◆ Mission Critical Communication Networks for Railways

Michael Liem and Veena B. Mendiratta

Communication networks in the railway sector are critical to the operation of the system and have stringent requirements for reliability and safety. These types of networks are commonly characterized as "mission critical." Further, rail communication networks have requirements for interoperability with legacy technology and long life cycle support. Many of the European railways operate trackside Global System for Mobile Communications-Railway (GSM-R) wireless networks; GSM-R is based on the GSM standard with railway-specific features. The railways have started to look at Long Term Evolution (LTE) as a potential future replacement system for GSM-R. This paper presents the role of communication networks in railway operations, the resulting unique requirements for mission critical rail networks, and current trends in railway telecommunications. A brief tutorial on GSM-R is provided. We then present the LTE network architecture and assess the suitability of LTE to meet the requirements of the railway sector, with a special focus on reliability. The paper focuses primarily on mainline rail networks; however, much of what is presented also is applicable to urban rail networks. © 2011 Alcatel-Lucent.

Role of Telecom in Railways: Moving Trains Safely and Efficiently

The railway transport system in many regions of the world, and in emerging countries in particular, is still the main transport system. But even in developed countries, in the overall context of worldwide efforts to reduce our carbon footprint, and coupled with the technological evolution of high-speed trains, railways are becoming a serious competitor to the airlines [8].

Telecommunications has always been a key enabling technology for railway transport. This is due to the nature of railway transport and the resulting operational constraints. An important characteristic of railway transport, with implications for the reliability and safety of the communication system, is the fact that due to the long braking distance, train drivers cannot normally drive by sight. Train movement in a railway network needs to be carefully controlled by ground personnel using a signaling system to avoid collisions and enable the efficient usage of the railway infrastructure. Operating the railway system involves:

- Setting of railway switches,
- Setting of trackside signals,
- Communicating movement orders directly to train drivers,

Bell Labs Technical Journal 16(3), 29–46 (2011) © 2011 Alcatel-Lucent. • DOI: 10.1002/bltj.20520

Panel 1. Abbreviations, Acronyms, and Terms

2G—Second generation
3G—Third generation
3GPP—3rd Generation Partnership Project
ASCI—Advanced Speech Call Items
ATC—Automatic Train Control
ATP—Automatic train protection
CCTV—Closed circuit television
CENELEC—European Committee for Electrotechnical Standardization
COTS—Commercial off-the-shelf
EIRENE—European Integrated Railway Radio Enhanced Network
EMC—Electromagnetic compatibility
EPC—Evolved packet core
ERTMS—European Rail Traffic Management System
ETCS—European Train Control System
ETCS L2—ETCS Level 2
ETSI—European Telecommunications Standards Institute
E-UTRAN—Evolved UTRAN
FRR—Fast Reroute
GSM—Global System for Mobile Communications
GSM-R—GSM-Railway
HSS—Home subscriber server
IMS—IP Multimedia Subsystem
IN—Intelligent network
IP—Internet Protocol
IT—Information technology
KPI—Key performance indicator
LTE—Long Term Evolution
MA—Movement authority
MME—Mobility management entity
MPLS—Multiprotocol label switching
MRF—Multimedia resource function
MSC—Mobile switching center
NE—Network element
PA—Public announcement
PABX—Private automatic branch exchange
PCRF—Policy and charging and rules function
PGW—Packet data network gateway
PSTN—Public switched telephone network
P-to-P—Peer-to-peer
QoS—Quality of service
RBC—Radio block center
RF—Radio frequency
S&C—Signaling and control
SCP—Signaling control point
SDH—Synchronous Digital Hierarchy
SGW—Serving gateway
SIL—Safety Integrity Level
SIP—Session Initiation Protocol
SONET—Synchronous Optical Network
TAS—Telephony application server
UE—User equipment
UMTS—Universal Mobile Telecommunications System
UTRAN—UMTS Terrestrial Radio Access Network
VoIP—Voice over Internet Protocol
WiMAX—Worldwide Interoperability for Microwave Access
WLAN—Wireless local area network

- Determining that railway tracks are clear, and
- Taking coordinated action in the case of unplanned events, to limit impact on passengers.

We distinguish functionally between the signaling level (which ensures the safe movement of trains) and the operation control level (which optimizes train movement in the overall system context) [19]. Railway signaling can be represented as a control loop, as shown in **Figure 1**.

In this control loop model, train movements are monitored by clear track-detection elements such as axle counters or track circuits. Train positions are logically processed with the position of railway switches and other control information to generate control actions. These control actions consist of setting switch positions and providing movement commands to the train driver via a series of visual signals, through an Automatic Train Control (ATC) system, or through direct oral orders to the driver. An example of a practical implementation of this control loop in a modern railway is shown in **Figure 2** [18].

The interlocking centers contain the vital safety logic of the signaling control loop. They are connected to the field elements either directly or through remote controllers. The interlocking centers are interconnected and connected to the operation centers which

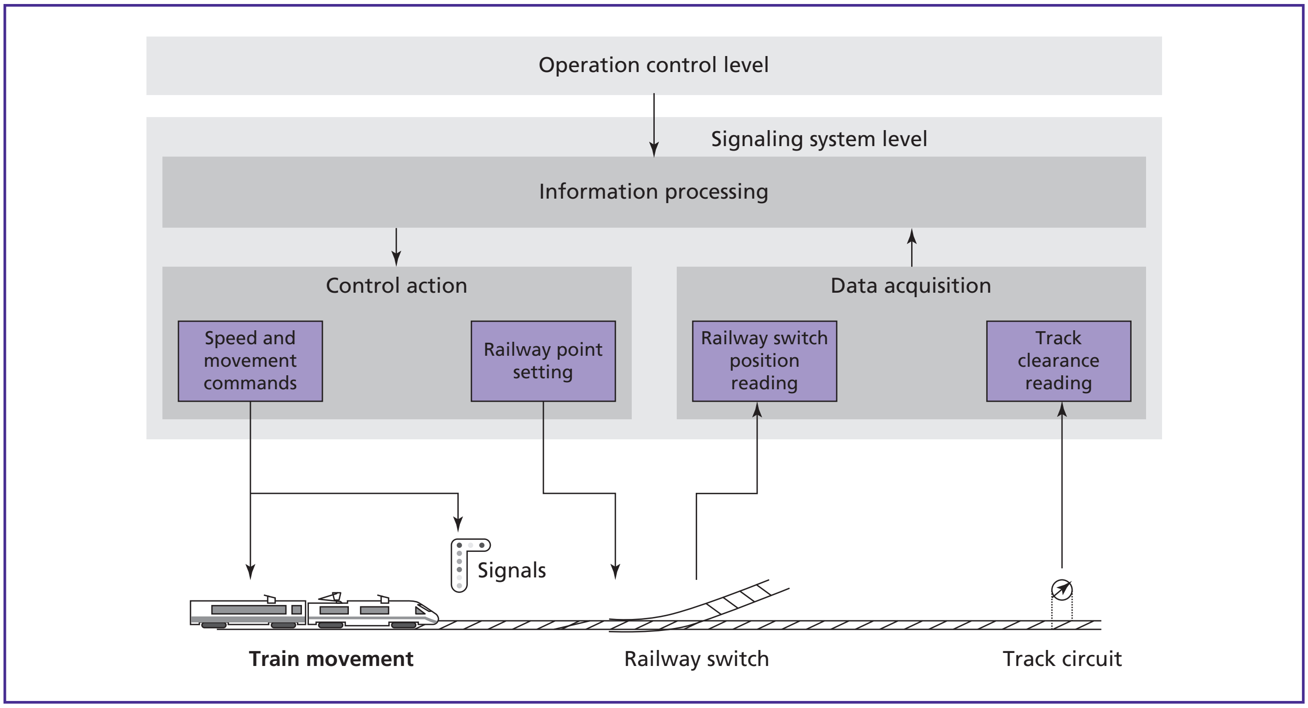


Figure 1.
Railway signaling control loop: controlling safe train movements.

centralize coordinating control functions. The signaling and operation control systems are geographically separated entities and require telecommunication systems to overcome the distances. As such, the telecommunication connections are an important part of the control loop of the railway transport system without which train operations would not be possible. This is a very good example of a mission critical network which is operated solely to support mission critical and often safety-relevant applications.

