Abstract

This chapter describes the main features of the wireless communication systems of urban rail and related applications. The perspective will be complete: application, network and physical layers will be discussed. Moreover, to properly address some of the challenges that these systems face, we will provide a deep insight into propagation issues related to tunnels and urban areas. Finally, a detailed survey on the directions of research on all these topics will be provided.

Keywords: communications, LTE, mass transit, transport, urban, wireless, WLAN

1. Introduction

In the last few years, wireless communication systems for transport systems, in particular for urban transport, are gaining importance. The so-called intelligent transportation systems have a major impact on the performance of transport, because they can increase the capacity of the lines (and even add some extra safety), provide added-value features to operators (i.e., remote maintenance) and also additional services to passengers, such as the very popular access to the Internet. However, commissioning all these features implies facing many relevant challenges at every conceivable layer, not only technical, but also regulatory, financial, etc. In this chapter, we have adopted a practical approach to all these issues, trying to focus on what the railway engineer really needs to know: the essentials, but also the challenges and where the Gordian knot is.

In this chapter, starting from the application layer and ending on the physical one, we cover the most important aspects of these systems. The chapter structure is based on the
well-known open systems interconnection (OSI) stack and, in our opinion, is very helpful to discuss all the details of very complex systems such as wireless in urban transport. We start in Section 2 by presenting the most relevant systems from the point of view of the criticality of the service: safety services such as communications-based train control (CBTC) or public safety (PS) radios; operator-oriented, like closed-circuit television (CCTV) or operation and maintenance (O&M) systems; and, finally, services only focused on the passenger (with no safety consequences), like infotainment or the Internet access. In Section 3, we provide a deep insight on networks and technologies that make all the previous services possible. In this section we discuss some aspects that require our attention when deploying services such as these, with all the mobility issues and many other problems that may arise. Section 4 is dedicated to explain the physical layer related to urban transport. Two main scenarios need to be covered: strict-sense urban (city landscape, mostly for tramways) and tunnels (mostly for subways). These two scenarios present significant differences between them, tunnels especially implying a challenge for wireless systems. In this section we also discuss the ‘spectrum problem’ in order to provide some key ideas on the frequency allocation of wireless systems. ‘Wireless communication systems for urban transport’ is a fast-evolving field and, as mentioned before, it is far from being finished; and a lot of research work is still in progress. The European Union, through its H2020 initiatives (particularly, Shift2Rail [1] and Roll2Rail [2]) is investing a lot of funding in some railway ‘hot topics’. In Section 5 we introduce all the lines of research that are meant to change the face of railways in the next 20 years. Finally, Section 6 hosts the main conclusions of this chapter.

2. Urban rail applications based on wireless communications

The next step in this chapter is to know about the railway services that demand wireless connectivity. This is the most visible part for railway operators, stakeholders and any other railway-related agent. Service must be understood as the supply of something needed. Therefore it is a synonym of application and in the current text we will use both of them interchangeably.

There is no general consensus on the classification of services, but railways always have paid a lot of attention to safety, so we use this approach to classify services. Accordingly we may say that we have three types of services: mission-critical (safety-related); operator-oriented, with no safety implications but operators heavily rely on them (like CCTV); and finally customer-oriented: every service that is intended only for the passengers, having no relation with operational or safety issues. Sometimes it is very complex to state clearly which domain a function belongs to. For example, consider a CCTV camera placed in the front of a train which is used by an operator to detect if somebody has fallen onto the track (letting the driver to stop the train before running over the individual). In strict terms, CCTV is an operator service (with no safety implications) but if this system fails, it could cost lives.
2.1. Mission critical services

Two taxonomies fell within this category: signaling-related, with a strong safety commitment and public-safety systems. The main difference between the two of them is that signaling systems take care of the safe movement of the train (telling the train how far it can go and how fast; helping the train to report its position and speed to the wayside equipment; etc.) and public-safety communication systems that do not control any movement of the train but provide a safe communication channel that enables railway agents (drivers, operational control center (OCC) controllers, maintenance and security staff, etc.) to communicate appropriately.

Railway signaling is a very broad field whose extent exceeds by far the purpose of this chapter. Here we focus on the implications on wireless communication systems that carry the data of the signaling service. For urban rail, the most relevant transmission-based signaling system is CBTC (communication-based train control). Using the definition of the very CBTC standard [3], this system is “a continuous automatic train control system utilizing high-resolution train location determination, independent of track circuits; continuous, high capacity, bidirectional train-to-wayside data communications; and train borne and wayside processors capable of implementing vital functions.” There are many keywords in this definition, but we need primarily to focus on three: continuous, high capacity and bidirectional. Continuous means that the communication between train and wayside could be everywhere, not only at certain points; high capacity obviously means that the system is able to handle all the data needed with some restraints (delay, jitter, data rate, etc.); and finally, bidirectional implies that not only the train will receive data from the wayside but the wayside will also receive from the train. These three aspects are essential to understand how a CBTC system works.

To explain CBTC properly an entire book like this would be required, so we use a simplified model of it, with two basic use cases. In our model, the train is able to know (in a very precise and reliable way) its speed and its position (by reading beacons located on the track and using odometry techniques). The measured speed and position needs to be transmitted to the wayside equipment, which is responsible to tell every train how far it can go and how fast (this is called the ‘authority movement’). So, the train transmits its position and speed, and receives the authority movement. Given that this information is critical, it needs to be delivered in a very short time, otherwise it would be useless. If a train experiences some kind of difficulty and finds itself unable to receive authorities of movement or to transmit its position, it will stop before going into a dangerous state (i.e., the risk of crashing into the next train).

