We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Transmission-Based Signaling Systems

Cesar Briso-Rodríguez,
Juan Moreno García-Loygorri and Lei Zhang

Additional information is available at the end of the chapter

Abstract

In this chapter, we describe the principal communication systems applied to the transmission-based signaling (TBS) systems for railways. Typical examples are communication-based train control (CBTC), European Rail Traffic Management System (ERTMS), and distance to go (DTG). Moreover, to properly address some of the challenges that need to face these systems, we will provide a deep insight on propagation issues related to all the environments (urban, suburban, rural, tunnel, etc.). We will highlight all the communication-related issues and the operational as well. Finally, a detailed survey on the directions of research on all these topics is provided, in order to properly cover this interesting subject. In this research, hot topics like virtual coupling are explained as well.

Keywords: communications, engineering, railways, services, signaling, ETCS, CBTC, GSM-R, LTE, CBTC, propagation, tunnels

1. Introduction

The new signaling systems used for the control and operation of high-capacity and high-speed railways demand the use of advanced communications to guarantee their operation and safety, decrease maintenance costs, and improve the user experience. For this reason, communications in the transport environment, especially in the railway environment, have developed dramatically in the last decade.

The communications applied to railway signaling have special requirements that are very dependent on the characteristics and performances of the railway system. Consequently, it is necessary to know in detail the characteristics and performance of the transportation system and apply very strict requirements to the communication system.

This chapter describes the most important characteristics of current signaling systems and the communication systems used in combination with them.

2. Railway signaling systems

2.1. ERTMS

The European Rail Traffic Management System (ERTMS) is the most advance signaling system available nowadays. The system was developed for high-speed trains. When the train goes faster than 200 km/h, the driver is unable to see the wayside signals, so it is necessary to transmit this information to the train cabin. Moreover, this information needs to be continuous in order to have an increased safety. European Rail Traffic Management System (ERTMS) was designed to solve these two requirements and one more, the need of interoperability, letting trains to cross borders without being equipped with all the regional signaling systems (EBICAB, TBL, AWS, ASFA, LZB, and a very large etcetera). Today, there are 38 countries (most of them in Europe, but some of them not, like China, Saudi Arabia, South Korea, Taiwan, and Australia, among others), 62,000 km of track and 7500 vehicles, equipped with ERTMS [1] which is a good measure of the popularity of this system. Therefore, ERTMS leads to an increase of the interoperability, the safety, and a reduction of the costs (only one system is required).

ERTMS was specified in the European Union Agency for Railways (ERA) in a task force composed by railway manufacturers and operators (UNISIG). ERTMS is structured in three levels: level 1 is a punctual ATP with balises or loops in the track, placed 500 meters in advance of the signals. These balises/loops can provide static or variable data. Level 2 is similar to a distance-to-go (DTG) system, and the communication is now bidirectional (through the GSM-R radio), so trackside signals could be removed but not track circuits. Finally, level 3 is a "moving block" system, and both track circuits and signals could be removed. Therefore, levels 2 and 3 require GSM-R and a network of active transmitters (balises) in the track. In **Figure 1**, there is a schematic depiction of the three ERTMS levels.

Moreover, ERTMS is able to provide the driver with other information of interest, like transitions between supply phases, viaducts, and tunnels. In these areas, the train shall not stop.

2.2. CBTC

There are five grades of train automation (called GoA levels, for Grades of Automation, as it is depicted in the EN 62290 standard [2]): GoA 0, which means no automation at all; GoA 1, automatic train protection (ATP), where the train driver brakes and accelerates but under the constraints given by the system, which protects the train; GoA 2, automatic train operation (ATO), where the system regulates the speed of the train and the driver is still in the cabin doing auxiliary functions like opening and closing doors—and many others; GoA 3, driverless train operation (DTO), where no driver is needed but some staff is needed onboard the train; and, finally, GoA4, unattended train operation (UTO) mode, no staff is needed onboard.
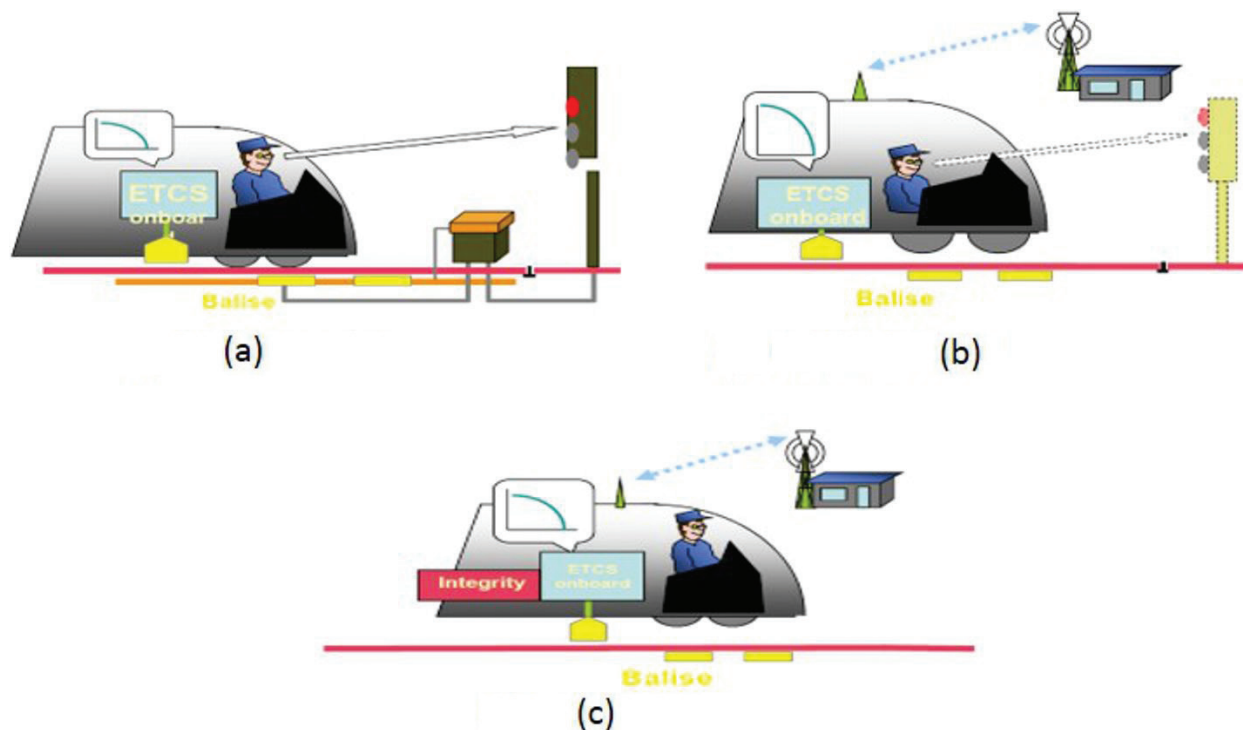
Figure 1. ERTMS levels. a) ERTMS level 1, b) ERTMS level 2, c) ERTMS level 3.

Meanwhile, metro systems have achieved all the GoA levels; mainline and high-speed trains are still in GoA 1 (or ATP). The reason behind this is that implementing higher GoA levels does not imply a significant benefit for a high-speed train operator as it is for a metro operator (headways very short, operation intensive in workforce, etc.) (**Figure 2**).

