
https://doi.org/10.1109/ICIT.2017.7915554

©2016
IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.
Will 5G Become Yet Another Wireless Technology for Industrial Automation?

Mikael Gidlund  
Mid Sweden University, Sweden  
Email: mikael.gidlund@miun.se

Tomas Lennvall  
RISE SICS Västerås, Sweden  
Email: tomas.lennvall@sics.se

Johan Åkerberg  
ABB Corporate Research, Sweden  
Email: johan.akerberg@se.abb.com

Abstract—With the growing interest in adoption of wireless technologies in industrial automation and the continuous search for new revenue streams new players are entering the arena in order to make their business grow further. One of the emerging wireless technologies aiming to support industrial automation applications is 5G, targeting anything from extreme throughput (> 10 Gbit/s) to extreme low latency (≤ 1 ms) to ultra high reliability (≥ 99.999%). In this article we intend to discuss the potential and challenges of adopting 5G in real industrial environments and give a more balanced picture compared to previous articles mainly written by telecom researchers and vendors as a way to promote their technology. Specifically, this article will discuss and provide some real industrial requirements, describe the main technical features of 5G and try to assess what applications it will support that are not by already supported by existing technologies. In the end, the success of 5G will depend on appealing business models and scalability, i.e., whether or not the same equipment can be deployed worldwide without any changes and spectrum rules, and the migration paths beyond 5G for reasonable business risks.

I. INTRODUCTION

Wireless communication systems are becoming more and more frequently used in industrial automation, especially within the domains of process automation and factory automation, mainly due to installation and maintenance costs, but also because they provide the possibility to predict maintenance and perform measurements where it previously was not possible, or very expensive (rotating machinery etc.) [1]. Initially, condition monitoring applications were targeted for the use of wireless technologies and it is still the predominant application, although closed-loop control applications are starting to emerge. Today there are several mature (released as early as 2007) wireless standards for process automation targeting monitoring and closed-loop control applications. The market leading standard in this area is WirelessHART [3], followed by ISA100.11a [4] and then Wireless Networks for Industrial Automation - Process Automation (WIA-PA) [5]. For factory automation the dominant technology is Wireless Interface for Sensors and Actuator (WISA) from ABB [7] and IWLAN by Siemens [6]. The PROFIBUS and PROFINET user organization (PNO) published in 2012 the WSAN-FA standard [8], based on WISA. WSAN-FA utilizes the PHY layer of IEEE 802.15.1 (Bluetooth) and the raw data rate is 1 Mbps and transmit power is 1 mW. The system supports frequency hopping and the hopping sequences have a period of 77 frames where each hopping frequency is visited exactly once per period. WSAN-FA allows to support up to 120 input slaves or 60 I/O per cell in different time/frequency slots with a cycle time \( T_{cyc} = 2.4 \, \text{ms} \) and maximum delivery times of acyclic data below 15 ms. The IWLAN is a proprietary solution based on 802.11 PHY using a point coordination function (PCF) on the medium access control (MAC) layer. PCF enables to assign regular network access to clients and the resulting polling scheme is intended for transmission of asynchronous data traffic as well as real-time traffic.

The next (5th) generation wireless systems are expected for initial deployment in 2020. After the initial 5G release (release 15) targeting mobile broadband, 3GPP (3rd Generation Partnership Project) will focus its efforts on new use-cases envisioned for wireless access, and the associated requirements and technologies. Some expected use-cases involves applications requiring delays close to 1 ms in combination with high reliability. In general, 5G will consist of three generic services:

- **Extreme Mobile Broadband (xMBB)** aims to deliver extremely high data rates and low-latency communication, and extremely coverage improving the Quality of Experience (QoE) by providing reliable moderate rates over the coverage area.
- **Massive Machine-Type Communication (mMTC)** aims to support tens of billions of network-enabled devices to be wirelessly connected. Scalability and wide-area coverage is prioritized over peak rates as compared to xMBB.
- **Ultra-reliable Machine-Type Communications (uMTC)** aims to provide ultra-reliable links with low latency in order to offer new applications with extreme requirements on availability, latency, and reliability.

