Millimeter-Wave Network Architectures for Future High-Speed Railway Communications: Challenges and Solutions

Hao Song, Xuming Fang, and Yuguang Fang

Hao Song and Xuming Fang are with Southwest Jiaotong University.

Yuguang Fang is with the University of Florida.

Digital Object Identifier: 10.1109/MWC.2016.1500255WC

Abstract

The shortage and congestion of lower spectra motivate the exploration of the broadband and underutilized mmWave to be used in future public mobile communications, and subsequently future HSR mobile communications, which has the potential to provide multi-gigabit rate radio access for train passengers. However, due to the inferior propagation characteristics of mmWave and particularity of HSR scenarios, there are many design challenges ahead. In this article, we tackle those challenges by developing technical solutions for mmWave broadband HSR systems. We first propose feasible multiple access techniques and frame structures based on OFDM and SC communications, respectively. We then present promising train–trackside network architectures based on different MIMO techniques, including BF and SM, which have been viewed as the key enabling technologies to realize mmWave communications in outdoor environments. Moreover, we discuss the inherent defects of each architecture and offer the corresponding solutions or recommendations. Finally, we conduct performance evaluation, and discuss the advantages and disadvantages of the proposed architectures. We hope this article will stimulate further research on the innovative use of mmWave in high-speed rail systems.

Introduction

According to a Cisco report published in 2015 [1], global mobile data traffic had grown by 69 percent, reaching 30 exabytes, and mobile equipment had increased by 500 million, reaching 7.4 billion, by 2014. Moreover, the global mobile data traffic will increase nearly tenfold with a compound annual growth rate (CAGR) of 57 percent between 2014 and 2019. To keep pace with this exponential growth only relying on the current common capacity enhancement approaches, spectrum efficiency and network density, is inadequate; therefore, a novel approach is needed. One approach, called spectrum extension, has attracted tremendous attention lately. By examining global spectrum allocation, we find that the spectra used by most radio communications systems, such as AM/FM, cellular, satellite, and 802.11 WLAN, concentrate in the range between 300 MHz and 5.8 GHz, which have become increasingly congested with mobile data traffic growth, while a vast amount of millimeter-wave (mmWave) spectra between 6 GHz and 300 GHz is still underutilized and underexploited. Unfortunately, mmWave experiences higher propagation loss in free space. Moreover, mmWave signals suffer from more serious penetration loss in most common materials, such as glasses, metals, brick walls, and concretes, than lower frequencies, which depends not only on frequencies, but also on the electrical properties of material being penetrated. Furthermore, mmWave transmissions have more stringent requirements on electronic components, size, and power consumption [2]. For these reasons, most research works on mmWave communications still focus on indoor or fixed point-to-point systems, such as 802.11ad, WirelessHD technology, ECMA-387, and 802.15.3c. The research on mmWave communications for outdoor mobile systems has received intensive attention only recently, and thus more thorough and deep investigations are needed. As a special case of outdoor mobile systems, there are few significant reports or technical specifications on mmWave for HSR communications. Considering the fact that densely populated users tend to gather in a single railroad carriage and travel for extended periods of time, it is not hard to imagine that they would be very interested in Internet services, particularly multimedia services. Unfortunately, due to the limitation of spectrum bandwidth, even the latest generation high-speed railway (HSR) communication system, LTE for Railway (LTE-R), cannot provide broadband radio access for each user in high-speed trains. Hence, in this article, we explore the possibility of utilizing ultra-wideband mmWave to offer high-speed, say, multi-gigabit data rate, Internet services.