Once we have this CBTC simplified model we start to look at the requirements to the communication subsystem. As the data to be carried is critical, we want a robust communication system, able to work even under the eventual failure of some elements. This implies a redundant solution for the communication network, with separate paths for the wirings, duplicated base stations, switches, routers, power supplies, etc. Regarding the wireless part, this approach is called ‘double-layer’ coverage, and it is a very expensive solution, as it represents (more or less) twice the cost that you would have on a ‘single-layer’ coverage. Figure 1 is a picture of a double-layer deployment, with coverage by two different base stations at every single point of the track. Besides, this network also needs to be able to handle the bidirectional communication of the wayside with every train. Usually this communication between the wayside and
the trains is done through a polling routine, where the wayside iterates from the first to the last train, requesting them to communicate their positions/speeds and sending to them the calculated authority of movement. This polling routine impacts on the whole performance of the system (and limits the number of trains a single wayside equipment can handle).

As the data volume to be sent is not very large (64–128 bytes per message every 1–2 s, depending on the supplier) and the required delay is less than 800 ms the quantitative requirements for the network are not very demanding. Therefore, all the aggregated data rate usually does not exceed 10–50 kbps, a figure that can be reached by almost every state-of-the-art wireless system (and even by some not so modern). In Section 3 we discuss in more depth the available wireless technologies, but since its early years, almost every commissioned CBTC has adopted IEEE 802.11 [4] technologies for its radio subsystem. These solutions are deployed in unlicensed (Industrial, Scientific and Medical, ISM) bands, something that has security implications. Ultimately, the trend for some vendors is to move instead to long-term evolution (LTE), but these cases are still an exception. For further details on the discussion see [5].

There are only a few CBTC vendors with a tested solution: Alstom, Ansaldo STS (now part of Hitachi), Bombardier, Siemens and Thales. There are many other signaling companies working on this technology, but the homologation procedures are very arduous and the operators frequently requesting to prove a tangible previous experience, this is an entrance barrier for new players. CBTC is not very much standardized (at least not like its high-speed counterpart ERTMS [European Rail Traffic Management System]), which means that vendors have a lot of liberty in the implementation of their solutions and CBTC not being interoperable, unfortunately for operators, with CBTC lines as they usually find out that their trains are captive on these lines.

The second group of mission critical services is the public safety (PS) issue, sometimes called ‘professional mobile radio’ (PMR). These systems are the evolution of the analog radiotelephony systems that were put into service in the early 80s. Strictly speaking, they are not safety systems (they do not protect the train like the signaling systems) but most operators and stakeholders act as if they were. As the OCC loses the communication with the train driver, some operators react to a breakdown of the radio telephony by evicting the train. The most common functionalities are the voice calls between the OCC and one or more trains. The addressing of these calls is usually required to be not only direct but also “functional” (based on the train number that may change every day) or “regional” (based on the location of the train).
There are two major technologies for PS radios: the first one is GSM-R (global system for mobile communications – railway), which carries PS data and signaling data in high-speed lines, mainlines (out of the scope of this book) and sometimes in commuter lines (like C4 in Madrid). The second technology is TETRA (terrestrial trunked radio) or any of its competitors. In this chapter we focus just on TETRA, as it is the most used PMR technology in the world. This is especially true for railways but its use is also widespread among police, firemen, border security, search and rescue, etc. All over the world it is assumed that the future of PMR systems is LTE (especially TETRA, whose association has publicly supported the effort of 3GPP LTE to provide PS functionalities [6]). In the next section, we discuss with more detail the feasibility of LTE for public safety.

The TETRA standard is very useful for railways because it allows operators, stakeholders, maintenance people, etc., to benefit from the following functionalities:

- Group calls: not just one-to-one calls, but also one-to-many.
- Functional addressing: i.e., the OCC wants to talk to the driver of train 2.
- Geographical addressing: i.e., the OCC wants to transmit a message to all trains that are in a certain area.
- Preemption: In the case of emergency, calls with less priority can be preempted in order to release resources for the higher priority ones.
- Direct communication: in the absence of a fixed network, two terminals can communicate with each other directly.
- High security levels: end-to-end encryption, authentication, integrity, etc.
- Data: although TETRA is not a broadband system, it is also able to transport short pieces of data from the train to wayside and vice versa.
- Push-to-talk (PTT): calls should be established in a very short time, usually, less than 500 ms. This represents a very demanding requirement for a network.

There are many TETRA vendors around the world, but it is a niche technology if we compare it with public mobile solutions such as LTE or UMTS (universal mobile telecommunications system).

2.2. Operator-oriented services

There are many operational services, the most common being CCTV (both real-time and recordings), passenger information systems (PISs) and telemetry of the train.

Passenger information systems deliver multimedia messages announcing passengers the proximity of the next station, the end of the line, etc. Most of the time only text messages are shown on a LED display at the end of the vehicle and/or audio recordings are played. However, modern subways and trams also provide video content on screens with information about nonoperational issues and also, very often, advertisement. This type of PIS is often called
‘infotainment’. Some operators consider this system that provides a ‘customer-oriented service’ and not an operational one.

To provide an efficient passenger information service, the onboard system needs to know the location of the train with a certain grade of accuracy. This is done usually by integrating the PIS with the onboard signaling subsystem, deploying beacons in the track, or integrating the onboard PIS with a wayside system that knows the position of the train. The first method is the most accurate, it does not need a wireless system at least for the PIS (data remains in the train) but can be very expensive, because this integration can be very time consuming hard to put into service (many use cases and degraded modes). The second method is also costly, because beacons need to be installed and maintained, and sometimes it is prone to failures (a train misses a beacon). The third method can be less expensive if you already have a train-to-wayside radio, but this ‘general purpose’ radios are not as reliable as the PIS should be. Also, the integration difficulties remain. Of course, onboard infotainment systems need to download the multimedia contents to be shown and a broadband train-to-wayside radio is required.

