributed Ledgers in IoT, GSMA 2018, online at: https://www.gsma.com/iot/wp-content/uploads/2018/09/Opportunities-and-Use-Cases-for-Distributed-Ledgers-in-IoT-f.pdf

[99] BEHRTECH, "The Ultimate Guide io Wireless Connectivity for Massive Scale IoT Deployments", online at: www.behrtech.com

[100] Lavric, "LoRa High-Density Sensors for Internet of Things", Journal of Sensors, 2019.

[101] ETSI – European Telecommunications Standards Institute, "Short range devices; Low Throughput Networks (LTN); Protocols for radio interface A", ETSI TS 103 357, 2018.

[102] T. Lauterbach, "MYTHINGS vs LoRa: A comparative study of Quality-of-Service under external interference", 2019, online at: https://behrtech.com/resources/lora-vs-mythings/

[103] MIOTY, Radiocrafts AS, online at: https://radiocrafts.com/products/mioty-network/

[104] mioty® – The Wireless IoT Technology; Fraunhofer IIS, online: http://mioty.de/

[105] ANT/ANT+ Defined, online at: https://www.thisisant.com/developer/ant-plus/ant-antplus-defined/

[106] "ANT message protocol and usage: Application notes," Dynastream Innovations Inc., 2007, online at: https://www.thisisant.com/resources/ant-message-protocol-and-usage/

[107] S. Khssibi, H. Idoudi, A. V. D. Bossche, T. Val, and L. A. Saidane, "Presentation and analysis of a new technology for low-power wireless sensor network," International Journal of Digital Information and Wireless Communications (IJDIWC), vol. 3, no. 1, pp. 75–86, 2013, online at: https://oatao.univ-toulouse.fr/12426/1/Khssibi_12426.pdf

[108] "AN1200.22 LoRa^{TM} Modulation Basics", online at: https://semtech.my.salesforce.com/sfc/p/#E0000000JelG/a/2R0000001OJu/xvKUc5w9yjG1q5Pb2IIkpolW54YYqGb.frOZ7HQBcRc

[109] P. Pachuca, "LoRa Edge^{TM} Explained: How LR1110 Drives Smarter Geolocation", Semtech's Corporate Blog, 2020, online at: https://blog.semtech.com/lora-edge-explained

[110] F. Adelantado, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melia-Segui and T. Watteyne, "Understanding the Limits of LoRaWAN," in IEEE Communications Magazine, vol. 55, no. 9, pp. 34–40, Sept. 2017, online at: doi: 10.1109/MCOM.2017.1600613.

[111] N. Ducrot, D. Ray, A. Saadani, O. Hersent, G. Pop, and G. Remond, "LoRa Device Developer Guide," Orange, Connected Objects and Partnership. Technical Document., Apr. 2016, online at: https://developer.orange.com/wp-content/uploads/LoRa-Device-Developer-Guide-Orange.pdf

[112] Minaburo, A. Pelov, and L. Toutain, LP-WAN Gap Analysis, IETF Std., Feb. 2016, online at: https://tools.ietf.org/pdf/draft-minaburo-lp-wan-gap-analysis-00.pdf

[113] Z. Yoon, W. Frese, and K. Briess. "Design and Implementation of a Narrow-Band Intersatellite Network with Limited Onboard Resources for IoT". Sensors (Basel). 2019 Sep;19(19), online at: doi:10.3390/s19194212. PMID: 31569831; PMCID: PMC6806246.

[114] H. Urlings, "Satellite IoT: A Game Changer for the Industry?", 2019, online at: http://satellitemarkets.com/satellite-iot-game-changer-industry

[115] W. Webb, "Weightless: A Bespoke Technology for the IoT", online: http://www.weightless.org/news/weightless-a-bespoke-technology-for-the-iot

[116] Ingenu. How RPMA Works: The Making of RPMA, online at: https://www.ingenu.com/portfolio/how-rpma-works-the-making-of-rpma/

[117] Sigfox Technology, online at: https://www.sigfox.com/en/what-sigfox/technology

[118] Weightless Specification, online at: http://www.weightless.org/about/weightless-specification

[119] Gupta, R.K. Jha, A survey of 5G network: architecture and emerging technologies, IEEE Access 3 (July) (2015) 1206–1232.