Despite more severe free space propagation loss experienced by mmWave because of higher frequencies, the shorter wavelength of mmWave has the advantage that more antennas can be deployed in the same area. For this reason, researchers have reached the conclusion that multi-antenna is a feasible way to overcome the high free-space loss of mmWave, and is a key enabling technique for mmWave mobile systems. For example, Shu *et al.* studied spatial multiplexing

(SM) and beamforming (BF) for mmWave mobile systems, and showed that mmWave could achieve very high performance in outdoor scenarios [3]. Hence, in this article, we propose a train–trackside mmWave wireless system based on different multiple-input multiple-output (MIMO) technologies. In addition to higher free space loss, due to the particularity of mmWave spectrum and the HSR scenario, we have to address the challenges below.

High Penetration Loss: The electromagnetic waves must penetrate train shells if an mmWave base station (BS) directly communicates with users within a train. Due to the well-known fact that the penetration loss of mmWave is more serious than lower bands [2], the link of train–trackside transmissions should not rely on one-hop communication.

High Doppler Shift: According to the formula for Doppler frequency shift $f_{d_{\max}} = (\upsilon/\lambda)\cos\theta$, the short wavelength of mmWave and high speed of a train may cause more serious Doppler frequency shift.

Handover Problems (Group Handover and Highly Frequent Handover): Group handover is an inherent problem for HSR wireless systems, caused by highly concentrated users inside carriages that share similar channel quality. Highly frequent handover is also an inherent problem because of high mobility, which is aggravated by mmWave since the effective coverage of mmWave is generally smaller than that of microwave under the same power constraints to combat higher path loss.

Channel Status Information Feedback Delay: The communication quality is more sensitive to channel status information feedback delay because of the fast time-varying channel caused by high mobility. Furthermore, MIMO technologies have stricter requirements for feedback delay and measurement instantaneity.

In this article, we focus on multiple access techniques, frame structures, and train–trackside network architectures that enable HSR mobile systems to perform high rate transmissions by utilizing ultra-wideband mmWave. In addition, we provide several feasible solutions for solving or alleviating the negative effect of the aforementioned design issues.

Multiple Access Techniques and Frame Structure for mmWave in HSR

Generally speaking, orthogonal frequency-division multiple access (OFDMA) is a candidate for the multiple access scheme used in mmWave HSR wireless systems, and has been chosen as that of fourth generation (4G) systems. Due to the high mobility of HSR and short wavelength of mmWave, the frame structure with orthogonal frequency-division multiplexing (OFDM) should be different from the released LTE specifications. Subcarrier spacing is the compromise between symbol length and Doppler shift, the choice of which should be followed by $\Delta f \ll 1/T_{CP}$ and $\Delta f \gg f_{d_{\max}}$, where T_{CP} and $f_{d_{\max}}$ represent the length of cyclic prefix and the maximum Doppler shift, respectively [4]. For example, as for a 28 GHz mmWave band, the maximum Doppler shift is about 13 kHz under 500 km/h train