Telecommunications Systems in Railways

The telecommunications systems used within the railways can basically be divided into three groups. The first are the systems which take part in the train operation and thus are part of the operational train control loop, and components of the signaling and control (S&C) system for the rail network. The second type of system is not linked to train operations, but is used to support the corporate business processes. The third type of system is the telecommunications system that provides services to travelers, e.g., Internet access on trains. The latter are described in the section on Advanced Passenger Services. Standard corporate telecommunications systems are not discussed in this paper, as they are essentially the same as those for any major national company or travel business. The following sections focus specifically on the operational or mission critical telecommunication systems for railways. Note that since railway operation principles can vary widely between different countries, correspondingly, there is considerable variation in the underlying telecommunication systems as well. It is therefore difficult to offer a general description of train operations and derive from there a general requirement for telecommunications systems. In order to provide the reader with a high-level understanding, we nevertheless attempt to do so in the following sections with the caveat that the description may be incorrect for certain countries.

Voice Telephony

Voice telephony is still one of the main means of communication that railway staff around the world rely on to manage the movement of trains. On many

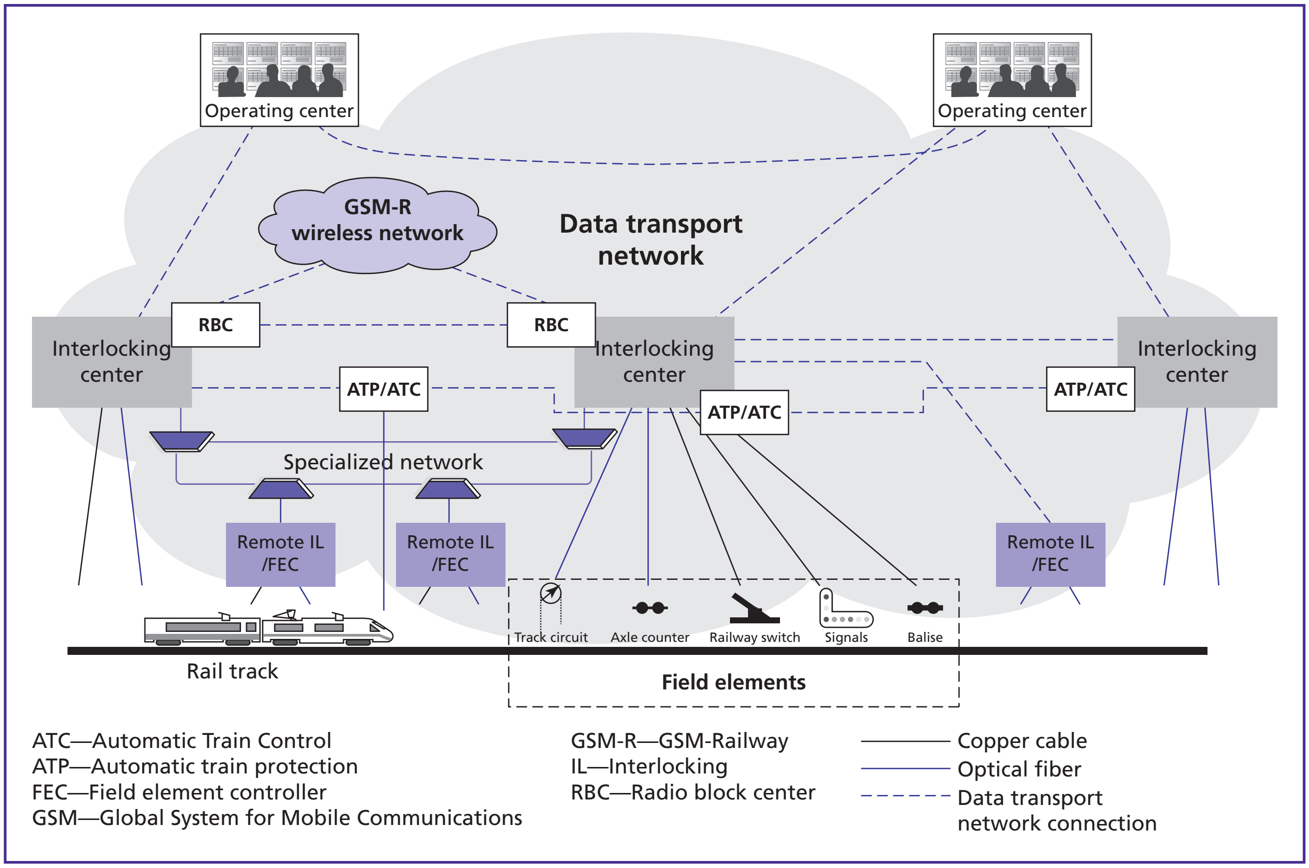


Figure 2.
Electronic interlocking systems with functional levels.

railways, signalers and dispatchers have key operational roles [19]. The signaler is the person responsible for safe train movement on a part of a track section. He controls the track signals and interlocking systems and gives the trains the authorization to enter a particular track section, once he has verified that it is free. The dispatcher has more of an overall coordination role. He has to be cognizant of unplanned changes to train schedules and determine the impact on train movements in near real time. Signalers and dispatchers communicate frequently. Signalers, who are often located in stations, have to hand over responsibilities for a train from one section of track to the next. Dispatchers have to give instructions to the signalers. Beyond these highly critical applications, communications with other ground staff are part of daily operations. Many railways have phone boxes along the railway track, in particular near signals. They serve as a backup means of communication between the train driver and the signaler. Typically, railways operate special purpose private automatic branch exchanges (PABXs). These PABXs provide key features such as support for group calls and call pre-emption in emergency situations. They also support a large variety of interfaces such as analog phones, party-line phones, public address (PA) systems, and train radio systems [12]. These railway PABXs are also called dispatcher systems.

Train Radio Systems

Wireless networks are now standard on most railways. They are used to communicate with the train driver directly. Their application can differ largely depending on the region. In North America, train movement authorization is communicated to the train crew via the train radio system, which sees heavy use during normal operations [17]. In Europe however, train radio is mostly used in atypical situations since train movement authorization is communicated through line-side signals. On a global basis, analog systems are still the most widely used, but digital systems are slowly beginning to replace them. An example system is Global System for Mobile Communications-Railway (GSM-R), which will be described in detail later.

Data Transport Networks

Modern electronic interlocking systems comprise different functional levels. Field elements (track circuits, axle counters, and railway points) and their control elements are connected to the central interlocking system. The central interlocking system components are connected to each other and to the systems from which they are controlled and operated. In order to provide these interconnections, railways often have their own transport network infrastructure. The backbones of these networks typically are based on optical fiber deployed along the tracks. Synchronous Digital Hierarchy/Synchronous Optical Network (SDH/SONET) is a well-established transport network technology in modern railways and is very attractive due to its protection capabilities. More recently, Internet Protocol (IP) network technology is being introduced as the core transport network technology.