The penetration of mobile communication in industrial automation has so far been low and mainly focused on remote service application and alert systems. Recently, a few research articles on suitability of using 5G in industrial automation have been published [9] - [12]. However, many of the assumptions and requirements made in the research articles in question are not realistic, which also makes the results described in the articles questionable. Moreover, most of the work in [9] - [12] focus on physical layer solutions, which will only provide solutions to a few challenges based on the diverse set of requirements in industrial automation.

The aim of this paper is to give a brief overview of use cases, requirements, and challenges of using 5G in industrial...
A major challenge and key question is if the plant owners are willing to pay for a licensed spectrum instead of using available technology which uses an unlicensed spectrum that can be used worldwide without cost. Another important challenge for 5G is to serve mission-critical application with safety critical hard real-time requirements and very low end-to-end latency (<1ms) when transmitting data over networks. In this paper we also identify several challenges that needs to be solved before a wider market acceptance of the use of 5G in automation is reached.

II. REQUIREMENTS OF INDUSTRIAL APPLICATIONS

In the literature, a few attempts have been made to characterize the large set of requirements of industrial automation [1], [2]. Recently, a set of papers from telecom vendors were published, which included requirements of industrial automation for 5G [11], [12] were listed, but most of these fail since they assume that all use cases within industrial automation has the same set of requirements. Nevertheless, industrial automation can be divided into several sub-areas, building automation (BAS), process automation (PA), factory automation (FA), and substation automation (SA), which have different distinct communication requirements, see Table I. The most common use cases for communication in industrial automation are [2]:

- interconnection of automation equipment distributed over large distances.
- interconnection of dedicated real-time automation systems with operator work-places for control and supervision.
- closed-loop control, ranging from slow to fast control applications with sampling rate between seconds to milliseconds.
- interlocking and control applications.
- monitoring and supervision applications.

The classic automation pyramid is divided into several layers with different networks; each network has different demands and rate the importance of various properties differently. Wireless communication for the fieldbus level in the automation pyramid must guarantee high reliability and low and predictable delay of data transfer while the higher layers of the pyramid has more relaxed latency constraints and real-time properties. A less obvious communication system requirement is the ability to enable high resolution time synchronization in the system in order to achieve distributed measurements at the same time for later control, where the predictable latency can be compensated for. In addition, failover mechanisms are needed in order to prevent production outages or damage of the mechanical equipment, in the event of deadline misses due to lack of seamless failover mechanisms. A good source for real requirements for industrial automation is given in [2].

III. SOME USE CASES SUPPORTED BY 5G

It is expected that future 5G networks will enable new use cases and services that it successors (3G and 4G) were not able to deliver. The grand challenge for 5G will be to support different use cases in a flexible and ultra-reliable way. Future 5G systems promise to support maximum latencies on the order of milliseconds. However, this will still give too large delays in some scenarios. Below we list some use cases which are foreseen to create new revenue stream for mobile network operators and network equipment vendors.

- **Mobile broadband** is the key use case today, and will remain so the in the future. The main difference from today’s mobile broadband is that 5G will in the future bring truly immersive experience to the end users. Data is continuously growing and is predicted to grow at a rate of 30-50% annually, which is expected to continue until 2030 [13]. 5G is foreseen to be the carrier for distributing systems to the cloud, which requires low end-to-end latencies in order to maintain a good user experience.

- **Industrial monitoring and control** is a use case where wireless communication will become increasingly important in order to reduce wire installation costs, ease of maintenance, mobility, and the possibility to make measurements that were previously impossible because of wired communication and available wireless technologies lacking support of some of the stringent requirements [2]. The key challenges for 5G will be to deliver high availability and automatic failover while not violating the application deadlines, predictable delays, and viable business models not jeopardizing companies’ return of investment (ROI) in large-scale industrial production facilities.