speed; therefore, $\Delta f \gg 13$ kHz. Besides, as HSR belongs to a rural scenario with fewer multipath components; the multipath spread is less than 650 ns normally [5]. Assuming that cyclic prefix (CP) length is 1 μs, the subcarrier spacings are 200 kHz, 62.5 kHz, 40 kHz, and 20 kHz calculated by $\Delta f = 1/T_{symbol}$, under 1/5, 1/16, 1/25, and 1/50 CP proportion in symbol length, respectively (the normal and extended CP proportion are 5.2 μs/71.9 μs ≈ 1/14 and 16.7 μs/83.3 μs ≈ 1/5, 33.3 μs/166.7 μs ≈ 1/5, respectively, in LTE specifications [4]). Apparently, 200 kHz, 62.5 kHz, and 40 kHz can all satisfy the maximum Doppler shift constraint. However, in view of transmission efficiency for each symbol, the last two are more reasonable. In LTE specifications, the sub-frame length, the minimum time unit for resource re-allocation, is 1 ms, which includes about 62 symbols under 62.5 kHz carrier space and 16 μs symbol length. Considering the sensibility of channel quality feedback delay for HSR communications and MIMO technologies, the gap between channel quality measurements and the corresponding transmissions should be as short as possible. To achieve this goal, we propose two improvements to guarantee the real-time performance. First, we appropriately lower the time granularity of resource re-allocations by shortening the sub-frame length. Shorter sub-frame length can shorten the interval between reception of channel quality feedback and the corresponding transmissions. Moreover, according to the current release of LTE specifications, the gap between measurements and feedback must be longer than 4 sub-frames [6], which can be shortened by shorter sub-frame length. However, it is noticeable that shorter sub-frame length also causes negative effects in two main aspects: the system complexity and control signaling overhead. Obviously, shorter sub-frame length brings more sub-frames and more frequent system optimization opportunities, including adaptive modulation and coding (AMC) reselection, resource re-allocation, channel measurement, feedback, and so on. Accordingly, higher system complexity would be induced. Additionally, with more sub-frames, the system will need more wireless resources on control signaling, such as control format indication, acknowledgment/negative acknowledgment (ACK/NACK), scheduling, and so on, to guarantee the reliable reception of each sub-frame. Second, for time-division duplex (TDD) systems, the feedback must be performed in uplink sub-frames. If there is no uplink sub-frame available right after 4 sub-frames behind measurements, the feedback has to be postponed until an uplink sub-frame appears, which brings in more delay. Thus, we propose that the feedback should be performed in the first uplink sub-frame after measurements rather than waiting 4 sub-frames. Based on this observation, we design a feasible OFDM frame structure for 28 GHz bands used for HSR mmWave systems, as shown in Fig. 1a. Using 62.5 kHz carrier spacing, recommended for 28 GHz in HSR systems, as an example, we propose that the length of one sub-frame is 250 μs, which includes 15 symbols. Without loss of generality, the length of radio frame is 10 ms, the same as that defined in LTE. A radio frame consists of 40 sub-frames, in which there are totally 8 special sub-frames, uniformly distributed in the radio frame. The distribution density of reference signal (RS) is determined by coherence time and coherence bandwidth. Under the condition