Closed-circuit television (CCTV) is an operator-oriented service already available in almost every subway and tram of the world. Sometimes the recordings are kept in the train (if there is no train-to-wayside system available or there is no OCC integration to manage the video). Here we focus on the more general case where the video is available in the wayside and is also integrated with the train control management system (TCMS) of the train. There are two basic functions on CCTV systems for railways: real-time video and recordings. Real-time video can be watched in the cabin of the train by the driver and/or in the OCC. Some operators/stakeholders have ‘security centers’ where all the security aspects of the system are centralized (access controls, CCTV, security staff management, etc.) but in most cases this is part of the OCC. As real-time CCTV (video streaming) systems are very demanding in terms of resources, the train-to-wayside wireless system shall be properly designed. The other major function is the viewing and downloading of video recordings. Some operators force their onboard CCTV systems to download their recordings to be stored somewhere in the wayside (in a NAS [network-attached storage], probably), while others prefer to store it on the train itself and access them on demand through the wireless system.

Apart from the previous two functions, there are many other railway-related than usually enriching a CCTV system. Having the video of the upcoming platform in the cabin of the train can be very useful in the case of crowded stations; and integrating the smoke detectors with the onboard CCTV (via TCMS) so when a smoke alarm is triggered, the driver automatically can watch the nearest camera to the alarm/source of the smoke, are some examples. In both cases, a big effort on systems’ integration shall be done, so it is important to keep in mind that a railway-CCTV system is not only a CCTV system placed in the railway system. CCTV systems, besides being relevant for security purposes, could also be helpful for operational ones (i.e., for driverless trains, cameras placed at the front of the train pointing to the track, etc.).

Finally, there are telemetry services where some data is transmitted out of the train to some/a particular wayside system, in order to analyze them, take decisions based on them, etc. There are two main final users of this data: train maintainers and railway operators. All these data
needs to be obtained inside the train (sensorized), transmitted out of the train (usually through
a gateway from the sensors network to the IP network) and finally, handled at the wayside
(and maybe even stored on a database and integrated into the OCC or maintenance tools). In
modern trains, the digital buses of the trains contain a lot of data of almost every single piece
of the train, so it is rarely needed to deploy new sensors. This type of buses, deployed for TCMS
purposes, usually follow the IEC 61375 [7] family of standards. In Section 5 we will know about
the evolution of these standards.

2.3. Customer-oriented services

The last group of services is focused only on the client experience, with no safety implications
or added-value for the operators. The most important service in this category is the Internet
access for passengers within the train. At first sight it might seem that it is not a very challenging
service to be provided, but many difficulties may appear:

• It is an extreme environment for electronic devices, with vibrations, high temperatures,
electromagnetic interference (EMI), etc. These perturbations are usually handled together
requesting that every onboard device complies with the railway normative.

• Vehicle penetration loss (the losses of the signal strength due to the presence of the vehicle)
is usually between 15 dB and 25 dB, depending on the frequency of the signal and type of
vehicle. This can be avoided using mobile relays.

• Cybersecurity. Having user data and safety data in the same network implies handling
various security issues, much more complicated than having only safety and operational
data in the network.

3. Networks and technologies

In this section we discuss the technologies that lie between the services explained in Section 2
and the physical layer to be covered later. Independently from the technology, any train-to-
wayside system carrying data from different services should have some QoS (Quality of
Service) policy, in order to implement prioritization schemes. This is not a trivial task (espe-
cially among legacy technologies). Security issues shall be addressed too, more than ever if the
network is reachable for passengers (WiFi services, for example). Furthermore, to implement
a successful train-to-wayside system it needs to be properly integrated in the ground network.
Again, this is out of the scope of this book, but many networking issues may appear if both
technologies are not compatible. For instance, if base stations need layer 2 connectivity, but
this is not possible with the topology of the network, the performance of the train-to-wayside
system will decay.

We cover two major technologies the 3GPP LTE [8] and the IEEE 802.11 [4] family of standards
with a spin-off of the latter, but with proprietary modifications (this is, out of the open
standard).
3.1. LTE

Undoubtedly, 3GPP LTE (long-term evolution) has a key importance in railways. Also, due to its popularity in the public mobile world, it is also very frequent to find out that passengers have LTE coverage. Furthermore, the intention of the 3GPP (the world organization responsible of the standardization of LTE) is to include railway use cases and functionalities. However, the 3GPP specifications for LTE are a living thing, with frequent releases and with the vendors pushing to develop their solutions. This implies that what is written here can be outdated in a few years, so our suggestion is to learn here the basics of LTE, the philosophy behind it while also trying to keep up with all the novelties that the market could offer.

The first problem that appears when you consider LTE as a candidate for your wireless system is the spectrum. In radio communications you cannot transmit anything you want due to the regulation that you shall comply with. So, first of all, you need to know in which band your system is going to work. There are two possible types of LTE networks: public (a service typically provided by mobile operators like O2 or Vodafone), where the railway users compete with the passengers and everybody else for the resources; or private, where the railway services are the unique users of the network. An arrangement between both extreme options is also possible (for example, a public mobile operator being a piece of their spectrum for railway use, de facto creating a ‘private LTE’) or, simpler, the railway service being one plain customer of the mobile operator.

The LTE architecture is very flat, having only a core and the access network (see Figure 2). The access network consists of a set of base stations called eNB (‘evolved node B’) which handles the radio interface between the user equipment (UE) or terminals, on one side, and the interface with the core, on the other side. There are many possible implementations for the eNB, depending on the scenario, for example, one eNB, two radio heads and a microcore all in one box; in the other side, in a large area, many eNB connected through a dense fiber network to a centralized core. Therefore, flexibility is one of the key features of LTE.