Specific Requirements for Telecommunications in Railways

Mission critical railway telecommunications systems have specific requirements which often differ from telecommunications systems for public operators. These requirements are due to the nature of the railway domain and are linked to safety, reliability, lifecycle support, electromagnetic compatibility, and information technology (IT) security.

Safety

In the railway domain, there is a marked difference between the technical systems that are critical and vital for the safety of the train movements, and the systems that are not. Vital systems are differentiated by the levels of criticality and are assigned a Safety Integrity Level (SIL) [7]. SIL 4 is the level for the most critical systems, while SIL 0 designates systems without any vital functions. Since it is impossible in complex systems to achieve full test coverage during the system development phase, the approach in the railway domain is to make certain elements in the development process mandatory. These elements are linked to formal specifications and documentation, and thus the development costs for safety-critical systems may be significantly higher than those for non safety-critical systems. In order to leverage telecommunications systems from the public domain for railway use, the safety aspects need to be addressed on the functional layers above the standard telecommunications layer. For a data network this means that data integrity, authenticity, and access protection must be implemented in a data protocol running on top of the standard telecommunications layers [5, 6].

In the event of a failure in the telecommunications network, the vital railway system using it must automatically transition to a "safe" state. A very concrete example of this is the behavior of the European Railway Traffic Management System (ERTMS) in the case of a telecommunications network failure. This is described in the section titled GSM-R in Railway Operations.

Reliability

As described in the previous section, the telecommunications systems used by vital railway systems such as interlocking are not safety-critical (in the sense of having an SIL of 1 to 4). The predominant requirement towards telecommunications systems in this context is however their reliability behavior. Even though a telecommunication system failure does not lead directly to an unsafe mode, operating in a degraded mode (railway terminology) increases the risks and can indirectly degrade the safety of the system [4]. For that reason, railways put a high priority on the reliability and availability of the telecommunications system. Absolute availability targets are normally derived from an overall system safety case or from high-level operational targets. In the latter case, objectives in terms of minutes of delay per year, for example, are translated into a maximum acceptable level of system unavailability, of which a portion is allocated to the telecommunications network. The resulting end-to-end network availability targets set by the railways often exceed 99.99 percent and are in most cases higher than the requirements for public networks. The "Five Nines" objective often referenced in the telecommunications world refers to the node availability, not the end-to-end availability level. End-to-end availability of a network composed of several Five Nines elements is typically in the range of 99.9 percent to 99.99 percent. For example, the public switched telephone network (PSTN) has an end-to-end availability of 99.93 percent. Five Nines availability for a telecommunications node for the most part can only be achieved after several product releases and successive debugging in the field. The railways therefore tend to rely on field-proven technology. Deploying networks in the railways—in particular, if relatively new technology is used—requires a realistic understanding of telecommunication products and the resulting network reliability. In the absence of this realistic understanding, in many cases railways will use a very conservative approach with significant redundancy to be "on the safe side."

Lifecycle Support

Railway infrastructure has a long life cycle. Locomotives may be in use for at least 30 years or more. This global expectation of long life cycles is in sharp contrast to the pace of development within the telecommunications industry where systems are amortized, and frequently replaced, within 5 to 10 years. As telecommunications is only a supporting infrastructure for the railways—one which does not directly generate revenues—short life cycles are usually unacceptable.

Electromagnetic Compatibility

Electromagnetic compatibility (EMC) is of considerable concern in the railway environment [16]. Telecommunications equipment which is deployed close to the railway tracks is exposed to strong electromagnetic interference generated by high frequency power converters in modern trains and by the high-voltage current in the overhead power lines. Telecommunications equipment needs to comply with railway EMC standards.

Vibration and Temperature

In the railway environment telecommunications equipment is often mounted in very close proximity to the railway lines, and as a result is exposed to very high levels of vibration from passing rail vehicles. The temperatures in railway environments are clearly subject to the extremes of wherever the railway is based—from the deserts of the southwestern United States to tropical conditions in Southeast Asia, to Arctic weather in Scandinavia, Russia, and Canada. Railway operations as well as safety rely on telecommunications systems—and as a result they must be designed with these extremes in mind.

IT Security

Telecommunications networks transporting safety critical information are vulnerable to malicious attacks. This concern applies to all telecommunications and IT systems, but in the case of railways such attacks can have disastrous consequences. IT security is therefore of increasing concern. In a very conservative approach, this can lead to a total separation of networks for different applications to reduce the risks. Other approaches provide a separation between networks for vital versus non-vital applications through firewalls and access right functionalities [20]. The IT security aspect is becoming even more relevant with the trend towards movement to a consolidated IP network infrastructure as described in the section below.

Trends in the Railway Domain Impacting Telecommunications Systems

Even though the pace of development in the railway domain is much slower than in telecommunications carrier and enterprise markets, changes from both within the railway domain and outside of it impact the use of telecommunication systems. Recent pushes for reductions in government subsidies have pushed many railways to look for overall efficiencies through better use of telecoms/IT systems. Opportunities to generate revenues by creating a railway "telco" and serving both internal and external customers with high reliability trunk services is being exploited in some cases.

Increasing Competition in the Railway Industry

Since World War II, railway transport in many parts of the world has been a state-owned monopoly though there are notable exceptions, such as the United States. With the evolution of road and air transportation, most countries felt a need to increase railway performance by introducing additional competition. This is done either by privatization of the railways or by unbundling, i.e., separating the management of railway infrastructure from the train operating companies [8]. This, in turn, has put pressure on railway managers to increase efficiency and reduce costs. This pressure has both a direct and indirect impact on the deployed telecommunication systems.

Centralization

Many railways that were operating in a decentralized mode are moving to more centralization in order to reduce costs. For example local signalers in the stations may be relocated to central control centers, reducing the number of required staff. This more centralized control requires better data networks to control interlocking stations and field elements remotely. With this trend, telecommunications systems are becoming even more critical for train operations.

COTS

Traditionally many technical systems were custom-made for railways in general and even for specific national railways. Specific safety requirements, which are often written after disastrous accidents or events, were the reason for these specific developments. Increased cost pressure from investor-owned railways is now bringing the high cost of railway-specific system developments into question. Dramatic cost reductions achieved in the public telecommunication industries through a much larger market made it more and more attractive to use commercial off-the-shelf (COTS) telecommunication systems for the railways. The introduction of a regulatory framework such as the European Committee for Electrotechnical Standardization (CENELEC) EN50159-2 standard [6] in Europe laid the groundwork for using COTS telecommunications systems as operational railway systems (see the section above on Safety).

Internet Protocol

The dominance of IP in the telecommunications world, and the cost reduction potential expected from it, also has helped IP to penetrate the operational telecommunications systems of the railways. Ten years ago, IP was considered an unreliable technology for best effort Internet services and thus unsuitable for safety critical railway applications. This perception began to change with the introduction of IP technologies such as Multiprotocol Label Switching (MPLS) Fast Reroute (FRR), which offer similar protection performance as SDH/SONET, a globally accepted transport technology among the railways. In addition, IP networks offer the potential to merge networks for different applications onto a single common infrastructure, leading to considerable cost reduction as opposed to operating separate networks. However, at the same time it makes the networks more vulnerable to malicious attacks.

Advanced Passenger Services

The increased competition among railways, along with competition from other modes of transport, is forcing the railways to offer better services to their customers. As high-speed trains start to seriously compete with air transport for shorter distances, Internet access is a key service which the airlines simply cannot offer. Train travelers today have come to expect Internet access, and have been known to become annoyed if Internet service is not available onboard. These services are not necessarily offered directly by railways but often through collaboration with public operators. Current discussions regarding broadband-to-the-train are linked more closely to end customer services than to operational railway needs. Operational railway applications requiring broadband are only beginning to emerge, and currently focus on closed circuit television (CCTV)-related applications.