Therefore, ATO functions (speed regulation, special maneuvers, door control, etc.) are not safety related. Only ATP is, but an ATO system means 8–10% increase in the regularity of the trains compared to a human-driven ATP (plus the extra comfort for passengers due to the smoothness of speed curves).

ATP can be discrete or continuous: in the first case, the driver only receives "protection" from the system (speed monitoring and emergency brake if needed) in some certain points along the track. In the continuous mode, this is done at every single point of the track. UIC's recommendation is to implement continuous ATP systems when the maximum speed is higher than 220 km/h or the headway between trains is less than 120 seconds. Anyway, the trend is to implement this type of ATP on behalf of the safety.

Regarding the technology behind ATP functionalities, we can find for different ones: speed codes, distance to go, virtual track circuits, and moving block. In speed codes, the trackside system transmits to the onboard the maximum speed that it can achieve on a given track circuit. In distance-to-go systems, the train knows better its position in the track, which is transmitted to the next train, and this better knowledge of the position of the next train leads to a shorter distance between them. In virtual track circuits (or fixed block), the position of the train is known with a higher accuracy (less than the physical track circuit) by using odometry techniques. As it happened before, this
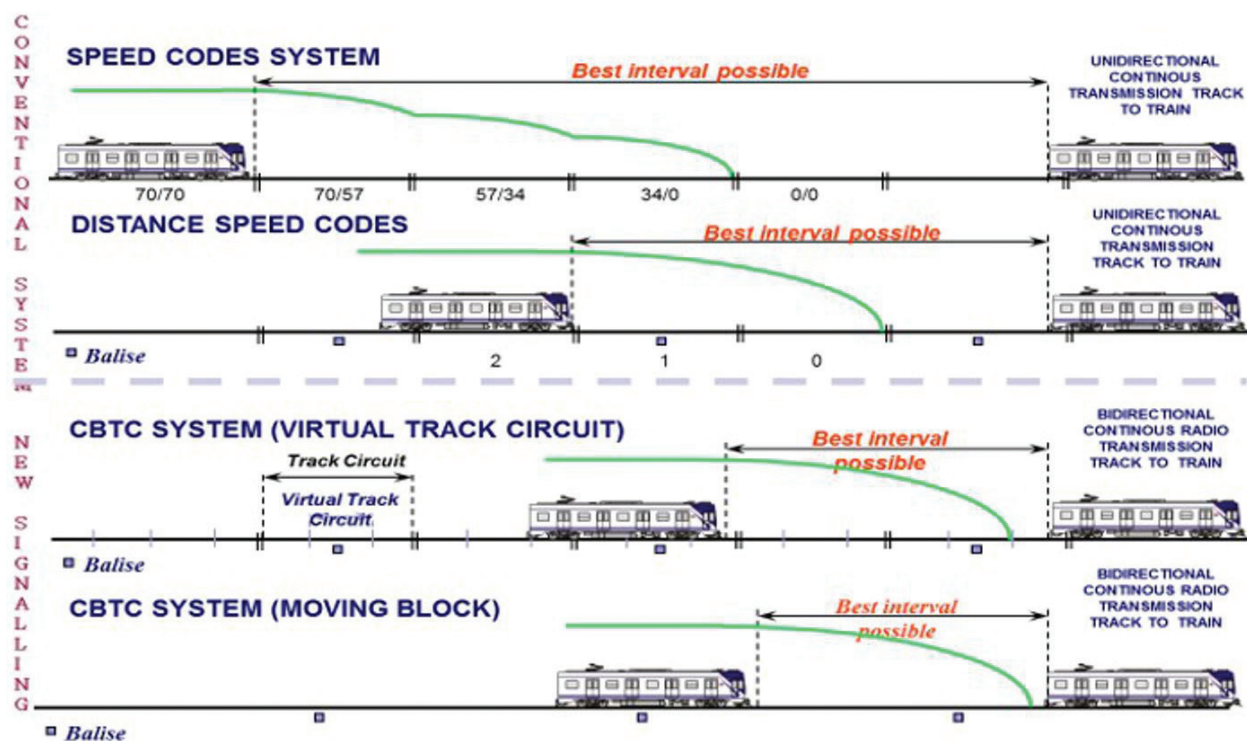
Figure 2. Comparative depiction of ATP technologies.

accuracy leads into a shorter headway. Finally, "moving block" techniques imply estimating the position of the train only using balises and odometry (no track circuits are needed), so the fragmentation due to track circuits is removed. This moving block system is the most advanced until now, and it represents 15–20% more capacity than the DTG ones [3].

DTO and UTO modes are generally implemented over communication-based train control (CBTC) systems. However, this is not a strict requirement, but an industry trend. CBTC is based on two pillars: the bidirectional communication between train and trackside equipment and the accurate positioning of the train. Train positioning is not standardized in CBTC, but it is very common to have redundant methods, like Doppler radars or tachogenerators. All CBTC vendors have their own implementation with some differences between them but with many common aspects (very high MTBFs, high positioning accuracies, etc.).

CBTC systems are able to provide headways (zero dwell, though) shorter than 60 seconds. This figure could be limited by external issues to the CBTC like delays in rail switches and other operational functions. An important remark is that there is no merit at all in having a short headway at a slow speed. The target is to have both high average speeds and short headway between trains (with no decrease in the safety).

2.3. CTCS

The Chinese Train Control System (CTCS) is a specification of train control systems of the People's Republic of China. The CTCS is based on ERTMS, and some forms are with the

European Train Control System compatibility. According to the functional requirements and equipment configuration, the application level of CTCS has divided into 0–4 levels [4] to define the equipment composition, information transmission, applicable section, track occupancy check, control mode, occlusion way, etc. in different levels.

CTCS-0: It consists of the existing track circuits, universal cab signaling, and train operation supervision system. With level 0, wayside signals are the main signals, and cab signals are the auxiliary signals. It is the most basic mode for CTCS. It is unnecessary to upgrade the wayside systems for CTCS level 0. The only way to realize the level 0 is to equip with the onboard system. CTCS level 0 is only for the trains with the speed less than 120 km/h.

CTCS-1: It consists of the existing track circuits, transponders, and ATP (automatic train protection) system onboard the train. It is used for trains with speed of 120–160 km/h. On this level, the block signals can be removed, and train operation and security are based on the onboard system, ATP, which controls the principal functions of the train: maximum speed of the track and doors opening. Transponders must be installed on the line. The requirements for track circuit in blocks and at stations are higher than that in the level 0. The control mode for ATP could be the distance to go or speed steps.

CTCS-2: It consists of digital track circuits (or analog track circuits with multi-information), transponders, and ATP system. It is used for the trains with the speed higher than 160 km/h. There is no wayside signaling in block for the level 2 anymore. The control mode for ATP is the distance to go. The digital track circuit can transmit more information than analog track circuit. ATP system can get all the necessary information for train control. With this level, fixed block mode is still applied.

CTCS-3: It consists of track circuits, transponders, and ATP with GSM-R. In the level 3, the function of the track circuit is only for train occupation and train integrity checking. Track circuits no longer transmit information concerning train operation. All the data concerning train operation information is transmitted by GSM-R. GSM-R is the core of the level. At this level, the philosophy of fixed block system is still applied.