The functions of the eNB belong to two different sublayers [10]: RRM (radio resource management) and MAC (media access control). First, RRM functions are user admission control, scheduling (dynamic assignment of resources to users), interference control between eNB to avoid interferences, mobility management (handovers) and many other minor functions.

The second sublayer media access control (MAC), with two main differences between the uplink (from the UE to the eNB) and the downlink (eNB to the UE). In the downlink, LTE uses OFDMA (orthogonal frequency division media access), a technique that separates the transmissions destined to different users by assigning them different blocks of data subcarriers. Given that every eNB manages limited resources in both time and frequency (the so-called bidimensional “grid” of time slots and subcarriers, see Figure 3), this technique gives eNB enough flexibility to manage their pool of transmissions. In the uplink, due to limitations in the dynamic range of the amplifiers, SC-FDMA (‘single-carrier FDMA’), which is a variation of OFDMA, is used instead. Another example of the flexibility of LTE is that the total bandwidth can be 1.4, 3, 5, 10, 15 and 20 MHz. LTE can work in paired mode (also called FDD mode with a band for the uplink and another one for the downlink, both with the same width) or
unpaired mode (TDD mode, with only one band for both up and downlink) [10]. Carrier aggregation can be used in the recent versions of LTE known as LTE-Advanced [10] and LTE Advanced Pro [9], combining two or more radio frequency (RF) channels of the same, or different, frequency bands to provide very high data rates.

Figure 2. LTE architecture, depicting both access stratum and core.

Figure 3. Radio resources on a LTE eNB.
The spectrum problem

An issue that shall be addressed when working with wireless is the spectrum allocation. Every single country in the world regulates the use of it, so we are not at liberty to transmit wherever we want. There are three main possibilities for this allocation:

Unlicensed bands: In these bands no license from the regulator is required. You only need to check that the emitted power does not exceed the maximum allowed, but as anybody can transmit, you need to take care of the interferences that they may cause. There are many unlicensed bands all over the spectrum; the most popular (for both railway and nonrailway use) are the 2.4 GHz and the 5 GHz.

Licensed bands for railway use: In some cases, the regulators allocate a piece of the spectrum for railway applications, but, unfortunately, this is not very common—for instance, the GSM-R band in Europe (873–880 MHz for the uplink; 918–925 MHz for the downlink).

Licensed bands for non-railway use: These bands represent the majority of the spectrum. No railway applications can use this band, unless an agreement with the owner of the rights is signed.

As mentioned in the previous section, when explaining public safety services, LTE is supposed to be the next dominant technology in this field. However, LTE as it is now is not able to provide the PS required functions [6] (future 3GPP LTE releases will come to solve this problem).

Also in 3GPP LTE’s roadmap [6] is planned to include a very useful technology for railway use: the mobile relay (MR) [11]. It is an onboard device that divides into two parts, the link between the UE and the eNB: one part between the UE and the MR (inside the train), and another one between the MR and the eNB (out of the train). This approach offers many advantages like reducing losses caused by the structure of the vehicle, the possibility of having better DSP (digital signal processing) techniques than in cheaper UEs to avoid multipath, Doppler and other undesired effects; group handovers can be performed (only the MR does the handover); obviously, it is an opportunity gap for railway operators to increase their incomes through a partnership with a mobile operator, etc. As usual, there are some drawbacks, because the MR increases the end-to-end latency (it implies one more hop); as every handover, it can fail resulting on a massive communication loss (every UE attached to the MR goes down if the eNB–MR link goes down); it shall be integrated with the TCMS of the train, which implies ad hoc designs and extra costs. However, as we write this text, no LTE vendor has on its portfolio anything like a MR.

3.2. IEEE 802.11

The IEEE 802.11 standard, or family of standards, is a group of requirements only affecting the physical and the MAC layers. It has many amendments, which will be briefly explained later, and it has gained widespread recognition because the popular ‘WiFi’ is based on this standard.
802.11 is a half-duplex technology, which means that both sides of the communication channel can transmit, although not at the same time.

In railways this technology is present in two different ways: to deploy private networks for safety services (like CBTC) and operator-oriented services, and also to provide access to the Internet for passengers (i.e., a public network in stations and inside the trains).

It is deployed in unlicensed bands (ISM bands) at 2.4 GHz and/or 5 GHz, depending on the amendment of the standard 802.11. These ISM bands are usually very crowded, with many APs interfering. This is the great challenge here, especially in the 2.4 GHz, which is far more crowded than the 5 GHz one. However, 802.11 technologies are very resilient and they are able to work even in noisy environments. In Figure 4, it can be observed that from the 13 possible channels at 2.4 GHz only three are nonoverlapping (typically channels 1, 6 and 11 are used). In the 5 GHz band the situation is different because there are more nonoverlapping channels (see Figure 5) and we can also aggregate bandwidths in order to form 40-, 80- and 160-MHz wide (and the 20-MHz ones, too). Please note that not every channel is available everywhere, because it depends on the country where you are in.

![Figure 4. 2.4 GHz channels for IEEE 802.11b/g/n.](image)

![Figure 5. 5 GHz channels for IEEE 802.11.](image)

Since 1997 there have been a lot of amendments to the baseline IEEE 802.11 specification. Here we only focus on the most popular for any application (and also railway-related): a/b/g/n/ac. From a practical point of view, amendments a/b/g/n/ac are standards on its own, so it is very common to specify an IEEE 802.11 technology only using the amendment identifier.

The architecture of this technology is similar in many aspects to LTE. We have user devices (called terminals or clients), base stations (access points, APs) and also controllers for the base stations. On IEEE 802.11 only the protocol between a client and the AP is standard. For example,
communication between APs and controllers (and also between controllers) does not need to be standard, so vendors have developed their own noninteroperable solutions.