End-of-Life for Signaling Systems and Telecom Systems

The first electronic interlocking systems which were deployed in the 1980s are now approaching end-of-life and need to be replaced. Many railways are using this opportunity to also upgrade the underlying telecommunications systems, which have much shorter lifecycles. Obsolescence of technologies presents a huge challenge for railway infrastructure managers and railway operators. It is not unusual to see requests for spare parts to be available for 20 or more years in new equipment tenders. This is an effort on the part of the railroads to safeguard their investment choice. Suppliers of technologies to railway companies can gain a competitive advantage by being able to demonstrate a clear roadmap for systems under consideration that protects the railways interests.

Example of a Railway Communication System: GSM-R

In this section we present an overview of GSM-R, a digital train radio communication system based on the GSM standard [14].

Background of GSM-R

GSM-R was specified as the European Integrated Railway Radio Enhanced Network (EIRENE) standard [10, 11] in the 1990s. It was envisioned as a digital wireless train communication system which could replace the legacy analog systems in place, and also to serve as the bearer system for the European Train Control System Level 2 (ETCS L2). The objective of the European Rail Traffic Management System (ERTMS = GSM-R + ETCS L2) was to enable a common train signaling system that would allow trains to cross countries in Europe without an engine change. As planned, it would cover approximately 70 percent of all rail tracks in Europe with GSM-R. Over one-third of those deployments are now in operation. In addition, GSM-R has since been rolled out in countries outside of Europe such as China, Australia, and India and in the Middle East.

GSM-R in Railway Operations

GSM-R is used for voice and data applications, as shown in **Figure 3**. GSM-R voice service is mainly used for the communication between the train driver and the dispatcher or signaler. On the European railways, train radio communication is used mostly in atypical situations, since train movement is controlled via line-side signals. Therefore the voice traffic generated in a GSM-R network in Erlangs per subscriber is relatively low compared to that of public wireless networks. GSM-R voice services are also used for shunting operations.

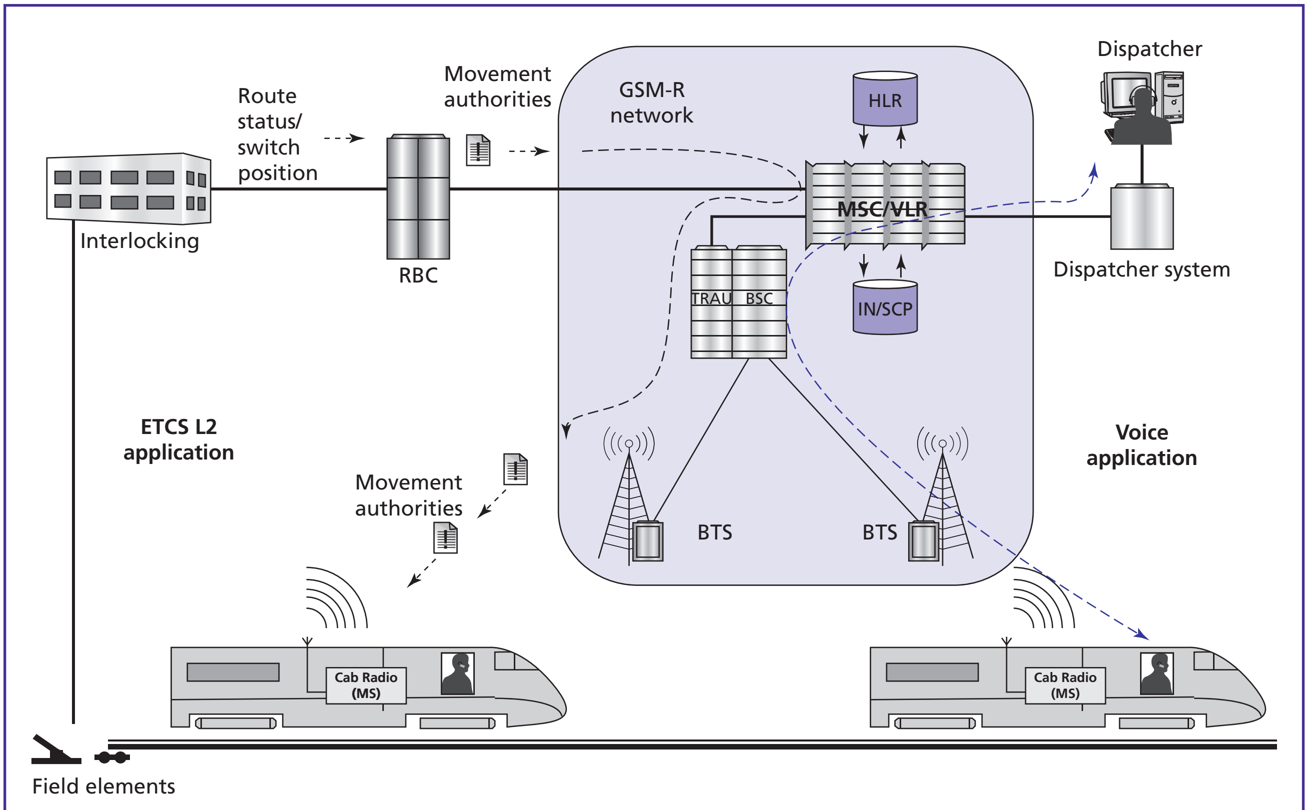
GSM-R circuit switched data service is used as a bearer for ETCS L2. The radio block center (RBC), which is connected to the interlocking and to the GSM-R network (mobile switching center (MSC)), plays a central role in communications. The interlocking provides route and track occupancy information which the RBC processes to produce movement authority (MA) messages. These messages are sent through the GSM-R network to the cab radio in the train. Using the MA messages, the onboard ETCS system calculates a dynamic speed profile for the train using positioning information as well as additional static data. In the event the GSM-R network is out of service, the train will either switch to a fallback signaling system or initiate an emergency break to get the train into a safe state (fail-safe principle).

Differences Between GSM-R and Public GSM Networks

GSM-R is based on the 3rd Generation Partnership Project (3GPP) European Telecommunications Standards Institute (ETSI) GSM standard and operates in the 900 MHz band with 19 dedicated GSM frequencies. The ETSI specification was enhanced with Advanced Speech Call Items (ASCI) features which are required to implement the following railway functionalities:

- Priority and preemption,
- Voice broadcast and voice group calls,
- Functional addressing,
- Location-dependent addressing,
- Fast call setup, and
- Railway emergency calls.

In most system implementations an intelligent network (IN) architecture is used in addition to the standard GSM architecture, with a signaling control point (SCP) controlling the call routing logic linked to the railway operations [15]. Beyond the differences in the system architecture, a major difference lies in the way the network needs to be engineered. The connection reliability for an operational train communication network needs to be significantly higher than the reliability of a public GSM network. When GSM-R is used as a bearer for ERTMS train control data, a communication link failure will lead to the emergency breaking of a train. As the most vulnerable component in the network, the radio link is therefore the biggest challenge when engineering a GSM-R network, and for ERTMS it needs a call drop rate 30 times better than that of public GSM network [9]. To increase system reliability, redundant elements are often used and double radio coverage layers are even deployed in some



BSC—Base station controller
BTS—Base transceiver station
ETCS—European Train Control System
ETCS L2—ETCS Layer 2
GSM—Global System for Mobile Communications
GSM-R—GSM-Railway
IN—Intelligent network

MS—Mobile station
MSC—Mobile switching center
RBC—Radio block center
SCP—Signaling control point
SDH—Synchronous Digital Hierarchy
TRAU—Transcoder and rate adaptation unit
VLR—Visitor location register

Figure 3.
Overview of GSM-R for voice and data applications.

networks to further increase the overall system availability for ERTMS applications. The radio engineering in GSM-R is particularly delicate. The high call drop rate performance requires engineering for a much higher radio frequency (RF) level and this in a difficult RF environment which includes metal bridges, tunnels, and cuttings. RF holes in linear track coverage are much more difficult to compensate through tilting of antennas than when covering surfaces in GSM. Location-dependent call routing to the right dispatcher or signaler based on the cell identity requires matching cell boundaries with operational dispatcher zones, an engineering challenge unknown in public GSM.