CTCS-4: Moving block system function can be realized by this highest level. The information transmission between trains and wayside devices is made by GSM-R. GPS or transponders are used for train position. Train integrity checking is carried out by onboard system. Track circuits are only used at stations. The amount of wayside system is reduced to the minimum in order to reduce the maintenance cost of the system. Train dispatching can be made to be very flexible for the different density of train operation on the same line.

The levels 2, 3, and 4 are backward compatible with the smaller level. The CTCS-3 is functionally equivalent to the ETCS level 2. Driving in migratory space distance or absolute braking distance of CTCS-4 is also specified in ETCS level 3.

CTCS-3/ETCS and ERTMS level 2 are to be used in the People's Republic of China on the nearly 1000 km long high-speed line between Wuhan and Guangzhou. The mid-2007 awarded contract to haul equipment has a volume of 66 million euros (for installation, delivery, testing, and commissioning) and consists of the line equipment and the equipment of
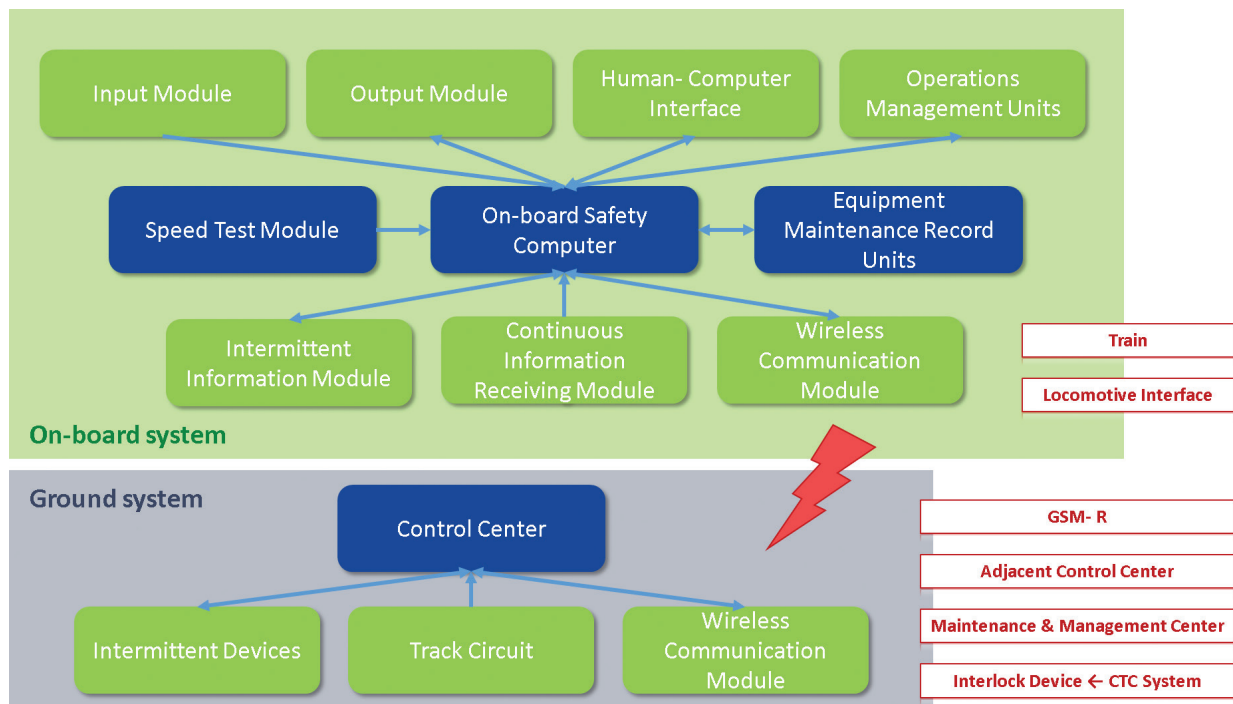
Figure 3. CTCS system architecture.

60 high-speed trains. The system has been commissioned in January 2010. The CTCS has the following features:

- **Openness**. The ETCS specification is the standard recognized by the European Union and the International Union of Railways; thus, all ETCS equipment suppliers can produce CTCS equipment in accordance with this standard.

- **Interoperability**. Since all ETCS devices are produced in accordance with uniform technical specifications, so different manufacturers of equipment can be conveniently integrated or even use directly.

- **Compatibility**. Although the vehicle equipment is different in locomotives with different levels of ETCS systems, the locomotives can operate on lines with various levels.

- **Scalability**. On the basis of the original equipment in low-level CTCS system, it can be easily upgraded to a higher level by adding some new equipment (modules); the original train control equipment can be continually used in high-level systems (**Figure 3**).

3. Communications in railway environment

The railway environment is very complex and represents a major challenge for communication systems. For this reason, to deploy a communication system in this environment, it is necessary to understand their special characteristics. The main features of the railway environment are summarized below:

- Speed influence with trains traveling up to 400 km/h.

- Breaking distance of trains up to 3 km.

- Environment of the track: cuts, hills, and crossed valleys. Tunnels >30% of the track.

- Base stations must be close to the track to access the infrastructure of the railway.

3.1. Requirements of railway communications

The general requirements of the communications network are very special to allow the improvement of the railway operation. Below, we summarize the main objectives of the communications network:

- The network must provide a high quality of service (QoS).

- The design must guarantee a high redundancy of coverage.

- Trains must have antennas and repeaters to guarantee radio coverage.

- Signaling and control data must be provided with security and passenger services.

- There are important limitations in the frequency bands that can be used.

Also, communication networks for railway applications must provide special services to enable railway operations and ensure safety. The main services are summarized below:

- Group broadcast call: multipoint link, emergency call.

- Fast call setup ⇒ setting calls <2 seconds (Eirene).

- The use of a special or proprietary frequency band.

- Improved equalization ⇒ mobile traveling at 350 km/h.

- Location-based addressing—training calls originated on trains depending on location.

- Functional addressing ⇒ numbers for each function of the train.

The principal specifications of the network based on the European Signaling System are summarized in **Table 1** [5]. These speculations are used to validate the communications network in

ETCS QoS parameter	Requirement
Connection establishment	<8.5 s (95%), ≤10 s (100%)
Transfer delay (TD)	≤0.5 s (99%)
Network registration delay (NRD)	≤30 s (95%), ≤35 s (99%), ≤40 s (100%)
Connection loss rate (CLR)	<10^{-2}/h
Transmission interference (TI)	<10^{-2}

Table 1. Main requirements of railway communication system.

a railway environment. The compliance of these specifications is verified by measuring along the track and averaging every 100 m. This guarantees the quality of service and reliability.

3.2. Network architecture

In order to guarantee the quality and reliability of the communications in the railway environment, it is necessary to design the networks to ensure an excellent coverage in percentages above 99% of the time and locations:

- Dual-layer or single-layer coverage with strong overlap.

- Two totally independent layers: 2 MSC, BSC, TRAU.

- Two totally independent redundant systems for the fixed network.

- Hot standby: change from layer A to layer B without loss of call.

One of the main requirements of the GSM-R network is to achieve a high quality of service. For this, two different strategies can be used:

- Two complete networks overlapped.