Amendment ‘a’ was the first of a long list and it had the same core protocol as the baseline IEEE 802.11, but introduced OFDM (orthogonal frequency-division multiplexing, with 52 subcarriers) and has a nominal data rate of up to 54 Mbps. As they are very dependent on the environment it is difficult to provide practical or average data rates. Unlike next amendments, the ‘a’ one works at 5 GHz. Amendment ‘b’ is also an evolution of the baseline IEEE 802.11 with some features in common with the latter (2.4 GHz band, direct-sequence spread spectrum [DSSS] modulation) while others not (‘b’ achieves nominal data rates of up to 11 Mbps and ‘baseline’ only 2 Mbps). This technology was the first IEEE 802.11 which was mainstream popular and also was implemented in some CBTC systems (i.e., Bombardier’s Cityflo 650) plus other general-purpose train-to-ground systems. The next step in the IEEE 802.11 evolution is the ‘g’ amendment, which is similar to ‘a’ but works in the 2.4 GHz band. IEEE 802.11n was the turning point for the IEEE 802.11: able to work in both the 2.4 and 5 GHz bands, it reaches higher data rates by using wider channels (40 MHz and also 20 MHz) and multiple input, multiple output (MIMO). The results are nominal data rates of 144 Mbps for 40 MHz-wide channels and 72 Mbps for 20-MHz ones. The last amendment covered here is the ‘ac’ one. It is disruptive with the earlier ones in the sense that it only works in 5 GHz and allows multiuser MIMO, channels of 20, 40, 80 and 160 MHz (or 80 + 80 MHz) wide, which leads to nominal data rates of 750 Mbps. Early tests of train-to-ground systems based on IEEE 802.11n and ‘ac’ proved data rates of 50 Mbps and 300 Mbps, respectively.

IEEE 802.11 is very frequently used in CBTC systems and operator-oriented services for train-to-ground communication. Obviously, ‘WiFi’ services for passengers in train stations, and also recently inside trains, implement this technology. Inside the trains there is also the need to backhaul the WiFi data from passengers out of the train with a train-to-ground link (which could be IEEE 802.11, LTE or a proprietary development). Some railway operators also use this technology to provide connectivity to two coupled consists, but sometimes the ‘train inauguration problem’ (discovery of the train topology to establish and maintain the train network) requires proprietary algorithms.

3.3. Proprietary solutions

IEEE 802.11 standard works in a very efficient way in static environments but in nomadic scenarios as well (punctual movements, at low speed, typical in walking areas like offices, homes or malls). But this is not the case for train-to-ground links where the train moves fast (subways up to 140 km/h). If a client wants to move to another AP area without losing connectivity, one AP needs to transfer the user to the other. This is generally called ‘handover’ and sometimes (mostly in WiFi context) ‘roaming’. This procedure is critical for the performance of the whole network and IEEE 802.11 solutions suffer when performing it, as they are not optimized for high mobility scenarios like subway lines or tramways. So, some modifications need to be done in the standard to make this roaming more efficient and fast. These modifications neglect the interoperability between the ‘modified’ solution and the standard one. Of course, handover optimization is not the only reason for performing
proprietary modifications to standard technologies. Security techniques, train inauguration algorithms, routing protocols, mesh architectures, etc. are sometimes the cause of these modifications.

Some examples of proprietary technologies for railways are Fluidity (developed by Fluid-mesh), Traincom (by Telefunken, now Siemens), etc.

---

**Railway certification**

Something very frequently missed when installing electronic equipment onboard is the importance of the railway certifications. Onboard equipment is exposed to many issues that threaten seriously their performance and reliability. The reason behind this phenomenon is that temperature usually ranges from −40°C to +85°C; due to the presence of high voltages and currents the electromagnetic fields are very large; trains also experience heavy vibrations, etc. Moreover, this equipment should be ready to work 24/7 for 20 years at least (trains are usually designed to last 30 years). Unless we address these threats making the onboard equipment more resilient, the impact on the performance would certainly decay. This means that no electronic device shall be installed on any train without the proper certification. The most common is the EN50155 [12], which references also almost 50 other norms. It covers temperature, power supply, humidity, vibration, shock, and electromagnetic compatibility (EMC) issues.

---

**4. Physical layer**

In this section we cover propagation and antenna issues. Propagation models are very much needed because knowing how the signal is going to behave is mandatory in order to design communication systems. From a system’s point of view it is not usually needed to know models in a very detailed way, but it is important to understand how the signal propagates, some basic principles of antenna design and how to operate with the models.

In general terms, the layer 1 (physical layer) design of a wireless system starts with some premises (frequency, maximum power, minimum signal in the receiver, desired overlap between cells, etc.) and then you place the base stations. Minimum signal strength is a parameter that is heavily dependent on the service. For instance, VoIP (Voice over IP) deployments require stronger signals than CBTC ones, as each service has its own KPIs (key performance indicators) and demands different QoS from the network. It is out of the scope of this chapter to provide precise figures about these values, because sometimes they are vendor-dependent (CBTC, for example) and sometimes they depend on subjective criteria (i.e., VoIP). Here we present some models to calculate the received power (given the transmitted power) to perform such a layer 1 design.