Lessons Learned From GSM-R

In general, the European railways consider the deployment of GSM-R for operational purposes a success. The system has been successfully deployed and operationally proven in more than 30 countries. A number of these networks are deployed on smaller lines, but some GSM-R networks have nationwide track coverage. Users acknowledge the improved quality of service (QoS) over the previous analog systems. As a mature technology coming from the public telecommunications world, the system reliability observed in the field meets the high objectives of the railways. However, the cost benefit of using

COTS technology could only be partially realized. The additional product features required for railway operations required the development of special versions of public GSM products for GSM-R. This led to relatively low volumes of these products and a reduced number of GSM vendors focusing on GSM-R, thus keeping the pricing for these products at a higher level than for equivalent GSM products. In future wireless train communication systems, the railways will try to avoid railway-specific versions of public systems in order to achieve the full cost and innovation benefits of using "real" COTS telecommunication systems. For a future wireless train communication system, however, this would require a clear separation of the basic COTS telecommunications bearer functions from any functionality which is used only in a railway context; this has not been the case for GSM-R.

Long Term Evolution

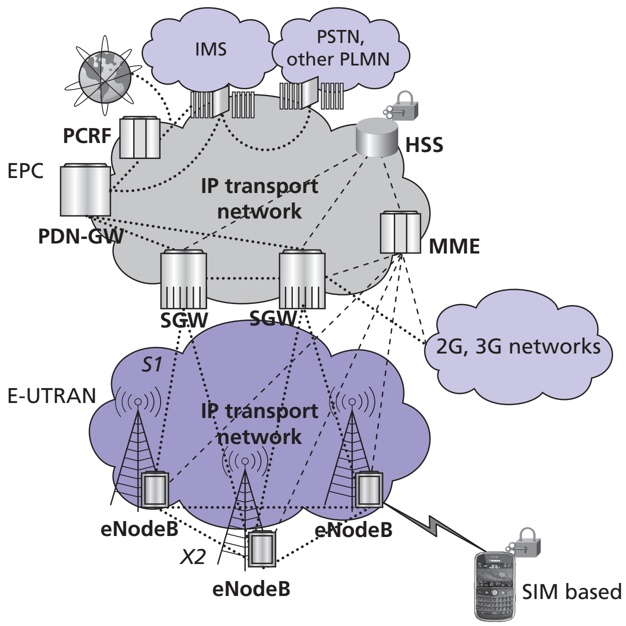
New GSM-R deployments are continuing in Europe while commercial mobile operators are evolving their GSM services to migrate to Long Term Evolution (LTE). LTE has been designed to be more efficient, to offer new services and still run on the same radio frequency bands as second generation (2G) and third generation (3G) systems. Therefore it is likely that at some point the railway industry will consider evolution to LTE to take advantage of the performance, throughput, reliability and cost benefits offered by this 4G wireless technology [13]. The key drivers for the railways are the costs and life cycle support aspects of GSM-R, of which the underlying GSM technology will itself start to become a legacy technology in the coming years. The introduction of new broadband services with LTE is also considered but is less important for the railways. For railways which have not yet moved to a digital wireless trackside system, the direct move from analog legacy systems to LTE will be a compelling alternative. The timeframe in which this will occur in practice depends on the telecom vendor's capability to demonstrate that the key requirements, listed below, for a wireless system-supporting railway operations are met [1].

- High availability for S&C functions.
- Support for low bandwidth operational applications (e.g., train control) with performance at least as good as GSM-R.
- Very low dropped call rate.
- Seamless handover and fast connection re-association time.
- Low sensitivity to high train speeds of up to 350 km/hour and beyond.

The standards allow LTE to be deployed in any 3GPP spectrum and in a variety of bandwidths (from 1.4 MHz to 20 MHz). That is, a 2G or 3G operator can allocate some of its existing spectrum to LTE. Wider bandwidth is mostly available in the higher frequency bands, allowing for higher throughput, and is well suited for areas where high data rates are required and propagation is not a concern, while lower frequencies have a better propagation profile and thus provide better coverage.

LTE Architecture

LTE consists of the evolved packet core (EPC) and Evolved UMTS Terrestrial Radio Access Network (E-UTRAN). The EPC is based on IP and is a multi-access core network that enables operators to deploy and operate a common packet core network for 3GPP radio access (LTE, 3G, and 2G), non-3GPP radio access (wireless local area network (WLAN), Worldwide Interoperability for Microwave Access (WiMAX)), and fixed access (Ethernet, Digital Subscriber Line, cable, and fiber). The E-UTRAN is connected only to the packet switched domain of the core network. The E-UTRAN protocols and user plane functions have therefore been optimized for the transmission of traffic from IP-based real time and non-real time applications and services. The E-UTRAN includes the eNodeB network elements. The EPC includes the following: serving gateway (SGW), packet data network gateway (PGW), mobility management entity (MME), policy charging and rules function (PCRF), home subscriber server (HSS), and IP Multimedia Subsystem (IMS) for voice and other applications. LTE is an all-IP packet-switched architecture where all services are delivered through packet connections including voice. Voice and other services such as video are implemented through the use of an IP Multimedia



E-UTRAN: evolved UMTS radio access network
- eNodeB

EPC: evolved packet core
- SGW: serving gateway
- PGW: packet data network gateway
- MME: mobility management entity
- PCRF: policy charging and rules function
- HSS: home subscriber server

Multi-homing

Interworking with legacy networks
- 3GPP, 3GPP2

All-IP architecture, compatible with legacy mobile networks

2G—Second generation
3G—Third generation
3GPP—3rd Generation Partnership Project
3GPP2—3rd Generation Partnership Project 2
eNodeB—Enhanced NodeB
EPC—Evolved packet core
E-UTRAN—Evolved UTRAN
GW—Gateway
HSS—Home subscriber server
IMS—IP Multimedia Subsystem
IP—Internet Protocol
LTE—Long Term Evolution
MME—Mobility management entity
PCRF—Policy charging and rules function
PDN—Packet data network
PLMN—Public land mobile network
PSTN—Public switched telephone network
SGW—Serving gateway
SIM—Subscriber identity module
TAS—Telephony application server
UMTS—Universal Mobile Telecommunications System
UTRAN—UMTS terrestrial radio access network

Figure 4.
LTE reference network architecture.

Subsystem. Railway-specific applications can be provided by a special-purpose IMS application server. **Figure 4** shows the LTE reference network architecture. It is outside the scope of this paper to provide details of the LTE network architecture; the reader is referred to references [2] and [3] respectively for a detailed treatment of the LTE core network and the LTE radio access network.

LTE incorporates reliability-enhancing features such as the following:

- Intra-eNodeB handovers between sectors.
- Inter-eNodeB handovers on the X2 interface with very low interruption time and no user session packet loss. The X2 interface improves handover performance and reduces loading on the MME.
- S1-Flex feature—an eNodeB is connected to a pool of MMEs so there is minimal impact on service in the case of an MME failure.
- Network elements such as the MME, SGW, and PGW can be deployed in pools or clusters in active load-shared mode and any element of the pool can be used to service a request from the eNodeB.