- A single network with a strong overlap of the radio coverage of base stations (**Figure 4**).

In the first configuration, two overlapped networks are used. This allows that if one or several base stations of one layer fail, an automatic roaming occurs to the other layer, thus ensuring communications. In the second design, a network with a double number of base stations is used. This allows a strong overlap of the radio coverage, so that if one cell fails, neighbor cells compensate its coverage automatically. In both cases, a high redundancy is achieved that guarantees the operation of the system.
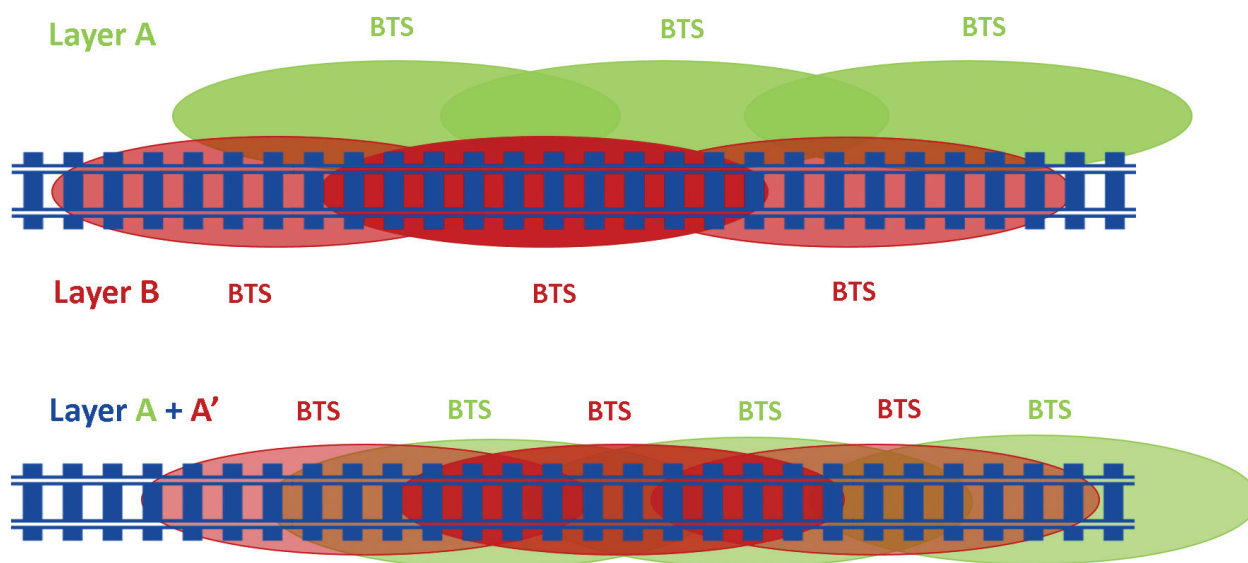
Figure 4. Radio coverage for high QoS and redundancy.

### 3.3. Communication problems in railway environments

As already mentioned, railway environment is very complex, and it is necessary to provide communications in a variety of environments and conditions that make it very difficult to plan and deploy a radio communication system with the quality of services demanded by a signaling system. The different communication systems used on railway environment have reported several problems on the different test trials made. The principal problems reported are summarized as follows:

- Failure of handovers: at high speed (250–350 km/h), some handovers can be lost. These failures depend also on the direction of the train. The main cause of the loss of handover is the little overlap of the radio coverage of neighbor cells. For this reason, the train loses coverage too fast, without completing the handover that requires 6 seconds.

- Ping-pong: this phenomenon is due to the realization of several handovers with the neighboring cell. It happens because the overlap with the neighboring cell is not uniform, so the coverage of the cell is better in some points and worst in others. For this reason two or more handovers are made when only one should happen. This significantly deteriorates the QoS and causes link losses.

- Loss of data during the handover: it is due to the cut of the link when the handover takes place between two neighbor cells. Depending on the type of communication system can have a length from some millisecond to 500–700 milliseconds.

- Bit error rate (BER) degradation with speed: communication systems employ channel equalizers to mitigate propagation problems and reduce transmission errors. However, these equalizers require the channel to remain stable for a short period of time. For this reason at high speeds, the channel varies too fast, and transmission errors occur, which increases the BER.

Another important aspect of railway communications is the problem of the railway environment. In this environment, providing a radio coverage with high QoS is complicated due to the characteristics of the environment, where cuts, trenches, tunnels, and viaducts are frequent and where the train travels at a high speed. Particularly, the study of communications in tunnels, which can currently account for up to 50% of the route of a railway line, is important.

Communication in tunnels is a major challenge for radio communication systems; so in this environment, it is necessary to use solutions other than those used in open areas. Basically, two different techniques can be used to provide coverage in tunnels: antennas and leaky feeder.

The first solution is based on the use of antennas. This solution requires much more engineering work, but it has less deployment and maintenance cost. First of all, it is necessary to ensure a careful planning of the radio coverage in this environment. Then, it is necessary to model the propagation in detail obtaining a propagation model adjusted to the characteristics of each tunnel and the communication system used. Then, it is necessary to validate the
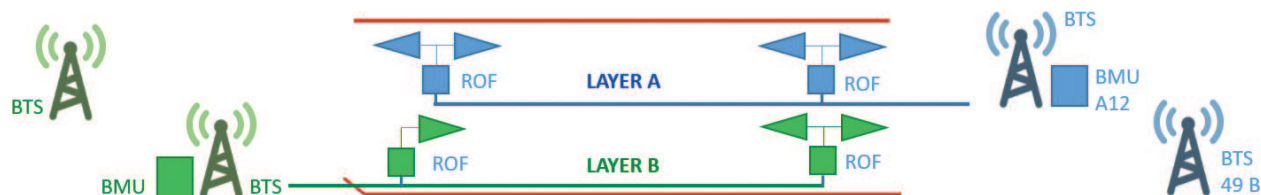
Figure 5. Radio coverage of a tunnel with repeaters and antennas.
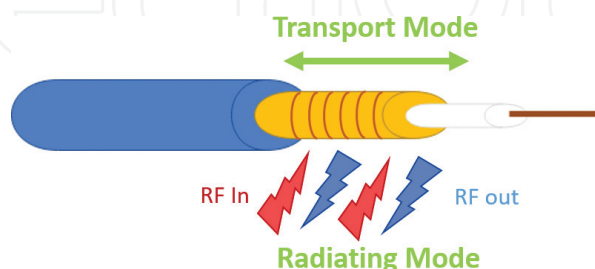
Figure 6. Leaky feeder.

approach with measurements on the target tunnel, and finally a solution is custom designed for each tunnel. In **Figure 6**, a solution for radio coverage of a railway tunnel is shown. In this case, the solution is based on the use of repeaters with technology radio over fiber [6] connected with an optical fiber. Each repeater retransmits the signal of the nearest base stations inside the tunnel, so the radio coverage is extended along the tunnel. In the case of **Figure 6**, two repeaters for each network layer are used to provide redundancy inside the tunnel. The handover is made when the train goes outside the tunnel. This solution requires the installation of a repeater every 1–2 km of the tunnel (**Figure 5**).