**4.1. Urban scenario**

Regarding urban transport we need to cover the so-called urban propagation models. These models tell us how signals behave between the transmitter and the receiver. Tramways usually
run in classic urban environment, that is, in surface areas surrounded by buildings. The modelization of these environments was started early in the sixties in Japan by Okumura and perfected in the next decade by Hata. The result is the famous Okumura-Hata model [13], where the path loss is:

\[
PL(dB) = 69.55 + 26.26 \log(f) - 13.82 \log(h_t) - a(h_r) + \left(44.9 - 6.55 \log(h_t)\right) \log(d)
\] (1)

where \(f\) is the carrier frequency in MHz, \(h_t\) and \(h_r\) denote the height in meters of transmitter and receiver, respectively, and \(d\) is the distance between them in kilometers. This model works well if the frequency is below 1.5 GHz, so for higher ones we shall use a different path loss expression. This is known as the COST 231 extension to Hata model [14]:

\[
PL(dB) = 46.3 + 33.9 \log(f) - 13.82 \log(h_t) - a(h_r) + \left(44.9 - 6.55 \log(h_t)\right) \log(d) + C_M
\] (2)

where \(C_M\) is 0 dB for medium cities and suburb areas and 3 dB for dense metropolitan areas. There are other limitations in the feasibility of these two models as they are intended for larger cells (typical in early mobile deployments) but it is still a good entry point to the design of a tram-to-wayside wireless system. However, the main limitation is that they do not work if the frequency is higher than 2 GHz. Working in higher bands require another models, like Winner II [15] which is specified for the band between 2 and 6 GHz.

### 4.2. Characterization of the tunnel scenario

The other environment where urban rail operates is the tunnel. The modeling of the propagation in tunnels is a task that is not trivial at all. There are some effects that make this characterization very difficult to accomplish and it is quite different from the general model.

As mentioned before, railways are a hostile environment for electronic devices in general. In the particular case of radio communications in tunnels, we need to add some extra problems that make communication even harder: presence of large metallic masses in movement (trains), arbitrary tunnel cross-sections, frequent changes on these sections, obstacles like cabinets, tracks, catenaries, etc.

To properly model this behavior, there are three types of models [16] based on modal-theory [17], general theory of diffraction (GTD) [18] and hybrid approaches [19]. Modal theory considers the tunnel an oversized waveguide, so in order to let signals to propagate the frequency it needs to be higher than the cut-off frequency of the tunnel [17]. The propagating signal is composed of the aggregation of many modes, each one of them with its own attenuation and phase constant. Another limitation is that properly modeling some details of the environment (a cabinet, the track, a pit, etc.) can be very arduous [16].

To overcome some of the limitations of modal-based models we use ray-tracing techniques [18], or more generally-speaking, those based on the GTD. These models consist on the
launching of a set of rays calculating the reflections and diffractions of each ray until they get to the receiver. To be done properly, significant computing resources are required, as well as a proper characterization of the surfaces of the elements in the tunnel. A good ray tracing-based tool is explained in [20].

The third group of models is the hybrid ones, which combine the advantages of the previous methods plus some corrections provided by measurements in the field. This is a hybrid approach of stochastic and deterministic models. Assuming that no model is perfect and site surveys are always needed, here we highlight an empirical method that, in our opinion, is accurate and easy enough to handle. It assumes that every tunnel is circular (if not, a correction term ‘κ’ is introduced for other shapes) and estimates the losses $\alpha$ (dB/m) as [16]:

$$\alpha = \kappa \lambda^2 \left[ \frac{\varepsilon_r}{a^3 \sqrt{\varepsilon_r - 1}} + \frac{1}{b^3 \sqrt{\varepsilon_r - 1}} \right]$$  (3)

where $\kappa$ is the shaping parameter, which varies with the cross-section of the tunnel (5.09 for circular; 4.34 rectangular; 5.13 arched), $\lambda$ is the wavelength, $\varepsilon_r$ the dielectric permittivity of tunnel walls and $a$ and $b$ are the width and the height of the tunnel, respectively. This estimation is very useful for physical layer designs in tunnel but its accuracy suffers when curves are present in the track.

To conclude this section we shall also refer to a scenario that is very much used in tunnels: the leaky feeder or leaky cable. This feeder is an alternative to discrete antennas especially useful in complex environments like winding corridors, staircases, castles and tunnels. It consists of a coax wire with many slots on its outer jacket that let the energy inside the wire to get out of it. The receiving process is the opposite, with signals getting into the slots, traveling across the leaky coax and reaching the receiver. The main benefit of this approach is to have a uniform coverage and that we do not need to perform complex calculations to estimate the coverage in the tunnel. It is true that this argument was usually given before the previous accurate models were developed, but it is still valid today in many subways. The main drawback is the high cost that has the installation of the leaky coax in the tunnel wall. The leaky feeder approach is usually followed when frequencies are under 2 GHz, but it is also possible to find leaky feeders that may work at higher frequencies.

4.3. Other scenarios: inside the train and vehicle-to-vehicle

Besides from the train-to-wayside, it is also important to emphasize other scenarios for wireless communication in urban rail: inside the vehicle (typically, for sensor networks or WiFi deployments); vehicle-to-vehicle (for the train backbone, replacing the Scharffenberger couplers); and also train-to-train (for future signaling systems). With the sole exception of the first one, the modeling of these channels is still on a very early stage of development. There
are some research papers [21] concerning this topic and some ongoing projects which will be introduced in the next section.

4.4. Antennas

Antenna engineering is a huge field that exceeds by much the scope of this chapter, but it is highly advisable to understand properly its basics. Here we cover two types of antennas: the wayside and the onboard. There are two main issues regarding both of them that require to be addressed: choosing the right antenna and placing them in the right location. Installing an antenna on a tunnel is usually arduous due to the gauge limitations. Onboard antennas tend to be omnidirectional and the installation tends to be duplicated (two cabins instead of just one) increasing the performance of the system, especially in tunnels, where the train blocks the signal when the wayside antenna is in the opposite side. Wayside antennas are directive in most of the cases but not always. It is important to highlight that some installations may require more than one antenna on a single spot. This may be needed in diversity, beamforming or MIMO setups. In these cases, the typical spacing between antennas is $\lambda$ (one wavelength) or $\lambda/2$.