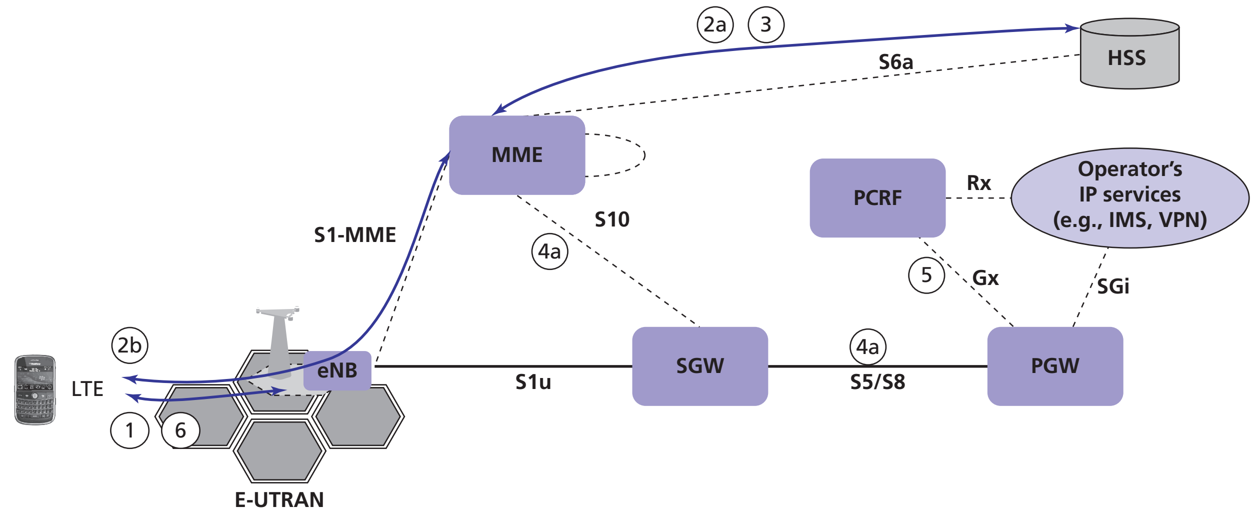
Reliability Analysis

One of the key aspects for railway telecommunication systems is the reliability of the bearer service as described in the section titled Specific Requirements for Telecommunications in Railways. In this section the system reliability of LTE is analyzed and then compared to that of GSM-R. We exclude in this analysis the reliability performance of the radio link which is difficult to assess analytically.

We list below some of the key performance indicators (KPIs) and reliability definitions that are important for assessing the availability and reliability of the LTE end-to-end solution.

Key performance indicators.

1. Probability of user equipment (UE) to successfully attach to the network, P (attach).
2. Probability of successful Session Initiation Protocol (SIP) registration, send INVITE, P (SIPreg).
3. Probability of successful service request for best effort data, P (ServData).
4. Probability of successful service request for Voice over Internet Protocol (VoIP), P (ServVoIP).
5. Probability of successful continuation of call or data session to completion/service request was successful and handovers were successful, P (completion).
6. Dropped call rate.



UE initiates attach (e.g., power on)

① Establish RRC connection

② MME initiates authentication and security mode with HSS and UE

③ Subscriber data query

④ MME selects SGW and PDN GW and establishes PDN connection

⑤ IP CAN session establishment with QoS policy query

⑥ Activate the default bearer

CAN—Connectivity access network
eNB—Enhanced NodeB
E-UTRAN—Evolved UTRAN
GW—Gateway
HSS—Home subscriber server
IMS—IP Multimedia Subsystem
IP—Internet Protocol
LTE—Long Term Evolution
MME—Mobility management entity
PCRF—Policy charging and rules function
PDN—Packet data network
PGW—Packet data network gateway
QoS—Quality of service
RRC—Radio resource control
SGW—Serving gateway
UE—User equipment
UMTS—Universal Mobile Telecommunications System
UTRAN—UMTS Terrestrial Radio Access Network
VPN—Virtual private network

Figure 5.
UE attach message flow.

Service availability in the LTE solution.

- *Voice/video service availability*: items 1, 2 and 4 are completed successfully.
- *Data service availability*: items 1 and 3 are completed successfully.

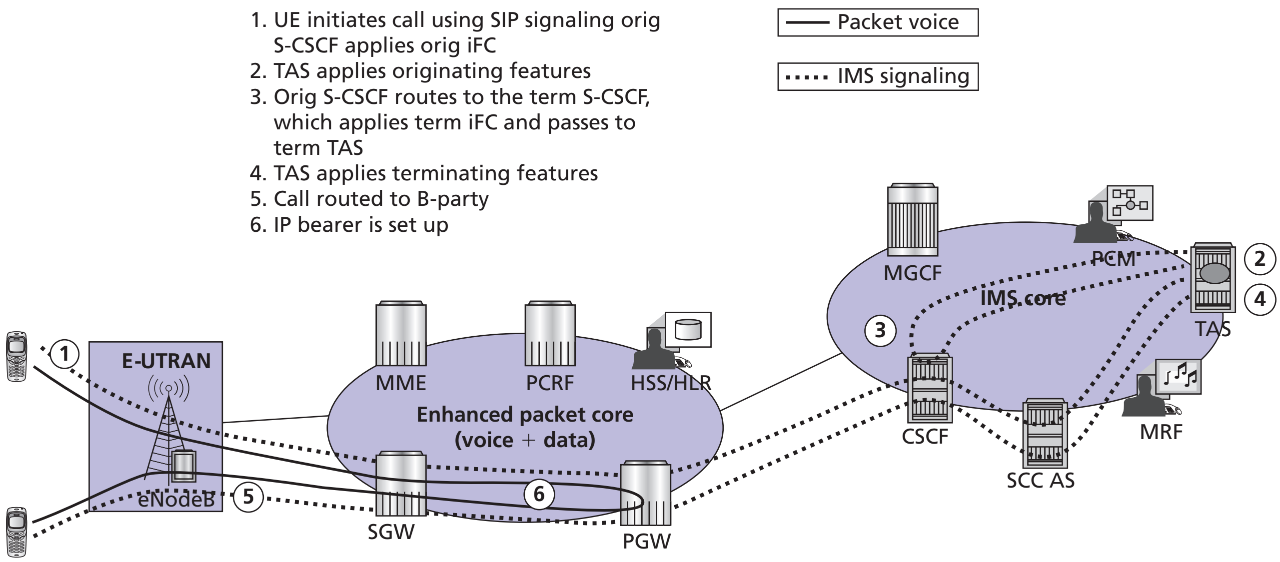
Service reliability in the LTE solution.

- *Voice service reliability (VoIP)*: item 5 is completed successfully.
- *Data service reliability*: item 5 is completed successfully.

The reliability analysis includes the reference connections for the following basic services:

- *Attach request.* UE registers with the LTE network; message flow is UE to local MME (eNodeB to MME for analysis purposes) including the local SGW, PGW, PCRF, and HSS.
- *Data service request* (best effort or with allocation of QoS bearers). Message flow is:
  — UE (eNodeB for analysis purposes) to the interface to the public data network/Internet including the local SGW, PGW, PCRF, and HSS.
  — UE-to-UE for peer-to-peer (P-to-P) data service including the LTE elements and networks at both ends of the call and the core IP network.
- *Voice service request* (VoIP on IMS). UE-to-UE (eNodeB to eNodeB for analysis purposes) including the LTE elements, IMS elements, and networks (local and metro) at both ends of the call, and the core IP network.