The second possible solution for radio coverage in tunnels is the use of leaky feeders [7]. In this case, a low loss coaxial with slots on the shieling (see **Figure 6**), is deployed along the tunnel wall. The antennas of previous case can be replaced by the coaxial. The result is a much stable radio coverage with a power level stable at a distance of 2 m from the cable. The disadvantages are the high cost of this solution that also needs a shorter distance between repeaters.

4. Wireless standard for railway

4.1. GSM-R

GSM-R system description is essentially the same system as the GSM but with railway-specific functionalities. It uses a specific frequency band around 800/900 MHz, as illustrated in **Figure 7**. In addition, the frequency bands 873–876 MHz (uplink) and 918–921 MHz (downlink) are used as extension bands for GSM-R on a national basis, under the name extended GSM-R (E-GSM-R). GSM-R is typically implemented using dedicated base stations (BSs) close to the rail track.
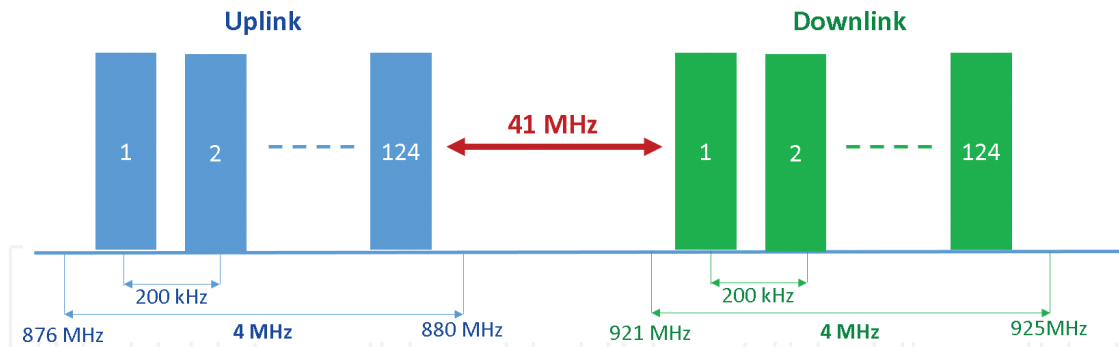
Figure 7. GSM-R frequency band and channels.

The distance between two neighbor BSs is 7–15 km; in China, it is 3–5 km because redundancy coverage is used to ensure higher availability and reliability. GSM-R has to fulfill tight availability and performance requirements of the HSR radio services.

The GSM-R network serves as a data carrier for the European Train Control System (ETCS), which is the signaling system used for railway control. The ETCS has three levels of operation and uses the GSM-R radio network to send and receive information from trains. On the first level, ETCS-1, the GSM-R is used only for voice communications. On the other two levels, ETCS-2 and ETCS-3, the GSM-R system is used mainly for data transmissions. The GSM-R is very relevant to ETCS-2 and ETCS-3, where the train travels at a speed up to 350 km/h, and it is thus necessary to guarantee a continuous supervision of train position and speed. When the call is lost, the train has to automatically reduce the speed to 300 km/h (ETCS-1) or lower.

The most typical HSR-specific services offered by GSM-R are as follows [8]:

- Voice group call service (VGCS): VGCS conducts group calls between trains and BSs or conducts group calls between trackside workers, station staff, and similar groups.

- Voice broadcast service (VBS): The BS broadcasts messages to certain groups of trains, or trains broadcast messages to BSs and other trains in a defined area. Compared to VGCS, only the initiator of the call can speak in VBS, and the others who join the call can only be listeners. VBS is mainly used to broadcast recorded messages or to make announcements in the operation of HSR.

- Enhanced multilevel precedence and preemption (eMLPP): eMLPP defines the user's priority and is used to achieve high performance for emergency group calls.

- Shunting mode: Shunting mode provides an effective means of communication to a group of personnel who are involved with a shunting operation, which regulates and controls user access to shunting communications (a link assurance signal used to give reassurance to the train driver).

- Functional addressing: A train can be addressed by a number identifying the function for which it is being used rather than a more permanent subscriber number.

- Location-dependent addressing: Calls from a train to certain functions can be addressed based on the location of the train as the train moves through different areas of BSs (**Figure 8**).
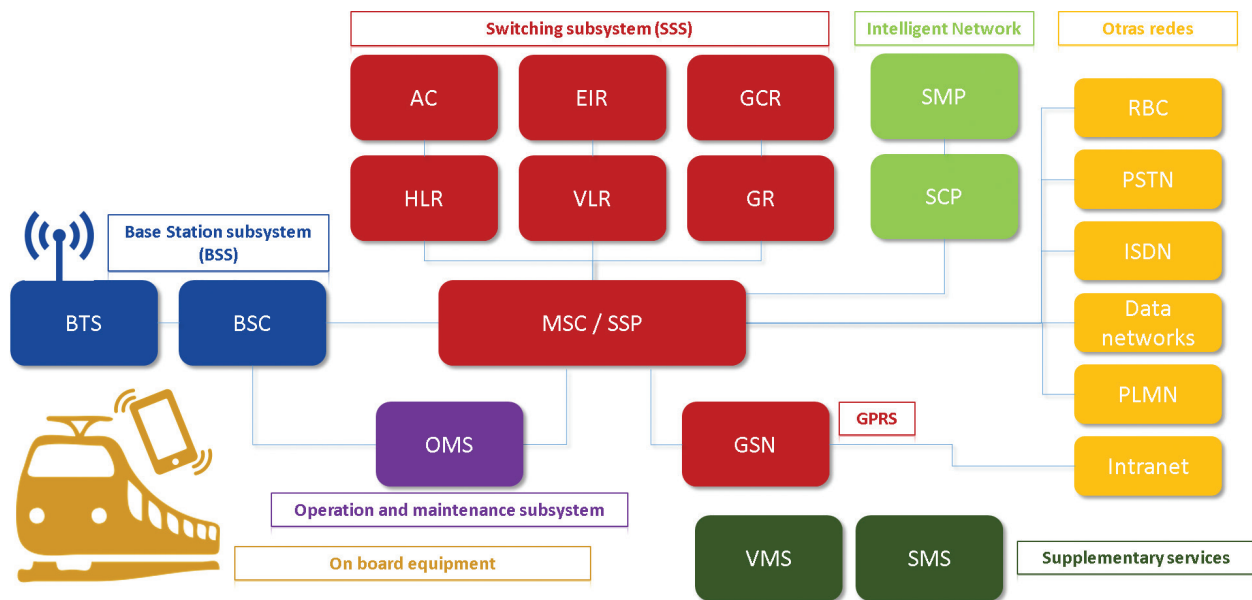
Figure 8. GSM-R network.

4.1.1. GSM-R limitations

Although the popularity of GSM-R is still growing, increasing interference from public networks is hampering the use of GSM-R, while the assigned radio frequencies limit its capacity. Several limitations are summarized in the following:

- **Interference**: The interference between GSM-R and other public networks increases because both railway and public operators want to have good coverage along the rail tracks. Instead of cooperating in network planning, railway and public operators fight for the coverage. The interference could result in severe impairment of voice and data communications as well as network loss over several hundred meters of track. Theoretically, such interference can be avoided if public operators do not use frequency bands adjacent to those of GSM-R for the areas close to rail tracks; however, this is not well implemented in practice. In the future, interference may increase owing to the growth of GSM-R network deployment and the potential growth of public networks.