The communication quality is more sensitive to channel status information feedback delay because of the fast time-varying channel caused by high mobility. Furthermore, MIMO technologies have stricter requirements for feedback delay and measurement instantaneity.

Undoubtedly, SC has lower system complexity than OFDM, but it cannot perform transmission parameter adjustment with fine granularity in the frequency domain as can OFDM.

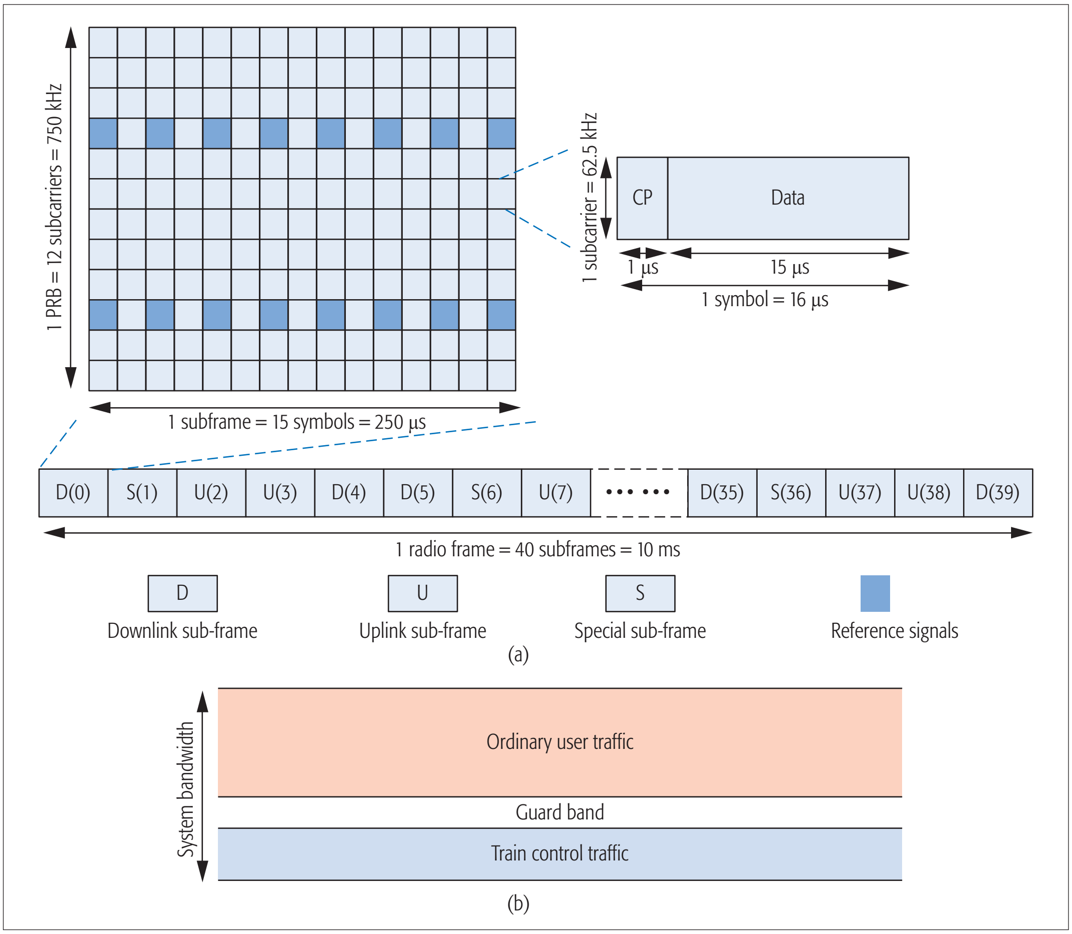

FIGURE 1. The feasible frame structures for mmWave in HSR: a) OFDM; b) SC.

of 1 µs delay spread and 500 km/h speed, the coherence time and coherence bandwidth of 28 GHz band are 33 µs and 500 kHz, respectively. As for the frame structure we designed, in each sub-frame, the frequency domain should be configured with two uniformly distributed RSs, and one RS should be inserted every two symbols in the time domain.

Besides OFDM, single carrier (SC) is another promising scheme to support mmWave transmission due to multiple advantages, such as lower peak-to-average-power ratio (PAPR), lower system complexity, and more available modulation and coding schemes. In order to avoid penetration loss, it is recommended to install onboard mobile relay stations (MRSs) to replace users inside the trains to directly connect with trackside BSs. Under the configuration of MRSs, the train–trackside system can be viewed as a single-user system, in which SC is more suitable than OFDM, because as main advantages of OFDM, multi-user diversity cannot be implemented in a single-user system. We suggest adopting single-carrier frequency domain equalization (SC-FDE) with a simpler frequency domain equalizer. Although the train–trackside system with MRSs can be treated as a single-user system in terms of geographical location and channel quality, two types of traffic, ordinary user traffic and train control traffic, should be considered because of their very different quality of service (QoS) requirements. As shown in Fig. 1b, the whole system bandwidth is divided into two parts, supporting ordinary user traffic and train control traffic, respectively, between which the guard band is inserted to prevent Doppler shift. Since the reliability of train control traffic directly concerns train safety, the link adaptive technologies (LATs) for control signaling should follow the reliability priority principle, which means that efficiency could be sacrificed to guarantee reliability. In contrast, the throughput priority principle followed by ordinary user traffic should pursue throughput maximization. Through the proposed SC frame structure, ordinary user traffic and train control traffic could be configured with different LAT types, depending on effectiveness and reliability, respectively, and performing asynchronous transmissions. Additionally, the system bandwidth could be divided into multiple parts, and the two kinds of traffic utilize multiple discrete parts to obtain frequency diversity. However, this method would result in more overhead of guard bands.