5. The future of wireless communication systems for urban rail: 5G and other research directions

There are many initiatives around the world that will influence very much the future of wireless systems for urban rail. In the final section, we will describe the most relevant lines of research actually in progress. The structure is the same as in the entire chapter: services, technologies and physical layer issues.

There are two main triggers for this research: major projects, with total or partial public funding and the participation of vendors, integrators, railway operators, universities, etc. on one side, and all the investigation that is privately done by companies. Obviously, it is more difficult to know any progress coming from the latter, so we will focus on the former. In the last five years, the European Union was developing numerous research projects on wireless aspects of railways. Some good examples are Roll2Rail, Shift2Rail, SECRET, Systuf, Tecrail, and Integrail, but there are many more. Roll2Rail and Shift2Rail are more initiatives than projects, acting as ‘containers’ for projects on many railway-related topics (for example, only work package 2 of Roll2Rail is focused on wireless, but it is still a very large project). Most of these trends are very unlikely to be implemented in real in-operation systems in the short term.

5.1. Upcoming services

Once you have a CBTC system on a subway line it could help the operator to provide many other services. Some good examples are passenger information (after all, the system that knows better the location of the train is the CBTC) and many other operator-related ones. It is very difficult to determine how CBTC will evolve (especially as vendors are very secretive con-
cerning their research), but we can predict that one characteristic that could be explored is the direct communication between trains (in addition to the train-to-wayside communication already available). If we introduce direct communications between trains, a train could transmit its position and speed to the precedent one (see Figure 6), which could calculate its authority of movement using this information and also its own position and speed. This is the same as in the conventional CBTC but it could be performed in shorter times, (end-to-end delay decreases). This concept needs a reliable on-board device-to-device communication link, which sometimes could be not available (when trains are not very close or are some obstacles between them). As far as we know, the pioneer here is Alstom, with its Urbalis Fluence technology, still under test.

Figure 6. Schematic representation of a CBTC system with V2V (vehicle-to-vehicle) capabilities (in blue) plus the ‘traditional’ V2I (in green).

Another service that is gaining importance is the sensorization of onboard systems, transmission to wayside servers in order to be analyzed for operational and maintenance purposes. This methodology is very much aligned with the Internet of Things (IoT) paradigm as well as with ‘Big Data’. It is still in progress in many subways and trams around the world, but is a very promising technology for operators and rolling stock maintainers, especially as it enables the opportunity to monitor, in real-time, the status of the entire fleet. This is also key for driverless trains. But many challenges lay ahead that need to be addressed. For example, onboard bus topologies are not in practice as standardized as they should be, with many legacy networks and buses still in use; data from the sensors do not reach the wayside; there are many security issues; the storage and CPU requirements in the wayside are enormous, especially in large subway networks with many trains. Despite of this, it is very likely that in the near future the IoT paradigm will be another history of success for urban rail.

Another trend that is being discussed at the time we write this chapter is the feasibility of a wireless train network. That is, replacing the wired buses carry all the TCMS data of the trains with wireless links. This is a move very similar to offices and homes, where many wires have been replaced by Bluetooth and WiFi links. However, the challenges are many: RAMS (reliability, availability, maintainability and safety), because railways are very demanding concerning this issue; security concerns too, due to the fact that the data is far more exposed in a wireless link; availability of a mature technology able to carry the data inside the train, between vehicles, etc. This is the main purpose of WP2 of the European Project Roll2Rail [2]. Public safety services are very likely to be provided in the future over an LTE network [6]. TETRA association has publicly supported this move [6] so it is only a matter of time. Huawei,
a major LTE player, has strongly supported this view with their technology eLTE, which is able to provide many public safety features over a nonstandard LTE core.

Finally, another important challenge for the future is the integration of tramways on smart cars’ platforms. All the work that it is being carried out on autonomous cars, smart highways, smart cities, etc. will also need to integrate tramways on it.

5.2. Future technologies

At the moment, the most promising wireless technologies for use in these environments are IEEE 802.11 and 3GPP LTE. Maybe they would require proprietary improvements, but it is clear that both of them are trying to be strong at railways use cases. In the case of IEEE 802.11, there are two main lines of research: one centered in the evolution of ‘ac’, which is the ongoing ‘ax’ and another one focused on the so-called ‘WiFi Gigabit’ which works on the 60 GHz band (that is, 802.11ay, the evolution of IEEE 802.11ad). The release of the first draft for both ‘ax’ and ‘ay’ standards is expected on 2016 and 2017, respectively. 60 GHz band is more likely to be deployed inside vehicles than for train-to-wayside or train-to-train scenarios.

As it was mentioned in Section 3, 3GPP LTE is introducing more railway-related use cases, like mobile relays. Besides from this standardization effort, vendors are very likely to introduce algorithms with optimizations for radio resource management tasks, especially handovers. This would happen when the 4G (Fourth Generation) coverage on high-speed lines gets denser. Another technology that could help the market penetration of LTE in railways is the unlicensed LTE (LTE-U). It is a technology in the roadmap for 3Q 2016 for the main LTE vendors. There are three different philosophies on LTE-U: (1) LTE on unlicensed bands but as a secondary best-effort carrier for user data (in the downlink only), remaining the licensed band as a primary carrier (for both user and control data). This is also known as LAA, license-assisted access. The second offload alternative for mobile operators is to send this traffic directly to WLAN (wireless LAN) networks (LTE–WiFi link aggregation or LWA). And the third one, “standalone LTE-U” does not require operator support or licensed bands to work.