[120] M. Iwamura, "NGMN View on 5G architecture", Proceedings of the IEEE Vehicular Technology Conference, 2015, pp. 1–5.

[121] NGMN 5G Initiative White Paper, February 2015, online at: https://www.ngmn.org/wp-content/uploads/NGMN_5G_White_Paper_V1_0.pdf

[122] Afolabi, A. Ksentini , M. Bagaa , T. Taleb , M. Corici , A. Nakao ,Towards 5G network slicing over multiple-Domains, in IEICE Trans. Commun. (11) (2017) 1992–2006.

[123] ITU-T Y.3011, Framework of Network Virtualization for Future Networks, January 2012, online at: https://www.itu.int/rec/T-REC-Y.3011-201201-ITRGPP23.799, Study on Architecture for Next Generation System, Rel.14 (December 2016), online at: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3008

[124] X. de Foy, A. Rahman , "Network slicing-3GPP Use case", InterDigital Communications, LLC, 2017, online at: https://tools.ietf.org/pdf/draft-defoy-netslices-3gpp-network-slicing-02.pdf

[125] J. Crawshaw, "Network Slicing: OSS/BSS Key to Commercial Success", Tech Mahindra, 2019 online at: https://cache.techmahindra.com/static/img/pdf/oss-bss-key-to-commercial-success.pdf
[126] Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, M. Ayyash, Internet of Things: a survey on enabling technologies, protocols, and applications, IEEE Commun. Surv. Tutorials 17 (June (4)) (2015) 2347–2376.
[127] M.R. Palattella, M. Dohler, A. Grieco, G. Rizzo, J. Torsner , T. Engel, L. Ladid , Internet of Things in the 5G era: enablers, architecture, and business models, IEEE J. Sel. Areas Commun. 34 (February (3)) (2016) 510–527.
[128] M. Hasan, E. Hossain, D. Niyato, Random access for machine-to-machine communication in LTE-advanced networks: issues and approaches, IEEE Commun. Mag. 51 (June (6)) (2013) 86–93.
[129] M. Maier, M. Chowdhury, B.P. Rimal, D.P. Van, The tactile internet: vision, recent progress, and open challenges, IEEE Commun. Mag. 54 (May (5)) (2016) 138–145.
[130] *Progress on 3GPP IoT*, February 2016, [Online]. Available: http://www.3gpp.org/news-events/3gpp-news/1766-iot_progress
[131] *Release 14*, February 2016, [Online]. Available: http://www.3gpp.org/release-14.
[132] S. Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, online at: https://bitcoin.org/bitcoin.pdf
[133] S. Popov, The Tangle, 2018, online at: https://assets.ctfassets.net/r1dr6vzfxhev/2t4uxvsIqk0EUau6g2sw0g/45eae33637ca92f85dd9f4a3a218e1ec/iota1_4_3.pdf
[134] IOTA, online at: https://iota.org
[135] Ripple, online at: https://ripple.com
[136] Sovrin, online at: https://sovrin.org
[137] BigchainDB, online at: https://www.bigchaindb.com
[138] Digital Twin consortium. Online at: https://www.digitaltwinconsortium.org/index.htm
[139] R., Stark, and T., Damerau, Digital Twin. Springer Berlin Heidelberg, Berlin, Heidelberg, 2019, pp. 1–8. Online at: http://dx.doi.org/10.1007/978-3-642-35950-7_16870-1.
[140] Digital Twin towards a meaningful framework-Arup. Online at: https://www.arup.com/perspectives/publications/research/section/digital-twin-towards-a-meaningful-framework
[141] Industrial Data. Online at: https://www.iso.org/committee/54158.html