- **Automotive** is expected to be one of the main key drivers for 5G with increased performance, increased security and D2D communication. For instance, existing onboard sensors and systems, can greatly improve performance and safety though 5G communications and the road utilization since vehicles can drive closer to each other. There are also discussions about traffic safety use cases in several EU-funded projects, nearly all car manufacturers are working towards finding solutions.

- **Smart grids** and merging communication networks with power networks in a smart grid is not a new technology. In order for 5G networks to be used in smart grids, it be able to monitor the grid operating status in real-time, isolate faults, and implement self-recovery, avoiding large-scale power failures. 5G is expected to play a significant role in the access networks and backhaul networks.

- **Health** sector will continue to grow and is already using mobile communication as enabler in several use cases such as tele-medicine, remote monitoring and sensors for measuring heart rate and blood pressure. 5G is expected to enable the Internet of medical things, and technologies such as cellular, WLAN and Bluetooth will enable IoT communication across use cases and 5G will be the network that connects these things.
IV. MACHINE-TYPE COMMUNICATION FOR INDUSTRIAL AUTOMATION

Automated devices such as meters or sensors are often referred to as M2M communication and in 3GPP, M2M is also called **machine-type communication** (MTC). MTC has been identified as an important use case if they can be efficiently supported, it will offer an attractive revenue stream to the mobile operators. One goal of MTC is to automatize everyday processes and take advantage of delegating task to machines. The main use cases identified for MTC are: track and trace, monitoring, control, and home automation. There are two types of MTC, **Massive MTC** and mission-critical MTC. Massive MTC assumes data services for typically infrequent data transmissions and is focuses on large volumes of data and connections using low cost devices that offer lower bit rate and low power consumption system design. The primary use case of MTC is utility metering.

For massive MTC the following requirements are considered important for 5G: 1) 10 years of battery life with AA battery; 2) Coverage enhancements of 20 dB; 3) 30000 devices per cell (most cells will contain smaller number of devices); 4) Low device complexity. There are a few MTC challenges and the key to addressing the challenges is to simplify the complexity of the network to support MTC devices without jeopardizing the security or quality of services for both MTC and H2H users that share the same network infrastructure. Supporting a massive number of MTC devices simultaneously will cause problems in connection establishment and radio resource allocation. Another major challenge is to accommodate small bursts of data. The design of LTE systems require that the user establishes the connection before sending the information data. This approach adds signaling overhead to the information data. A way to improve the efficiency of data transmission is to use data aggregation, but this approach could incur additional delays and is only applicable to non-delay-sensitive MTC applications.

In contrast, mission-critical MTC targets ultra-high reliability, very low latency, and real-time control of objects and processes. In mission-critical MTC, ultra-reliable communication is vital for services that provide very high reliability (reliability levels can be in the 99.999th percentile) and often very short latencies (end-to-end packet transmissions within 1 ms). The main application of mission-critical MTC is predicted to be Road safety and traffic efficiency, but industrial control and monitoring are also predicted as an area of application (mainly for factory automation). So far, the use of mission-critical MTC for industrial control has received little attention from the main telecom vendors. It is likely that business cases are not in place and requirements cannot easily be met by a generic wireless design (a real-time system managing hard deadlines is needed). Fulfilling the safety critical hard real-time requirements of industrial automation applications requires a pessimistic design, which is very costly.

V. SELECTED ENABLING TECHNOLOGIES FOR 5G WIRELESS COMMUNICATION

In order to fulfill the stringent requirements and use cases for 5G systems there are several enabling technologies that need to be combined. In this section we will highlight some of the key enabling technologies, but it should be noted that there are many more that can be used for 5G system. It is very important to understand that most of the topics below are in the early stage of research, and not at all ready for implementation in a real system.