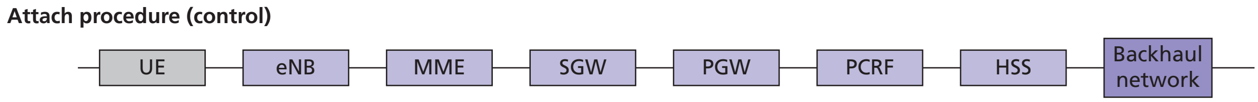
For illustration, **Figure 5** and **Figure 6** show the UE attach message flow path and the IMS call message flow path respectively.



1. UE initiates call using SIP signaling orig S-CSCF applies orig iFC
2. TAS applies originating features
3. Orig S-CSCF routes to the term S-CSCF, which applies term iFC and passes to term TAS
4. TAS applies terminating features
5. Call routed to B-party
6. IP bearer is set up

AS—Application server
eNodeB—Enhanced NodeB
E-UTRAN—Evolved UTRAN
HLR—Home location register
HSS—Home subscriber server
iFC—Initial filter criteria
IMS—IP Multimedia Subsystem
IP—Internet Protocol
MGCF—Media gateway control function
MME—Mobility management entity
MRF—Multimedia resource function
orig—Originating
PCM—Pulse code modulation
PCRF—Policy charging and rules function
PGW—Packet data network gateway
SCC—Session continuity control
S-CSCF—Serving call session control function
SGW—Serving gateway
SIP—Session Initiation Protocol
TAS—Telephony application server
term—Terminating
UE—User equipment
UMTS—Universal Mobile Telecommunications System
UTRAN—UMTS Terrestrial Radio Access Network
VoIP—Voice over Internet Protocol

Figure 6.
VoIP on IMS: call message flow.

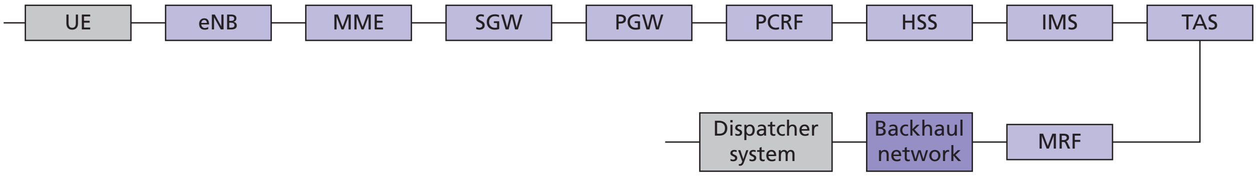


eNB—Enhanced NodeB
HSS—Home subscriber server
LTE—Long Term Evolution
MME—Mobility management entity
PCRF—Policy charging and rules function
PGW—Packet data network gateway
SGW—Serving gateway
UE—User equipment

Figure 7.
LTE attach procedure reliability block diagram.

For the reliability analysis for the attach, data service request, and voice service request procedures reliability block diagrams representing the respective reference connections—for the control and bearer paths—are developed to calculate the availability of each service. These are shown in **Figure 7** and **Figure 8** and respectively.

The availability of these services is dependent on the availability of the network elements in the path. Typically the telecommunications network elements have 99.999 percent or Five Nines availability. For the analysis, we assume that all elements have an availability of 99.999 percent, except for the eNodeB which has an availability of 99.995 percent. Note that the LTE network can be architected such that the network elements are in a cluster or pool configuration (with the exception of eNodeBs) and any element in the cluster can be selected based on the load-balancing scheme that is implemented. Therefore, in effect, the availability of a given type of network element, for example the MME, is greater than the individual node availability and is equal to the cluster availability. For this analysis we assume that the cluster size is 2 for all network elements, which provides cluster availability



(a) Voice service request (control): LTE IMS to dispatcher call

(b) Voice service request (bearer): LTE IMS to dispatcher call

eNB—Enhanced NodeB
HSS—Home subscriber server
IMS—IP Multimedia Subsystem
IP—Internet Protocol
LTE—Long Term Evolution
MME—Mobility management entity
MRF—Multimedia resource function
PCRF—Policy charging and rules function
PGW—Packet data network gateway
SGW—Serving gateway
TAS—Telephony application server
UE—User equipment
VoIP—Voice over Internet Protocol

Figure 8.
LTE VoIP on IMS call reliability block diagram.

Table I. LTE reliability model inputs.

Element	Availability
eNodeB	0.99995
MME	0.99999
SGW	0.99999
PGW	0.99999
PCRF	0.99999
HSS	0.99999
IMS	0.99999
TAS	0.99999
MRF	0.99999
Backhaul network	0.9999

eNodeB—Enhanced NodeB
HSS—Home subscriber server
IMS—IP Multimedia Subsystem
IP—Internet Protocol
LTE—Long Term Evolution
MME—Mobility management entity
MRF—Multimedia resource function
PCRF—Policy charging and rules function
PGW—Packet data network gateway
SGW—Serving gateway
TAS—Telephony application server

in the range of 99.9999 percent. The exact value depends on the coverage factor (see below). Further, to improve reliability, the network can be architected so that there is overlapping coverage between the eNodeBs—in effect providing eNodeB redundancy—where the degree of overlap determines the effective eNodeB availability. We also assume that there are two eNodeBs with overlapping coverage which is a typical wireless architecture for railway applications.

Cluster availability is a function of the fault recovery coverage of the cluster. Fault recovery coverage is defined as the conditional probability, given that an error has occurred in one of the network elements in the cluster, that the system recovers automatically from the failure within the designated recovery interval with minimal service impact. For a pool of LTE network elements (NEs), this means that the failed NE is removed from the pool and other NEs in the pool and network are aware of the working/non-working state of the NEs in the pool. Typically, mature telecommunications system NEs with high reliability have coverage factor values greater than 0.99; however coverage is never perfect, it is always less than 1.

The inputs for the reliability model are shown in **Table I,** and the expected availability results for the LTE attach and voice services are shown in **Table II** for cluster coverage values varying from 0.90 to 1.0. The UE, IP packet transport network in the core, radio interface, and dispatcher system are not included in the availability model; the backhaul network is included in the availability model. For comparison, the GSM-R network element and end-to-end voice connection expected availability is shown in **Table III**.

Table III. GSM-R network element and solution expected availability.

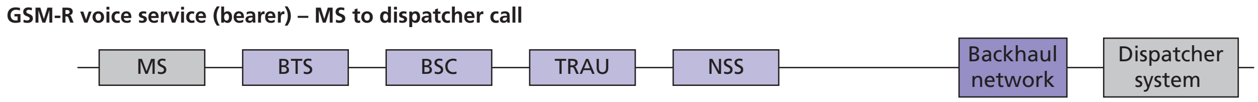
Element	Availability
NSS	0.99999
BSC	0.99999
TRAU	0.99999
BTS	0.99995
Backhaul	0.9999
End-to-end	**0.9998**

BSC—Base station controller
BTS—Base transceiver station
GSM—Global System for Mobile Communications
GSM-R—GSM-Railway
NSS—Network switching subsystem
TRAU—Transcoder and rate adaptation unit

Table II. Expected availability for LTE services.

Service	Coverage = 1	Coverage = 0.99	Coverage = 0.90
Attach—control	0.99990	0.999899	0.999890
Voice—control	0.99990	0.999898	0.999886
Voice—bearer	0.99990	0.999899	0.999892

LTE—Long Term Evolution



Note: The simplified assumption is made that all core network nodes (MSC, HLR, SCP) are within a single node.