[142] Digital Twin Manufacturing Framework. Online at: https://www.iso.org/standard/75066.html
[143] JETI. Online at: https://jtc1info.org/technology/jeti/
[144] ISO/TC 184 – Automation systems and integration. Online at: https://www.iso.org/committee/54110.html
[145] P2806 – System Architecture of Digital Representation for Physical Objects in Factory Environments. Online at: https://standards.ieee.org/project/2806.html#Standard
[146] Requirements and capabilities of a digital twin system for smart cities. ITU-T work programme. Online at: https://www.itu.int/ITU-T/workprog/wp_item.aspx?isn=16396
[147] buildingSMART International, Position Paper, "Enabling an Ecosystem of Digital Twins". Online at: https://www.buildingsmart.org/wp-content/uploads/2020/05/Enabling-Digital-Twins-Positioning-Paper-Final.pdf
[148] IoTwins H2020 project. Online at: https://www.iotwins.eu/
[149] Virtual Singapore. National Research Foundation (NRF), Singapore. Online at: https://www.nrf.gov.sg/programmes/virtual-singapore
[150] University of Cambridge. UK. National Digital Twin Programme. Online at: https://www.cdbb.cam.ac.uk/what-we-do/national-digital-twin-programme
[151] J Konečný, H.B., McMahan, F.X., Yu, P., Richtárik, A.T., Suresh, D., Bacon. "Federated learning: Strategies for improving communication efficiency". Online at: arXiv preprint arXiv:1610.05492
[152] OpenMined – open-source community. Online at: https://www.openmined.org/
[153] P., Kairouz, H. B., McMahan, et. al., "Advances and Open Problems in Federated Learning," Dec. 2019. Online at: https://arxiv.org/abs/1912.0497
[154] K. Panetta, 5 Trends Emerge in the Gartner Hype Cycle for Emerging Technologies, 2018, online at: https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/
[155] N. J. Nilsson, The Quest for Artificial Intelligence: A History of Ideas and Achievements, Cambridge, UK Cambridge University Press, 2010.
[156] IEEE, Tactile Internet Emerging Technologies Subcommittee, online at: http://ti.committees.comsoc.org/

[157] The Tactile Internet ITU-T Technology Watch Report ITU-T, 2014, online at: https://www.itu.int/dms_pub/itu-t/oth/23/01/T23010000230001PDFE.pdf

[158] G. Batra, A. Queirolo, and N. Santhanam, McKinsey & Company, 2018, Artificial intelligence: The time to act is now, online at: https://www.mckinsey.com/industries/advanced-electronics/our-insights/artificial-intelligence-the-time-to-act-is-now

[159] 5G Network Architecture A High-Level Perspective, White Paper, HUAWEI Technologies CO., LTD., 2016, online at: https://www-file.huawei.com/-/media/CORPORATE/PDF/mbb/5g_nework_architecture_whitepaper_en.pdf?la=en&source=corp_comm

[160] 5G Security Architecture White Paper, HUAWEI Technologies CO., LTD., 2017, online at: https://www-file.huawei.com/-/media/CORPORATE/PDF/white%20paper/5g_security_architecture_white_paper_en-v2.pdf?la=en&source=corp_comm

[161] Network Slicing Use Case Requirements, GSMA, April 2018, online at: https://www.gsma.com/futurenetworks/wp-content/uploads/2018/07/Network-Slicing-Use-Case-Requirements-fixed.pdf

[162] O. Holland et al., "The IEEE 1918.1 "Tactile Internet" Standards Working Group and its Standards," in Proceedings of the IEEE, vol. 107, no. 2, pp. 256–279, Feb. 2019. Online at: https://doi.org/10.1109/JPROC.2018.2885541

[163] Z. S. Bojkovic, B. M. Bakmaz and M. R. Bakmaz, "Vision and enabling technologies of tactile internet realization," 2017 13th International Conference on Advanced Technologies, Systems and Services in Telecommunications (TELSIKS), Nis, 2017, pp. 113–118. Online at: https://doi.org/10.1109/TELSKS.2017.8246242

[164] L. Zou, *et al.,* "Novel tactile sensor technology and smart tactile sensing systems: A review," *MDPI Sensors*, vol. 17, 2017.