A. Massive MIMO

A new paradigm, shown to provide efficient communication between base stations and terminal devices, is the use of massive numbers of antenna elements at the base stations. With such configurations it is possible to perform very efficient spatial multiplex of many users in the same time/frequency resource, focusing the radiated energy toward the intended direction and in the same time minimizing intra- and intercell interference. The idea of MIMO has been around for decades, but the idea of actually deploying base stations (BSs) with more than a handful of service antennas is relatively new and was initiated some six years ago [14].

There are several challenges connected to Massive MIMO, one is that more advanced signal processing capability is required in the nodes. Furthermore, each node must be able to determine the data transmitted from one antenna to that transmitted from another, otherwise network performance will be limited [15]. Other challenges urgently in need of attention is low-cost hardware, hardware impairments, fast and distributed coherent signal processing, internal power consumption, and channel characterization.

B. Network Slicing

Slicing in general is to use virtualization technology to architect, partition and organize computing and communication resources of a physical infrastructure to enable flexible support of various use cases. It should be noted that network...
slices is not a new concept, which is demonstrated by the fact that a VPN is a basic version of a network slice. In future 5G networks, different customized virtual network will exist simultaneously and without interfering each other. For example, a customized virtual network for an augmented reality application requiring very high throughput can co-exist with a customized virtual network for autonomous vehicle control which requires low latency. Such use case can be prearranged through agreements and provided by the mobile operator on demand. The critical aspect for such arrangement is the business agreements.

C. Full Duplex

Previously it has been assumed impossible to have simultaneously transmission and reception of overlapping signals in the same frequency channel due to the resulting self-interference. In existing cellular systems (4G), both frequency division duplex (FDD) and time division duplex (TDD), are used but it requires two separate channels to realize orthogonal transmission and reception. In 5G, full duplex allows radios to transmit and receive on the same frequency and effectively doubling the spectral efficiency. Hence, the most important challenge in order to realize full duplex communication is to effectively eliminate self-interference [16]. The canceling needs to be done either at the antennas, or before the A/D converter or in the digital baseband. Recently, there has been a significant progress in the self-interference cancellation, which presents a great opportunity to realize full duplex communications for the next generation of cellular networks [17].

D. Device-to-Device Communication

Device-to-device (D2D) communication can be described as direct communication between two devices which are in close proximity to each other, the base station can allow the users to directly communicate with other without data packets are transmitted via the base station. It is expected that D2D will an important communication method for 5G, especially to provide low latency between devices over short distances and minimal propagation delays. The potential of improving spectral utilization has promoted a lot of work in recent years, which shows that D2D can improve system performances by reusing cellular resources [18] and expected applications for D2D is peer-to-peer file sharing, voice, video streaming, and content-aware applications. Although there are benefits with D2D communication, some issues like direct discovery, interference management, and direct communication need to be properly addressed [19].

E. Millimeter Wave Communications

A major problem for future 5G wireless networks is the bottleneck of wireless bandwidth but recently is has been shown that Millimeter-wave (mmWave) communication can be a very promising solution for future 5G networks. What mmWave refers to is the part between 30 and 300 GHz of the spectrum, although the majority of current activities are targeting the 28 GHz band, the 38 GHz band, 60 GHz band and the E-band (71-76 GHz and 81-86 GHz) [20]. There is considerable research on mmWave propagation in the 60 GHz band [21]-[23]. The high available bandwidth in this range of the spectrum allows extremely high data rates. Hence, there exist several challenges on PHY and MAC level to solve before mmWave can make a big impact on the 5G wireless networks. The major drawbacks with using mmWave is huge propagation loss, directivity, and due to weak diffraction ability, it will become sensitive to blockage by obstacles such as humans and machines. For instance, blockage by a human penalizes the link budget by 20-30 dB [24]. The main application of mmWave in 5G will be niche applications and small-cells [25], [26], but for industrial automation it is more difficult to see its applicability because industrial environments evolve over time, as well as being cluttered with metallic equipment and obstacles.