5G (fifth generation) technologies are starting its standardization. There are many governments and organizations in China, Korea, Japan, and the European Union that already have their committees, where the railway industry is also very active with some research projects that include some railway-related use cases, like the Spanish Enabling5G. The first three generations of mobile communications turned their back on railways (with the remarkable exception of GSM-R); 4G started to look at railways, and the railway industry is really implied on 5G standardization.

Finally, an interesting debate is to dedicate some spectrum for railway use (excluding the 5 + 5 MHz of the GSM-R band, which is not available for urban rail). A joint task force composed by the ETSI (European Telecommunications Standards Institute) and the UITP (International Association of Public Transport) has recently launched a very interesting technical report [22] on this topic, where it is suggested to share part of the 5.9 GHz band (dedicated to ITS) with the CBTC systems (now on ISM bands, especially at 2.4 GHz). This is an interesting move for both organizations as it could open a door for a licensed band for railways.
5.3. Physical layer research trends

There are two main challenges: the research of better channel models and the development of more suitable antennas for railways. A recent paper [23] identifies the most relevant open issues on channel modeling in railway environments (modeling the influence of vehicles, not only for train-to-wayside channels but also for train-to-train, vehicle-to-vehicle and intravehicle). In the train-to-train channels the modeling of nonstationary scenarios is still an open issue [24]. In the antenna field, the most relevant aspect is how to design MIMO antennas for the onboard, with $4 \times 4$, $8 \times 8$ setups and also to integrate them appropriately in the car body shell.

6. Conclusions

Wireless communications has revealed as a key player in railways. It has improved both safety and security, provided a better quality of service to passengers, reduced operational costs, etc. In this chapter we have explained the most important urban rail services, technologies that make them possible and finally, the physical environment where the wireless communication happens. Wireless technologies tend to be more standard every day (3GPP LTE and IEEE 802.11 lead the way). We have seen why the physical layer is still complicated to handle, especially in tunnels. Finally, we have introduced the most relevant research lines in this field to let the reader know how the near future is likely to be.

Abbreviations

4G fourth generation
5G fifth generation
AP access point
CBTC communications-based train control
CCTV closed-circuit television
DSP digital signal processing
DSSS direct-sequence spread spectrum
EMC electromagnetic compatibility
EMI electromagnetic interference
eNB evolved Node B
ERTMS European Rail Traffic Management System
ETSI European Telecommunications Standards Institute
GSM-R Global System for Mobile Communications – Railway
GTD general theory of diffraction
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IoT</td>
<td>Internet of Things</td>
</tr>
<tr>
<td>ISM</td>
<td>Industrial, Scientific and Medical</td>
</tr>
<tr>
<td>ITS</td>
<td>intelligent transportation systems</td>
</tr>
<tr>
<td>KPI</td>
<td>key performance indicator</td>
</tr>
<tr>
<td>LAA</td>
<td>license-assisted access</td>
</tr>
<tr>
<td>LTE</td>
<td>long-term evolution</td>
</tr>
<tr>
<td>LTE-U</td>
<td>LTE unlicensed</td>
</tr>
<tr>
<td>LWA</td>
<td>LTE–WiFi link aggregation</td>
</tr>
<tr>
<td>MAC</td>
<td>media access control</td>
</tr>
<tr>
<td>MIMO</td>
<td>multiple input, multiple output</td>
</tr>
<tr>
<td>MR</td>
<td>mobile relay</td>
</tr>
<tr>
<td>NAS</td>
<td>network-attached storage</td>
</tr>
<tr>
<td>O&amp;M</td>
<td>operation &amp; maintenance</td>
</tr>
<tr>
<td>OCC</td>
<td>operational control center</td>
</tr>
<tr>
<td>OFDM</td>
<td>orthogonal frequency-division multiplexing</td>
</tr>
<tr>
<td>OFDMA</td>
<td>orthogonal frequency-division multiple access</td>
</tr>
<tr>
<td>OSI</td>
<td>open systems interconnection</td>
</tr>
<tr>
<td>PIS</td>
<td>passenger information system</td>
</tr>
<tr>
<td>PMR</td>
<td>Personal Mobile Radio</td>
</tr>
<tr>
<td>PS</td>
<td>public safety</td>
</tr>
<tr>
<td>PTT</td>
<td>push-to-talk</td>
</tr>
<tr>
<td>QoS</td>
<td>quality of service</td>
</tr>
<tr>
<td>RAMS</td>
<td>reliability, availability, maintainability and safety</td>
</tr>
<tr>
<td>RF</td>
<td>radio frequency</td>
</tr>
<tr>
<td>RRM</td>
<td>radio resource management</td>
</tr>
<tr>
<td>SC-FDMA</td>
<td>single-carrier FDMA</td>
</tr>
<tr>
<td>TCMS</td>
<td>train control management system</td>
</tr>
<tr>
<td>TETRA</td>
<td>terrestrial trunked radio</td>
</tr>
<tr>
<td>UE</td>
<td>user equipment</td>
</tr>
<tr>
<td>UITP</td>
<td>International Association of Public Transport</td>
</tr>
<tr>
<td>UMTS</td>
<td>Universal Mobile Telecommunications System</td>
</tr>
<tr>
<td>V2V</td>
<td>vehicle-to-vehicle</td>
</tr>
<tr>
<td>VoIP</td>
<td>Voice Over IP</td>
</tr>
<tr>
<td>WLAN</td>
<td>wireless LAN</td>
</tr>
</tbody>
</table>
Author details

Juan Moreno García-Loygorri*, José Manuel Riera and Carlos Rodríguez

*Address all correspondence to: juan.moreno@metromadrid.es

1 Engineering and R & D Department, Metro de Madrid, Madrid, Spain

2 Departamento de Señales, Sistemas y Radiocomunicaciones, Universidad Politécnica de Madrid, Madrid, Spain

References