- **Capacity**: The 4 MHz bandwidth of GSM-R can support 19 channels of 200 KHz width. This is sufficient for voice communication, as voice calls are limited in time and do not occupy resources continuously. However, the current capacity turns out to be insufficient for the next-generation railway system, where each train needs to establish a continuous data connection with a radio block center (RBC) and each RBC connection needs to constantly occupy one time slot. The radio capacity can be increased by using more spectrum resources.

- **Capability**: As a narrowband system, GSM-R cannot provide advanced services and adapt to new requirements. The maximum transmission rate of GSM-R per connection is 9.6 kb/s, which is sufficient only for applications with low demands; message delay is in the range of 400 ms, which is too high to support any real-time application and emergency communication [9].

4.1.2. Future of HSR communications

The future services of HSRs such as real-time monitoring require a wideband system to have larger data rate and short delay. Due to the above limitations, GSM-R must eventually evolve to eliminate the revealed shortcomings [10]. Long-term evolution (LTE)-R, which could be based on the LTE standard, is a likely candidate to replace GSM-R in the future for the following reasons:

1. LTE has many advantages over GSM in terms of capacity and capabilities. As a fully packet-switched–based network, LTE is better suited for data communications.

2. LTE offers a more efficient network architecture and thus has a reduced packet delay, which is one of the crucial requirements for providing ETCS messages.

3. LTE has a high-throughput radio access, as it consists of a number of improvements that increase spectral efficiency, such as advanced multiplexing and modulation.

LTE is also a well-established and off-the-shelf system and provides standardized interworking mechanisms with GSM [11].

4.2. TETRA

TETRA is an ETSI standard [12] for public safety, very popular in police, civil defense, border control, fireman, etc. It allows point-to-multipoint and point-to-point (direct) communications, of both voice and data. It is a very robust system, popular in urban railways but very limited because it is a narrowband system (up to some kbps). This technology usually works in the 380–470 MHz band, which needs a smaller number of base stations than GSM-R (which works in the 900 MHz band) for a given distance. TETRA has four 25 KHz channels, which is not enough for high data rates. The industry trend is to replace TETRA systems with LTE.

4.3. IEEE 802.11

It is almost impossible to gather all the complexity of the IEEE 802.11 standards [13] in so little space. IEEE 802.11 is a worldwide famous standard (its popular name is "Wi-Fi") which constitutes one of the most successful communication systems ever. It specifies both the physical and the MAC layer and has had many amendments since its beginnings. A summary of the most important features of the main standards is presented in **Table 2**.

| Characteristic	IEEE 802.11b	IEEE 802.11g	IEEE 802.11n	IEEE 802.11ac
Frequency band	2.4 GHz	2.4 GHz	2.4 and 5 GHz	5 GHz
Channel width	20 MHz	20 MHz	20 and 40 MHz (only in 5 GHz)	20, 40, 80, and 160 MHz
Transmission tech.	SISO	SISO	MIMO	MIMO
Max. rate (Mbps)	11	54	433	1300

Table 2. Main characteristics of the IEEE 802.11 family of standards.

IEEE 802.11 technologies are widely used in railways for two main reasons: they work in unlicensed bands (ISM, ETSI, etc.) and its low cost. The two main applications for them are train-to-ground communications (in workgroup bridge mode) and access to the Internet for passengers (client: access point (AP) mode). It is worth mentioning that the workgroup bridge mode is not standardized, so it has some proprietary aspects that could break the interoperability that is assumed in the usual client: AP mode.

4.4. 4G/5G

In this section, we will cover the feasibility of the fourth and the fifth generation of mobile communications to provide railway signaling services. LTE (long-term evolution), the most relevant 4G technology, is standardized by the 3GPP. LTE has a flat architecture, to suppress bottlenecks and also to achieve low end-to-end delays. LTE vendors like Huawei, Nokia, or ZTE have a lot of interest in railways, and some years ago, the UIC identified LTE as a key technology for railway communications [14]. Explaining LTE in some detail is out of the scope of this chapter, but there are very good references in the literature worth reading [15, 16]. The most remarkable features of LTE are that it is a full-IP technology; is very flexible that can work either in TDD or FDD mode; and supports carrier aggregation and several MIMO configurations. Its general architecture is shown in **Figure 9**, with an access stratum, composed by eNBs, and the core. More details can be found in [16].

There are still many challenges to have a successful LTE deployment for signaling. The first one is the spectrum, because almost everywhere there is no frequency band allocated for this purpose. Almost every single LTE band is licensed, so the railway operator would need to reach an agreement with a mobile operator in order to provide this service with the needed requirements for signaling services. Among the other challenges, we can find the cybersecurity (almost every single signaling network is owned by a railway operator) and the cost (generally speaking, LTE is more expensive than other technologies that do not need a core, but this is a controversial statement).
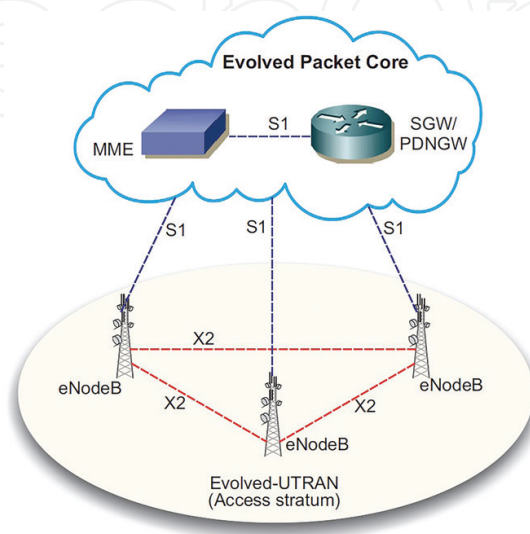
Figure 9. LTE architecture.

Regarding 5G mobile communications, it is still a technology not in operation but being massively researched. It is very promising not only for railways but also for the general industry and civilian uses. For railways, it is expected that 5G will be an enabler for the next generation of services, like "virtual coupling" (which needs low latencies and high reliability), "remote driving" (both high capacity and low latencies needed), and a large etcetera.

5. Wireless channel in railway

On conventional railway communication, the high penetration losses caused by the train's body, the harsh electromagnetic environment, and diverse scenarios have brought many difficulties. In recent years, the rapid development of HSTs makes demands on broadband transmission, high capacity, and reliability services regardless of the locations and speeds. Thus, modern wireless communication systems for railway have to overcome additional challenges, such as fast handover, fast travel through diverse scenarios, and large Doppler spreads, resulting from the high speed of the train (>250 km/h) [17].

In this section, the basic knowledge of the large-scale fading and small-scale fading of radio propagation is presented. Also, the channel characterization and modeling for railway are discussed to provide a brief overview of the research status for wireless channel in various railway scenarios.