VI. MAIN CHALLENGES FOR ADOPTION OF 5G IN INDUSTRIAL AUTOMATION

There are many challenges, both from a technical and business perspective, before 5G systems can be successfully adopted in industrial automation. This section will highlight a few of the challenges that have to be overcome.

A. Business Models and Spectrum Availability

A good business model is key for future use of 5G technology in industrial automation. Especially since 5G is based on using licensed frequency bands, i.e. plant owners need to buy or lease a spectrum. An important question that needs an answer is why plant owners should invest in new 5G technology when most of the considered use cases already have a low-cost solution in the form of current wireless technologies. If we consider smart meters, they already come with cellular connectivity via GPRS and 3G systems. An argument for using 5G is that in the future the meters will be more complex and communication-intensive, and enable real-time estimation and control of the grid state. However, if this becomes reality, there are already several wireless systems, such as industrial WSNs and WLANs, that fulfill the requirements. Within industrial automation, current wireless systems uses one of the unlicensed frequency bands (488 MHz, 868/915 MHz or 2.4 GHz). The main reason for using unlicensed frequency bands is that there is no additional cost for frequency license and it offers flexibility in terms of worldwide applicability. The general assumption of the ISM band is that it is more mature and will not change at the same pace as the spectrum for the mobile telecom system, and it currently fits better with the life time of automation equipment. A licensed frequency band might only work in certain countries. A benefit of using a licensed frequency band is that interference from other devices is less than in an unlicensed band and it might become easier to guarantee some of the hard requirements on reliability, availability, and latency. Hence, methods and tools to handle coexistence between wireless systems in unlicensed frequency bands [29] have been
developed. A selling point from 5G vendors is the use of higher frequency spectrum, but the question remains how well this will work in industrial environments, which as previously mentioned are filled with materials that will have a high impact on the radio propagation. Plants can also be very large (several kilometers and still growing) making the range of wireless communication an important factor as well.

Using 5G in process automation or factory automation is questionable since the number of connected sensors is in the range of a few hundred to a few thousands. This makes a 5G solution expensive, especially since the plant owner needs to buy one license or at least lease one license. A related question is if the mobile operators see any business opportunities for such a small number of few connected sensors. If 5G is sold in a way similar to the current telecom technologies (GPRS, 3G, 4G), with operators charging per bit, it would be extremely expensive for a plant owner to operate the plant compared to the cost of using wireless technologies such as WSNs and WLANs.

B. Latency

The unanimous opinion is that 5G standards should support less than 1 ms end-to-end round trip delay (latency) in order to support future industrial applications. However, there is no clear unified view which parts that are included in this end-to-end delay. According to some research articles [11] and white papers [31], the latency budget (based on Tactile Internet ITU watch report) is given as the time delay between data being generated at the sensor and being correctly received by the actuator [31]. This requirement will become very difficult to achieve since the law of physics puts some hard limitations on this. The speed of light and electricity is limited. In 1 ms, light can travel around 200 km through an optical cable. If we consider user equipment processing delay to be 4 ms, base transceiver station delay is 4 ms, core network delay is 1 ms, the scheduler request and grant will have some delay, so this will already adds up to more than 7 ms. Even if the network equipment does not add any latency, the maximum round trip distance is 100 km which make it impossible to remote control a robot with a latency of 1 ms in one part of the world from a place halfway around the world. Today LTE uses 1 ms transmission time interval (TTI) and 8 ms waiting time at every retransmission which results in end-to-end latencies of 20 – 40 ms. In order to decrease the latency for LTE (and 5G) several modifications are needed. First, reduced TTI, e.g., down to 100μs, and shortening the OFDM symbol durations will enable fast and efficient data transmission. Secondly, redesign of physical channels allowing early channel estimation. Furthermore, use of convolutional codes (e.g., for data channels) and block codes (e.g., for control channels) will provide fast and reliable decoding, and the use of high diversity levels will improve the reliability of signal detection and decoding.