[165] F. Dressler, "Towards the Tactile Internet: Low Latency Communication for Connected Cars," online at: http://conferences.sigcomm.org/sigcomm/2017/files/tutorial-c2c.pdf

[166] W. E. Forum, "Industrial internet of things: Unleashing the potential of connected products and services," Jan. 2015, World Economic Forum, Geneva, Switzerland, White Paper REF 020315.

[167] T. H. Szymanski, "Securing the Industrial-Tactile Internet of Things with Deterministic Silicon Photonics Switches," in IEEE Access, vol. 4, pp. 8236–8249, 2016. Online at: https://doi.org/10.1109/ACCESS.2016.2613512

[168] M. Maier, M. Chowdhury, B. P. Rimal and D. P. Van, "The tactile internet: vision, recent progress, and open challenges," in IEEE Communications Magazine, vol. 54, no. 5, pp. 138–145, May 2016. Online at: https://doi.org/10.1109/MCOM.2016.7470948

[169] M. Maier, "FiWi access networks: Future research challenges and moonshot perspectives," 2014 IEEE International Conference on Communications Workshops (ICC), Sydney, NSW, 2014, pp. 371–375. Online at: https://doi.org/10.1109/ICCW.2014.6881225

[170] M. Chowdhury and M. Maier, "Collaborative Computing for Advanced Tactile Internet Human-to-Robot (H2R) Communications in Integrated FiWi Multirobot Infrastructures," in IEEE Internet of Things Journal, vol. 4, no. 6, pp. 2142–2158, Dec. 2017. Online at: https://doi.org/10.1109/JIOT.2017.2761599

[171] Spectrum of Seven Outcomes for AI, online at: https://www.constellationr.com/

[172] ITU-T, Internet of Things Global Standards Initiative, http://www.itu.int/en/ITU-T/gsi/iot/Pages/default.aspx

[173] International Telecommunication Union – ITU-T Y.2060 – (06/2012) – Next Generation Networks – Frameworks and functional architecture models – Overview of the Internet of things

[174] O. Vermesan, P. Friess, P. Guillemin, H. Sundmaeker, et al., "Internet of Things Strategic Research and Innovation Agenda", Chapter 2 in Internet of Things – Converging Technologies for Smart Environments and Integrated Ecosystems, River Publishers, 2013, ISBN 978-87-92982-73-5

[175] Parks Associates, Monthly Wi-Fi usage increased by 40% in U.S. smartphone households, online at https://www.parksassociates.com/blog/article/pr-06192017

[176] Gluhak, O. Vermesan, R. Bahr, F. Clari, T. Macchia, M. T. Delgado, A. Hoeer, F. Boesenberg, M. Senigalliesi and V. Barchetti, "Report on IoT platform activities", 2016, online at http://www.internet-of-things-research.eu/pdf/D03_01_WP03_H2020_UNIFY-IoT_Final.pdf.

[177] IoT Platforms Initiative, online at https://www.iot-epi.eu/

[178] IoT European Large-Scale Pilots Programme, online at https://european-iot-pilots.eu/

[179] S. Moore, (2016, December 7) Gartner Survey Shows Wearable Devices Need to Be More Useful, online at http://www.gartner.com/newsroom/id/3537117