BSC—Base station controller
BTS—Base transceiver station
GSM—Global System for Mobile Communications
GSM-R—GSM-Railway
HLR—Home location register
MS—Mobile station
MSC—Mobile switching center
NSS—Network switching subsystem
SCP—Signaling control point
TRAU—Transcoder and rate adaptation unit

Figure 9.
GSM-R voice call reliability block diagram.

Similar to the LTE analysis, the UE and the radio interface are not included in the availability model for GSM-R; the backhaul network is included in the availability model. **Figure 9** shows the GSM-R voice call reliability block diagram.

As seen from the results in Table II and Table III, the LTE solution provides higher service availability than the GSM-R solution based on the reliability enhancing features in LTE, in particular the pooling of NEs which effectively increases the availability of an NE type. LTE results are shown for coverage factor values of 0.90, 0.99, and 1.0. At initial deployment the coverage is likely to be lower than 0.99 but with a concerted effort at root cause analysis and fixing of faults found in the field in the early deployment phase, coverage factor values of 0.99 can be achieved rapidly.

Conclusions

In this paper, we provided an overview of the key role telecommunication systems play in railway safety and operations. These mission critical telecommunications systems must meet specific requirements with respect to reliability, IT security, EMC, and lifecycle support. Voice is still the dominant application in the railways today, but data networks are becoming more important with centralization and automation of railway operations. The pressure to reduce costs is driving the use of commercially available communication systems from the public operator domain. One successful example is GSM-R, which is based on the GSM standard and specifically adapted for the railway domain. GSM-R deployment has also illustrated the difficulty of achieving cost reductions through COTS while at the same time supporting railway-specific features and requirements, in particular long life cycle support. LTE will become the dominant wireless technology for public operators in the coming years and could be a successor to GSM-R. We analyzed the reliability of LTE voice service and compared it to GSM-R, showing that LTE will meet the current reliability performance of GSM-R, thus confirming LTE's position as a potential successor to GSM-R. Further work has to focus on the field radio link performance of LTE in high-speed conditions. The implementation of railway-specific features in conjunction with LTE requires further analysis to avoid a railway-specific version of LTE and thus to achieve the maximum cost benefit possible.

References

[1] Alcatel-Lucent, "Changing Track: Ground-to-Train Broadband Communications and New Train-Borne Applications Are Considering a Move to 4G/LTE," Railway Strategies, Oct.-Nov. 2009, <http://www.railwaystrategies.co.uk/article-page.php?contentid=8909&issueid=289>.
[2] K. Bogineni, R. Ludwig, P. Mogensen, V. Nandlall, V. Vucetic, B. K. Yi, and Z. Zvonar (eds.), LTE Part I: Core Network, IEEE Commun. Mag., 47:2 (2009), 40–98.
[3] K. Bogineni, R. Ludwig, P. Mogensen, V. Nandlall, V. Vucetic, B. K. Yi, and Z. Zvonar (eds.), LTE Part II: Radio Access, IEEE Commun. Mag., 47:4 (2009), 40–88.

[4] European Committee for Electrotechnical Standardization, "Railway Applications—The Specification and Demonstration of Reliability, Availability, Maintainability and Safety (RAMS)," CENELEC EN 50126, 2000.
[5] European Committee for Electrotechnical Standardization, "Railway Applications—Communication, Signalling and Processing Systems—Part 1: Safety-Related Communication in Closed Transmission Systems," CENELEC EN 50159-1, 2001.
[6] European Committee for Electrotechnical Standardization, "Railway Applications—Communication, Signalling and Processing Systems—Part 2: Safety-Related Communication in Open Transmission Systems," CENELEC EN 50159-2, 2001.
[7] European Committee for Electrotechnical Standardization, "Railway Applications—Communication, Signalling and Processing Systems—Safety Related Electronic Systems for Signalling," CENELEC EN 50129, 2003.
[8] J. A. Gomez-Ibanez and G. de Rus (eds.), Competition in the Railway Industry: An International Comparative Analysis, Edward Elgar, Cheltenham, UK, Northampton, MA, 2006.
[9] International Union of Railways, "ERTMS/GSM-R Quality of Service Test Specification," UIC O-2475 3.0, 2007.
[10] International Union of Railways, "UIC EIRENE System Requirements Specification," v15, 2010.
[11] International Union of Railways, "UIC Project EIRENE Functional Requirements Specification," v7, 2010.
[12] U. Lange and M. Rost, "Neue Betriebsfernmeldetechnik der Ungarischen Eisenbahnen" ("New Communication System for Hungarian Railways"), Signal und Draht, Dec. 2007, 22–25.
[13] D. Mandoc and K. D. Masur, LTE/SAE—The Future Railway Mobile Radio System? Long-Term Visions on Railway Mobile Radio Technologies, International Union of Railways (UIC), 2009.
[14] M. Mouly and M.-B. Pautet, The GSM System for Mobile Communications, Telecom Publishing, Palaiseau, Fra., 1992.
[15] A. F. Neele and S. T. G. Wootton, GSM-R Procurement Guide, International Union of Railways (UIC), 2007.
[16] S. Niska, Measurements and Analysis of Electromagnetic Interferences in the Swedish Railway System, Ph.D. Thesis, Luleå University of Technology, 2008.
[17] J. Pachl, "Kommunikation im Bahnbetrieb" ("Communication in Railway Operations"), Deine Bahn, July 2007, 35–40.
[18] H.-W. Renz, "Open Communication Standards for Safe and Secure Railway Control," Alcatel Telecommun. Rev., 2Q (2004), 208–214.
[19] J. Trinckauf, "Basic Characteristics of Railway Systems and the Requirements for Signalling," Railway Signalling and Interlocking: International Compendium (G. Theeg and S. Vlasenko, eds.), Eurailpress, Hamburg, 2009, Chapter 1.
[20] T. Wickinger, "Sicherer Datenaustausch zwischen den Integritätsbereichen von Betriebszentralen" ("Safe Data Exchange Between the Control Systems in the Operations Centres"), ETR: Eisenbahntechnische Rundschau, 50:6 (2001), 355–359.

(Manuscript approved June 2011)

MICHAEL LIEM is a senior solution architect in Alcatel-Lucent's Services Division in Velizy, France. After a military career in the German Air Force, Mr. Liem worked 12 years for Nortel in research and development (R&D) and Network Engineering for second and third generation wireless systems. He joined Alcatel-Lucent four years ago where his focus is today the design and deployment of mission critical networks for the railway domain. His technical interests include new systems engineering methods and network reliability. Mr. Liem holds a Diplom-Ingenieur degree in electrical engineering from Helmut Schmidt University in Hamburg, Germany, and an advanced masters degree in telecommunication networks from Telecom ParisTech, France.

VEENA B. MENDIRATTA leads the next-generation solutions, services and systems reliability work in the Bell Labs Network Performance and Reliability department at Alcatel-Lucent in Naperville, Illinois. She began her career at AT&T Bell Labs over 25 years ago. Her work has focused on the reliability and performance analysis for telecommunications systems products, networks, and services to guide system architecture solutions. Her technical interests include architecture, system and network dependability analysis, software reliability engineering, and telecom analytics. Current *work is focused on predictive analytics for the telecommunications domain, LTE solution reliability engineering, and service reliability modeling for railway solutions. She holds a B.Tech. in engineering from the Indian Institute of Technology, New Delhi, India, and a Ph.D. in operations research from Northwestern University, Evanston, Illinois, USA. Dr. Mendiratta is a senior member of the IEEE, a member of INFORMS, a member of the Alcatel-Lucent Technical Academy, and a recipient of the Bell Labs President's Award.* ◆