Undoubtedly, SC has lower system complexity than OFDM, but it cannot perform transmission parameter adjustment with fine granularity in the frequency domain as can OFDM. The future train–trackside system may utilize ultra-wide bandwidth, even more than 1 GHz, on mmWave bands. On such wide bandwidth, the channel quality on different frequencies may differ greatly, especially in an environment with interference. In this case, the data rate of OFDM would significantly exceed SC. Thus, we suggest that train–trackside mmWave systems should deploy both OFDM and SC, and dynamically select transmission schemes according to actual radio environments.

Train–Trackside mmWave Network Architecture

In our proposed train–trackside network architecture, MIMO technologies and MRSs are adopted to overcome high path loss and penetration loss, respectively. Additionally, configuration of MRSs can solve the group handover problem completely. For efficient signal processing and network optimization, our proposed architecture builds on the cloud radio access network (C-RAN), which is a promising architecture for future mobile networks. The key idea of C-RAN is that by leveraging cloud computing technologies, the relevant procedures related to storage, computation, and management, originally operated by RF units, could be migrated to cloud computing to reduce system complexity, improve flexibility, perform more efficient signal processing, and achieve centralized resource allocation and network optimization [7]. For our design, in trackside segments, dense ground remote radio heads (GRRHs) provide linear coverage along railways, and are only in charge of RF signal transmissions. A ground baseband unit (GBBU) pool, in charge of centralized signal processing and network optimization, connects to GRRHs by fronthaul networks, which can use fibers as the ideal links. Each GRRH is a MIMO antenna or antenna array, which is determined by the MIMO technologies employed in the architecture. Due to the potential gain caused by MIMO technologies, mmWave should not be a significant inherent disadvantage compared to lower frequencies not only in terms of coverage range but also in terms of received signal quality. Similarly, for onboard segments, to mitigate the interference between beams for BF and attain the spatial diversity for SM, MRSs, as onboard RRHs (ORRHs), are deployed on the train in a distributed way. All MRSs connect to the onboard BBU (OBBU) pool by fronthaul networks. According to the current research findings, the most likely mmWave bands that could be used in future mobile systems fall between 10 GHz and 60 GHz. However, 60 GHz band has high oxygen absorption and has been allocated as unlicensed uses by 802.11ad, wireless HD, and so on [2]. Considering that the path loss increases with frequency, we recommend that the train–trackside system used in the outdoor scenario should choose the mmWave bands below 30 GHz.

To compare the performance of the train–trackside mmWave systems designed based on different MIMO technologies, the major network parameters should be the same. It is assumed that the length of a train is 200 m with 4 ORRHs deployed on it. The parameters of the distance between ORRHs and between GRRHs should be determined by the compromise between throughput and system complexity in addition to handover frequency. Under the assumption of the train length and the number of ORRHs deployed, we adopt 60 m and 50 m as the distance between ORRHs and between GRRHs, respectively, which could perform the appropriate compromise [8]. Besides, the power constraint P_t for each GRRHs is 33 dBm, and the carrier frequency is 28 GHz with 500 MHz system bandwidth. Other simulation parameters are listed in Table 1.