5.1. Radio propagation mechanism

The fading in mobile communication can be classified into two main groups: large-scale fading and small-scale fading. Literally, the large-scale fading denotes a large distance (hundreds of wavelengths) that the radio signal traveled. Generally, the slow dissipation of energy due to the separation of transmitter and receiver within such a large distance is defined as path loss. In the meantime, the penetration or diffraction around large objects in the actual propagation channel results in the fluctuations on the local mean power, which are the so-called large-scale fading (or "shadowing"). Small-scale fading is used to describe the self-interference of the arrived signals from different paths with different amplitudes, delays, and phases at the receiver over a short period (in the order of 10 ns depending on mobility) or a travel distance on the same order of the wavelength. **Figure 10** illustrates the large-scale fading that superimposed with the small-scale fading.

When radio signal travels through the wireless channel in the railway environment, the presence of trees, people, buildings, mountains, and other obstructions results in the signal
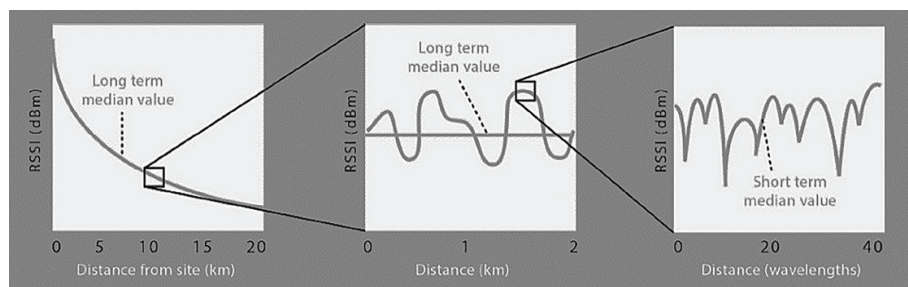
Figure 10. Path loss, shadow fading, and small-scale fading.

undergoing many kinds of propagation effects that result the radio signal reaches at the receiving antenna by two or more paths. This multipath denotes that the radio signal is transmitted from the transmitter and then "interacted" (reflected/diffracted/scattered) with the objects (also known as "scatterers") presented in the channel. The multiple copies of the transmitted signal travel from several different paths and then arrive and combine at the receiver. This combination can be either constructive or destructive that can cause random fluctuations in the received signal's amplitude, phase, and angle of arrival, which is termed as multipath fading.

5.2. Channel characterization and modeling for railway

A channel model is known as an abstract and simplified approach to mathematically demonstrate the main characters of an actual channel and evaluates the influences on the performance of a specific wireless technology in this channel. Fundamentally, the modeling approaches fall into two categories: physical models and analytical models. The physical models are established on the basis of electromagnetic wave propagation and independent of antenna configurations, such as antenna pattern, number, etc. It models the bidirection propagation between the transmitter and receiver based on the measurement or simulation in a specific scenario, which means that it changes as the locations of transmitter and receiver changes, whereas the analytical models mathematically/analytically reproduce the statistical properties of the MIMO matrix in the corresponding domain, which is limited by the computation complexity. The physical channel models are widely adopted and generally classified into two types: (1) **Deterministic modeling** characterizes the physical propagation in a completely deterministic way, i.e., calculating the received signal from knowledge of geometry, the electrical properties of the medium of propagation, cross section of objects, and antenna radiation pattern. Therefore, the deterministic models are approximately exact. However, the deterministic approach, such as ray tracing, is subject to the computation complexity and complication of environment reconstruction. (2) **Stochastic modeling** is void of exact geometrical assumption and provides the statistical manner and conditional dependencies between different channel parameters with less computation resources. This kind of channel model is extracted from the large amounts of measurement data, which is practically useful to present the general characteristic of a typical scenario. The propagation characterization in the specific scenarios is modeled by the statistical analysis of the channel impulse response (CIR), extracting the key parameters, such as path loss, delay spread, angular spread, etc., to parametrize the channel (**Figure 11**).

The wireless channel characteristics in railway are of importance for train control and passenger services. The deep and comprehensive knowledge of the propagation channel is the premise of the communication system design. A convincible solution to support wireless communication system design for railway requires the characterization of four different channels, as shown in **Figure 10**: BTS-to-train (B2T)/train-to-ground (T2G), carriage-to-carriage (C2C), train-to-train (T2 T), intra-carriage (InC), and optionally satellite communication train-to-satellite (T2S). In the last years, research on communication systems for railway has been focused specially on the T2G communications in different environments, as presented in [18, 19] with the complete surveys on results and challenges of high-speed trains (HST) and metro trains in tunnels, respectively.
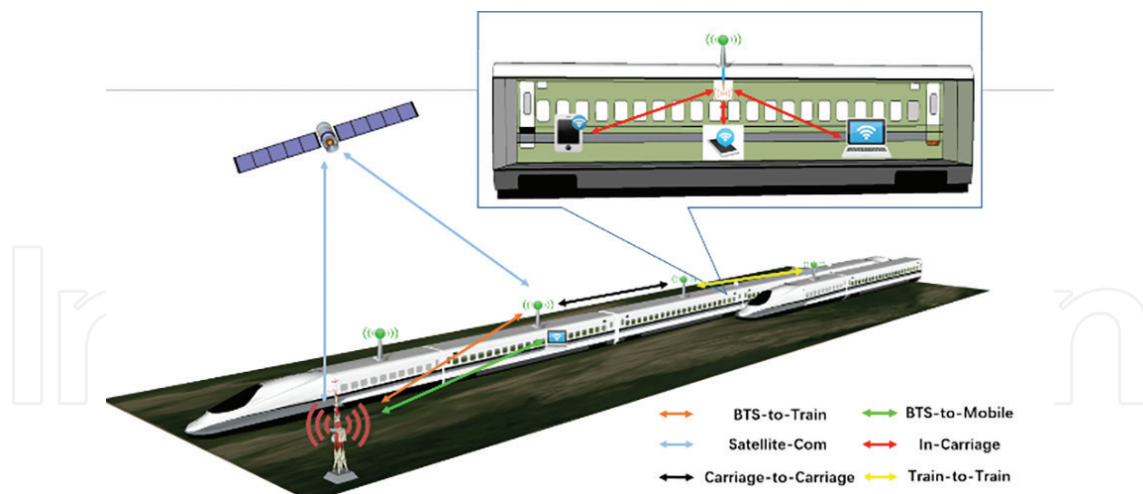
Figure 11. Wireless links in railway.

| Scenario	Path loss	Shadowing	K-factor	Delay	Doppler	AoA/AoD	EoA/EoD
Open space	M & S	M & S	M & S	M & S	M & S	M & S	—
Viaduct	M & S	M & S	M	M	M	M	—
Cutting	M & S	M & S	M	M	M	—	—
Metro	M	M	M	M	—	—	—
Intra-carriage	M	M	M	M	—	—	—

Table 3. The research status of the channel characterization and modeling for railway.

In recent years, the channel characterization and modeling have become the focus of a lot of studies based on the measurement campaigns. Subject to the robustness, scalability, hardware redundancy, and traceability of the particular hardware, measurement-based approaches cannot keep the spatial consistency which is crucial in time-varying channel modeling in railway. Thus, some deterministic channel models have been proposed based on the ray tracing. Here, we briefly review and classify the measurement (**M**)-/simulation (**S**)-based research status for railway communications according to the scenarios, approaches, and key channel parameters, including path loss, time delay, and so on, as shown in **Table 3** [17, 20].