[180] Digital Economy Collaboration Group (ODEC). Online at http://archive.oii.ox.ac.uk/odec/
[181] Connect building systems to the IoT, online at http://www.electronics-know-how.com/article/1985/connect-building-systems-to-the-iot
[182] S. Kejriwal and S. Mahajan, Smart buildings: How IoT technology aims to add value for real estate companies The Internet of Things in the CRE industry, Deloitte University Press, 2016, online at https://www2.deloitte.com/content/dam/Deloitte/nl/Documents/real-estate/deloitte-nl-fsi-real-estate-smart-buildings-how-iot-technology-aims-to-add-value-for-real-estate-companies.pdf
[183] ORGALIME Position Paper, 2016, online at http://www.orgalime.org/sites/default/files/position-papers/Orgalime%20Comments_EED_EPBD_Review%20Policy%20Options_4%20May%202016.pdf
[184] The EU General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679).
[185] RERUM, EU FP7 project, www.ict-rerum.eu
[186] Digital Agenda for Europe, European Commission, Digital Agenda 2010–2020 for Europe. Online at http://ec.europa.eu/information_society/digital-agenda/index_en.html
[187] O. Vermesan, P. Friess, G. Woysch, P. Guillemin, S. Gusmeroli, et al., " Europe's IoT Strategic Research Agenda 2012", Chapter 2 in The Internet of Things 2012 New Horizons, Halifax, UK, 2012, ISBN 978-0-9553707-9-3
[188] O. Vermesan, et al., "Internet of Energy – Connecting Energy Anywhere Anytime" in Advanced Microsystems for Automotive Applications 2011: Smart Systems for Electric, Safe and Networked Mobility, Springer, Berlin, 2011, ISBN 978-36-42213-80-9
[189] M. Yuriyama and T. Kushida, "Sensor-Cloud Infrastructure – Physical Sensor Management with Virtualized Sensors on Cloud Computing", NBiS 2010: 1–8
[190] Test Considerations for 5G New Radio, White Paper, Keysight Technologies, April 2018, online at: http://literature.cdn.keysight.com/litweb/pdf/5992--2921EN.pdf
[191] M. A. Ferrag, L. A. Maglaras, H. Janicke, J. Jiang, and L. Shu, "Authentication Protocols for Internet of Things: A Comprehensive Survey", in Security and Communication Networks, Hindawi, 2017, doi: 10.1155/2017/6562953
[192] V. Hassija, V. Chamola, V. Saxena, D. Jain, P. Goyal and B. Sikdar, "A Survey on IoT Security: Application Areas, Security Threats, and

Solution Architectures," in *IEEE Access*, vol. 7, pp. 82721–82743, 2019.

[193] M. Banerjee, J. Lee, K.-K. R. Choo, "A blockchain future for internet of things security: a position paper", in Digital Communications and Networks, Vol. 4, Issue 3, 2018, pp. 149–160, doi:10.1016/j.dcan.2017.10.006

[194] N. Chikouche, P.-L. Cayrel, El H. M. Mboup, B. O. Boidje, "A privacy-preserving code-based authentication protocol for Internet of Things", in Journal of Supercomputing, Springer Verlag, 2019, doi:10.1007/s11227-019-03003-4

[195] Schmitt, C.; Noack, M.; Stiller, B. TinyTO: Two-way authentication for constrained devices in the Internet of Things. In Internet of Things; Elsevier: Amsterdam, The Netherlands, 2016; pp. 239–258.

[196] M. Jan, P. Nanda, M. Usman, and X. He, "PAWN: A payload-based mutual authentication scheme for wireless sensor networks," Concurrency Computation, 2016.

[197] Y. Chung, S. Choi, Y. S. Lee, N. Park, and D. Won, "An enhanced lightweight anonymous authentication scheme for a scalable localization roaming service in wireless sensor networks," Sensors, vol. 16, no. 10, article no. 1653, 2016.

[198] Maarof, M. Senhadji, Z. Labbi, M. Belkasmi, "Authentication protocol for securing internet of things", In Proceedings of the Fourth International Conference on Engineering & MIS 2018. ACM, pp 29:1–29:7

[199] A. Esfahani et al., "A Lightweight Authentication Mechanism for M2M Communications in Industrial IoT Environment," in IEEE Internet of Things Journal, vol. 6, no. 1, pp. 288–296, Feb. 2019. Online at: https://doi.org/10.1109/JIOT.2017.2737630

[200] X. Li, J. Niu, M. Z. A. Bhuiyan, F. Wu, M. Karuppiah and S. Kumari, "A Robust ECC-Based Provable Secure Authentication Protocol With Privacy Preserving for Industrial Internet of Things," in IEEE Transactions on Industrial Informatics, vol. 14, no. 8, pp. 3599–3609, Aug. 2018. Online at: https://doi.org/10.1109/TII.2017.2773666