Beamforming-Based Architecture

Figure 2a shows a potential BF-based architecture, in which each active ground antenna array (GAA) communicates with one unique onboard antenna array (OAA) by a narrow beam. BF gain to improve received signal-to-interference-plus-noise ratio (SINR), coupled with the spatial-division multiple access (SDMA) realized by multiple beams, can significantly enhance spectral efficiency. To maintain communication quality, inter-beam handover is inevitable. When the channel quality of a beam (shown by red dotted lines) received by one OAA reaches the handover threshold, the system will choose an idle GAA with the best channel condition for this OAA (shown by bold green lines) as target handover GAA. Under the control of the GBBU pool, the procedure of inter-beam handovers is simpler than that proposed in [11], which just executes BF training between the target GAA and OAA. Nevertheless, the transient transmission termination caused by a hard handover still exists. The negative impact will become much more serious with frequent inter-beam handovers because of small coverage of each beam and high mobility. Moreover, handover failure will increase with the growth of handover frequency. To solve this problem, we improve our proposed BF architecture by adding redundant transmission components. As Fig. 2b illustrates, one OAA selects N GAAs with the best channel quality to connect with, and each of these selected GAAs sends a beam with the redundant information toward this OAA. Accordingly, one active GAA may send up to four beams toward four different OAAs, and the total power of these beams is limited by the power constraints. For inter-beam handovers, when one of N beams received by one OAA triggers a handover, the system will select a GAA (idle or active) as the target handover GAA, which is not communicating with this OAA and has the best channel condition for this OAA. Obviously, when one of N beams is executing the handover, there are still N – 1 beams to provide redundant transmissions. As a result, no transient transmission termination happens during the inter-beam handover, called seamless handover. Moreover, the spatial diversity brought by multi-beam redundant transmissions can enable each OAA to always maintain high data rate. The throughput of this architecture can be expressed as

$$B \cdot \sum_{i \in \Omega_i} \log \left(1 + \frac{\sum_{k \in \Psi_i} h_k \cdot P_k}{N_0 + \sum_{l \notin \Psi_i} h_l \cdot P_l} \right)$$

where Ψ_i stands for the set of beams that points toward OAA i; P_k stands for the power of beam k; $h_k = c_k \cdot \alpha_k(\theta_k) \cdot \beta_k(\varphi_k)$ represents the channel gain of beam k, in which c_k, $\alpha_k(\theta_k)$ and $\beta_k(\varphi_k)$ are propagation loss, transmit BF gains, and receive BF gains, respectively; $h_l = c_l \cdot \alpha_l(\theta_l) \cdot \beta_l(\varphi_l)$ is the interference caused by beam l. In simulation, it is assumed that the power is equally distributed for all beams transmitted by one GAA, N = 3, and the beamwidth of the main lobe is 10° in simulation. Moreover, the threshold of the inter-beam handover is 12 dB, which means that when the received SINR of a beam is less than or equal to 12 dB, a handover is

We propose a beamforming algorithm based on the Grey Markov chain method which is suitable for HSR because the Grey theory component can precisely reflect the strong regularity of HSR and Markov chain component corresponds to irregular factors, that is, the abrupt speed change moments.

Parameters	Value	Parameters	Value
Height of GRRH	10 m	Path loss	157.4 + 32log 10*d* km)
Distance between GRRH and track	30 m	Thermal noise density N_0	–174 dBm/Hz
Height of ORRH	1 m	Lognormal shadowing standard	4 dB
Antenna array model	Uniform linear array	Channel model	Rician fading channel

TABLE 1. System parameters [9, 10].

triggered. There are 11 BRRHs in the range of 500 m, located in 50*k* (*k* = 0, 1, ..., 10) positions of the X-axis. From the simulation results presented by Fig. 2c, we observe that without redundancy, the throughput declines severely in some positions where the beam is close to handover or executing handover, shown by the phenomenon that all lower throughput (below 2000 Mb/s) points are distributed around the handover points. In particular, this phenomenon would be worse when the distance between OAAs is close to an integer multiple of the distance between GAAs, such as 60 m and 50 m in the simulation of this article, which may cause all the beams close to handover point to have bad channel quality, simultaneously. Obviously, our proposed redundancy scheme can solve this problem effectively because of the spatial diversity brought by multi-beam redundant transmissions and seamless handovers.