It should also be noted that today there are wireless solutions for factory automation that can support latencies below 5 ms (i.e., WISA and iWLAN), and for process automation there are several mature industrial technologies, such as WirelessHART, that already supports monitoring and control (<1s) applications. Recent research results also show that WirelessHART can support safety-critical applications with latencies below 50 ms [27, 28].

C. Reliability

In addition to promising very low latency, 5G also promises to deliver ultra high reliability, defined as 99.999% or higher. In the literature on mission-critical MTC, no difference is made between the concepts of availability and reliability. They are referred to with the all-inclusive term "reliability". The most important requirement for an industrial automation system is availability. I.e., the plant must be able to produce or construct as intended do with minimal downtime. Downtime means no revenue and only costs for the plant owner(s). In order too maximize the availability, the reliability of various system components (e.g., networks) becomes a critical requirement. Today Distributed Control Systems (DCS) can in the extreme case promise a reliability of 99.9999% [30]. This includes all the components of the DCS, instruments, I/Os, controllers, and the communication. It is important to remember that availability is the most important requirement, but in order to achieve safe and economical viable production there are also requirements on deterministic systems in order to achieve safe production facilities as well as maintain extremely low variation in production quality to be attractive on the global market. One way of achieving higher availability and to eliminate single-point-of-failures in 5G is to use multiple access points since using a single AP can cause 100% network outage. Some technical challenges that arise when using several access points is that inter-cell interference can decrease the reliability and latency. The problem with inter-cell interference can be partly solved by using perfect interference cancellation schemes and power control schemes. Nevertheless, including more access points will become a costly solution.

D. Determinism

There are other important aspects which seems to be missing from the discussion on how 5G will support industrial automation applications with hard real-time requirements. There is much more to achieving this than just supporting low latency and high reliability. Determinism and predictability, the ability to foresee and know how the system will behave is critical for the support of all the components in the system. There are many different ways to achieve a high reliability for communication, e.g., retransmissions, error codes, multi-path routing, however, many of these solutions do this at the cost of sacrificing determinism and introducing jitter. E.g., keeping on retransmitting a packet in order to successfully transfer it may cause it to miss its deadline. For 5G, the focus seems to be on achieving high reliability on the PHY layer, but a complete communication solution also consists of higher layers in the stack, e.g., TCP/IP, which must behave deterministically. If 5G uses TCP/IP it will inherently be non-
deterministic, regardless of how good the PHY is (apart from delivering 100% reliability).

VII. CONCLUDING REMARKS

This article discusses the use of 5G wireless communication in industrial automation scenarios and their challenging communication requirements. 5G is expected to deliver extreme throughput, be ultra-reliable and provide low latency communication that can create new applications and unlock new revenue sources for telecom operators. It is important to emphasize that nothing comes for free. Absolute zero latency and/or 100% reliability can never be achieved due to inherent physical properties of any data transmission link, wireless or using fiber optics. Physics implies that there will be an achievable limit for latency/reliability with given resources (e.g. bandwidth, transmission power, residual error rate). Thus, the cost for reducing latency and improving reliability on the physical layer is expected to grow exponentially and diverge unless all layers in the communication stacks are considered. The main challenge for 5G, for receiving market acceptance within industrial automation, and particular in process- and factor automation, is a sound business model and a strong argument for using licensed frequency bands, which are costly, instead of using license-free bands, which are free of charge. From a technical point of view, the challenge will be to present a solution that covers the wide range of requirements and for 5G to outperform the existing wireless technologies that already work well with many of the targeted applications. Nevertheless, it should be mentioned that 5G will find new applications within industrial automation, mainly where wireless communication links with high throughput is of importance, and where demands on latency, predictability and availability are more relaxed.

REFERENCES