6. Future trends in signaling systems

It is always hard to predict the future, even the more immediate one. Signaling, like the entire railway sector, goes in the same direction: more efficiency, more integration between systems, open standards (where it is possible), and more safety and security. When we described ERTMS, the lack of implementation of its level 3 was noteworthy. The reasons for this are many, but perhaps the most relevant is that it implies a translation of costs from infrastructure managers to railway operators (less wayside equipment needed, but some functions to be done by trains, like guaranteeing the train integrity). However, in the last
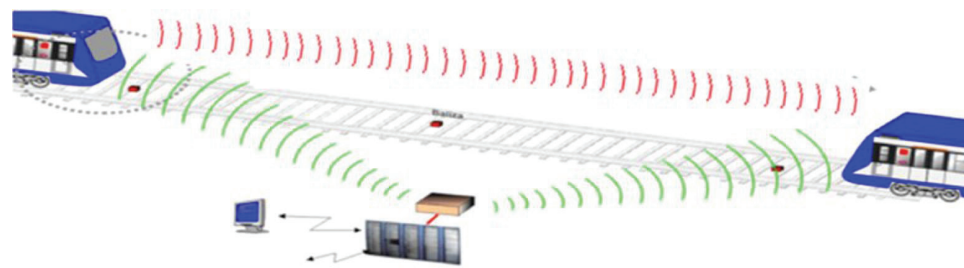
Figure 12. Schematic of the V2V-CBTC with both a direct link between trains and a train-to-infrastructure link.

years, some tests [21] have been carried out to check the feasibility of the level 3 of ERTMS, but not following a satisfactory approach (some parts were missing in these proofs of concept). Moreover, this is one of the main projects within Shift2Rail IP2 [22].

Another trend for the future of ERTMS is its "low-cost versions." Under this denomination we can find some initiatives that are focused on the implementation of a standardized and interoperable ERTMS but with some limitations to lower its CAPEX and OPEX. The most remarkable initiative in this direction is the "regional ERTMS" [23] where the line is divided into "dark zones"; only one train is allowed into each one of them. More relevant, there is no need of track circuits, and GSM-R is needed only in a punctual way.

When we explained CBTC, we saw that it can work in two modes: virtual track circuit and "moving block." The latter is very similar to ERTMS-L3; so in metro lines, this paradigm has been achieved some years ago. The next step in CBTC systems is to have a direct link between trains, in order to decrease the communication delay and the headway between them as well. Some signaling companies are working in this idea, but, as far as we know, only Alstom has shown some interest on it. In **Figure 4**, the basic concept under this "vehicular-to-vehicular CBTC" is shown. It would not remove the existing link between trains and wayside, but to be complemented by a direct link (V2 V) between trains when they are closer enough (**Figure 12**).

Finally, the sublimation of the signaling models is the so-called virtual coupling [23], where a large set of trains not coupled together could act as if they were. This implies sharing the same braking and power information for the entire "virtual train" that could be a large number of trains. To achieve this futuristic model, we need an ultrareliable communication link, with a very low latency. This line of research was raised 20 years ago but now is being researched in the EU-funded Shift2Rail IP2 initiative [22], with the aim of having a demonstrator with real trains in 2022.

Author details

Cesar Briso-Rodríguez[1]*, Juan Moreno García-Loygorri[1] and Lei Zhang[2]

*Address all correspondence to: cesar.briso@upm.es

1 Universidad Politecnica de Madrid, Spain

2 Key Lab of Wireless Sensor Network and Communication, Shanghai Research Center for Wireless Communications, China

References

[1] http://www.railway-technology.com/projects/european-rail-traffic-managementsystem-ertms

[2] EN 62290-1:2014 Railway applications. Urban Guided Transport Management and Command/Control Systems. System Principles and Fundamental Concepts

[3] González FJ. Señalización ferroviaria: del guardagujas a la operación sin conductor; Fundación Dialnet. 2008

[4] WIT press. Advanced Train Control Systems. 1st ed. WIT press; 2010

[5] European Railway Agency. ERTMS GSM-R QoS Test Specification [Internet]. July 2016. Available from: http://www.era.europa.eu/Document-Register/Pages/O_2475.aspx

[6] Tetsuya K, Atsushi K, Pham TD, Naokatsu Y. Seamless access networks using radio-over-Fiber. In: IEEE, editor. 2016 IEEE Sixth International Conference on Communications and Electronics (ICCE); India. 2016

[7] Motley AJ, Davis QV. Reliable high-speed data transmission from a mobile vehicle via a leaky feeder. Radio and electronic. Engineer. 1980;**50**(7):337-344

[8] Ruisi H, et al. High-speed railway communications: From GSM-R to LTE-R. IEEE Vehicular Technology Magazine. 2016;**11**(3):49-58

[9] NOKIA Networks. Railways. [Internet]; 2017. Available from: https://networks.nokia.com/railways

[10] Jaime C, et al. Long term evolution in high speed railway environments: Feasibility and challenges. Bell Labs Technical Journal. 2013;**18**(2):237-253

[11] Aleksander S, Jose. LTE for railways: Impact on performance of ETCS railway 999. IEEE Vehicular Technology Magazine. 2014;**9**(2):69-77

[12] Terrestrial Trunked Radio (TETRA). Voice plus Data (V+D); Designers' Guide; Part 1: Overview, Technical Description and Radio Aspects http://www.etsi.org/deliver/etsi_etr/300_399/30001/01_60/etr_30001e01p.pdf

[13] IEEE 802.11 standards. http://standards.ieee.org/about/get/802/802.11.html

[14] LTE/SAE -The Future Railway Mobile Radio System: Long-Term Visions on Railway Mobile Radio Technologies, UIC (Unión International Communications), V 0.4 Draft 14.09; 2009.

[15] [15GPP] Specifications [Online]: http://www.3gpp.org/dynareport/36-series.htm

[16] Dahlman E, Parkvall S, Skold J. 4G: LTE/LTE-Advanced for Mobile Broadband. 2nd ed; Elsevier. 2013

[17] Wang CX, et al. Channel measurements and models for high-speed train communication systems: A survey. IEEE Communications Surveys & Tutorials. 2016;**8**(2):974-987

[18] Ai B, et al. Challenges toward wireless Communications for High-Speed Railway. IEEE Trans. Intelli. Transp. Sys. Oct 2014;**15**(5):2143-2158

[19] Hrovat A, et al. A survey of radio propagation Modeling for tunnels. IEEE Communications Surveys & Tutorials. 2014;**16**(2):658-669

[20] Chen B, et al. Channel characteristics in high-speed railway: A survey of channel propagation properties. IEEE Vehicular Technology Magazine. 2015;**10**(2):67-78

[21] Furness N, van Houten H, Arenas L, Bartholomeus M. ERTMS level 3: The game-changer. IRSE News. April 2017;**232**:2-9

[22] Innovation Programme 2 – Shift2Rail. https://shift2rail.org/research-development/ip2

[23] ERTMS Regional – Status and Future Enhancements, UIC ERTMS World Congress Istanbul, Turkey 1st–3rd. April 2014. http://www.ertmsconference2014.com/assets/SESSION-PRESENTATIONS/S2/ERTMS-Regional-Status-andfuture-enhancements-2014-01-27-ed01.pdf.