[201] X. Yao, H. Kong, H. Liu, T. Qiu and H. Ning, "An Attribute Credential Based Public Key Scheme for Fog Computing in Digital Manufacturing," in IEEE Transactions on Industrial Informatics, vol. 15, no. 4, pp. 2297–2307, April 2019. Online at: https://doi.org/10.1109/TII.2019.2891079

[202] P. Dunphy and F. A. P. Petitcolas, "A First Look at Identity Management Schemes on the Blockchain," in IEEE Security & Privacy, vol. 16, no. 4, pp. 20–29, July/August 2018. Online at: https://doi.org/10.1109/MSP.2018.3111247

[203] G. Kornaros, O. Tomoutzoglou and M. Coppola, "Hardware-Assisted Security in Electronic Control Units: Secure Automotive Communications by Utilizing One-Time-Programmable Network on Chip and Firewalls," in IEEE Micro, vol. 38, no. 5, pp. 63–74, Sep./Oct. 2018. Online at: https://doi.org/10.1109/MM.2018.053631143

[204] D. Sethia, D. Gupta and H. Saran, "NFC Secure Element-Based Mutual Authentication and Attestation for IoT Access," in IEEE Transactions on Consumer Electronics, vol. 64, no. 4, pp. 470–479, Nov. 2018. Online at: https://doi.org/10.1109/TCE.2018.2873181

[205] D. Mbakoyiannis, O. Tomoutzoglou, and G. Kornaros, "Secure over-the-air firmware updating for automotive electronic control units", In Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing (SAC '19), pp. 174–181, 2019. https://doi.org/10.1145/3297280.3297299

[206] L. Zhang, C. Hu, Q. Wu, J. Domingo-Ferrer and B. Qin, "Privacy-Preserving Vehicular Communication Authentication with Hierarchical Aggregation and Fast Response," in IEEE Transactions on Computers, vol. 65, no. 8, pp. 2562–2574, 1 Aug. 2016. Online at: https://doi.org/10.1109/TC.2015.2485225

[207] G. Kornaros, et al., "Towards Holistic Secure Networking in Connected Vehicles through Securing CAN-bus Communication and Firmware-over-the-Air Updating", Journal of Systems Architecture, vol. 109, pp. 101761, 2020. https://doi.org/10.1016/j.sysarc.2020.101761

[208] J. Bernal Bernabe, J. L. Canovas, J. L. Hernandez-Ramos, R. Torres Moreno and A. Skarmeta, "Privacy-Preserving Solutions for Blockchain: Review and Challenges," in IEEE Access, vol. 7, pp. 164908–164940, 2019. Online at: https://doi.org/10.1109/ACCESS.2019.2950872

[209] Decentralized identity foundation," 2018, https://identity.foundation/.

[210] C. Allen, "The path to self-sovereign identity," 2016.

[211] IETF, https://datatracker.ietf.org/doc/draft-ietf-ace-oauth-authz/?include_text=1

[212] W3C, A Primer for Decentralized Identifiers", Draft Community Group Report 19 Jan. 2019, https://w3c-ccg.github.io/did-primer/

[213] Wi-Fi Alliance, https://www.wi-fi.org/
[214] S. Behrens. "What's the Status of Wi-Fi 6?". Paessler, 3 May 2019.
[215] M. Kapko. "5G, WiFi 6 Set to Battle for Control of Enterprise Private Networks. SDxCentral, 29 January 2020.
[216] The 3rd Generation Partnership Project (3GPP), http://www.3gpp.org/
[217] Narrow Band IoT & M2M – IoT Global Ecosystem. GSA, April 2020, https://gsacom.com/paper/iot-global-ecosystem-april-2020-summary/
[218] R. Merritt. IoT Nets in Two-Horse LPWAN Race. EETimes, 7 May 2019, https://www.eetimes.com/iot-nets-in-two-horse-lpwan-race/#