Proposed Beamforming Algorithms and Training Scheme for HSR

However, the simulations perform under the assumption that accurate BF weights could be obtained, which are difficult to realize in fact by traditional BF algorithms based on channel reciprocity at the high speed of HSR (350 km/h currently and maybe 500 km/h in the future). Maiberger *et al.* [12] introduces a location assistant BF algorithm for HSR, and the simulations provided testify that the performance is outstanding. However, this method must work under the premise that location information is known by other equipments, such as GPS. Since the train only runs along the track, the change of BF weights may possess strong regularity. We propose a beamforming algorithm based on the Grey Markov chain method which is suitable for HSR because the Grey theory component can precisely reflect the strong regularity of HSR, and the Markov chain component corresponds to irregular factors, that is, abrupt speed change moments [13]. For a TDD system, the downlink (DL) beamforming weights are generally estimated by the angle of arrival (AOA) of uplink (UL), which can be accurately estimated using the Multiple Signal Classification (MUSIC) or Estimation of Signal Parameters via Rotational Invariance Technique (ESPRIT) algorithms [14]. For example, for a line of sight (LoS) channel, the typical channel model for HSR, the channel vector of UL, can be expressed as $\boldsymbol{h}^*(\theta) = a[1, e^{-j\pi\cos\theta}, \cdots\cdots, e^{-j\pi(M-1)\cos\theta}]$, where a, M, and θ stand for channel gains, the number of received antennas, and AOA of UL, respectively. For maximal SINR, the BF weights can be derived by $\boldsymbol{w} = \boldsymbol{h}/||\boldsymbol{h}||$. In this article, we use the Grey Markov chain method to eliminate the residual error between the estimated AOA by UL and the optimal angle for DL angle of departure (AOD) caused by high mobility.

The Grey Markov chain prediction consists of four main steps [13]. Assume that $T - i (i = 1, 2, \cdots)$, T, and $T + 1$ represent the past, present, and future, respectively. First, by inputting the actual value array to the GM(1,1) model of Grey systems, we obtain predicted values of each time, $\cdots \bar{X}_{T-1}$, $\bar{X}_T$, $\bar{X}_{T+1}$. Second, we calculate the residual error between the actual value and the predicted value of the past and present time by $e_t = X_t - \bar{X}_t$, $t = \cdots T - 1, T$. Third, by dividing the residual error state, calculating the state transition probability and the sum of transition probability of each state, we find the state with the greatest sum of transition probabilities. Finally, we calculate the predicted value of $T + 1$ time by $\hat{X}_{T+1} = \bar{X}_{T+1} + \bar{e}_{T+1}$, where $\hat{X}_{T+1}$ and $\bar{e}_{T+1}$ are the predicted value of $T + 1$ time output by the Grey Markov chain and maximum expected residual error obtained from the discovered state, respectively. Building on different feedback conditions, two kinds of prediction methods are put forward here. As shown in Fig. 3a, we assume that a TDD system is configured with the distributed DL and UL sub-frame structure. Additionally, the X-axis stands for the location of the train, and θ is defined as the angle of clockwise rotation from the horizontal axis to line of sight (LoS). The UL AOAs of the past are assumed to be accurately estimated by MUSIC or ESPRIT. The main approaches for the algorithm are developed as follows.

•Input the UL AOAs of the past history $[\theta_{T-2n}, \cdots\cdots, \theta_{T-2}, \theta T]$, $n = 1, 2, \ldots$, into the Grey Markov chain model to obtain the prediction value $\hat{\theta}_{T+2\_UL}$.

•Calculate the estimated angle of DL in $T + 1$ time by

$$\hat{\theta}_{T+1\_DL} = \operatorname{arccot} \frac{\cot\theta_{T\_UL} + \cot\hat{\theta}_{T+2\_UL}}{2}.$$

•Update the channel vector by $\mathbf{h}^*(\hat{\theta}_{T+1\_UL})$, which is utilized to obtain beamforming weights. It is not difficult to see that the train in $T + 1$ time is assumed to be located in the midpoint between the locations of T time and the predicted locations of $T + 2$ time. Considering that the interval between two UL sub-frames is extremely short (2 ms according to TD-LTE specifications), this method can obtain precise predicted angle values in most cases, while acutely abrupt speed changes can still degrade accuracy. However, as for the HSR mobile systems, the deployment of OAA raises the possibility of providing DL AOA estimation and feedback for systems, aiming to obtain more precious predicted angle values. The algorithm under the DL AOA estimation is illustrated as:

- Estimate the DL AOA by OAA and perform the feedback in the next UL sub-frame.
- Calculate the residual error between the UL AOA and DL AOA by $e_{T-(2n-1)} = \theta_{T-(2n-1)\_DL} - \theta_{T-(2n)\_UL}$, and input the residual error array, $\mathbf{e} = [e_{T-(2n-1)}, \cdots\cdots, e_{T-3}, e_{T-1}]$, into the Grey Markov chain model to obtain the predicted residual error value $\hat{e}_{T+1}$.
- Calculate the DL AOA in $T + 1$ by $\hat{\theta}_{T+1\_DL} = \theta_{T\_UL} + \hat{e}_{T+1}$, and update the channel vector and the BF weight vector.

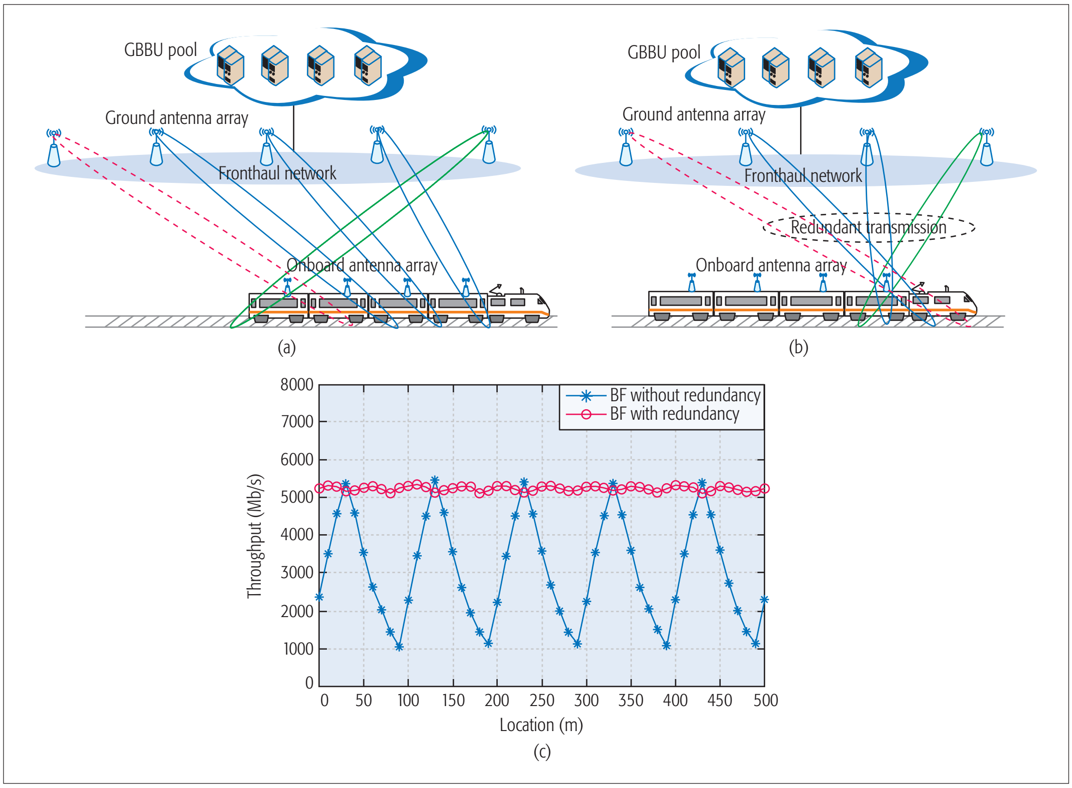

FIGURE 2. The potential BF-based architectures: a) without redundancy; b) with redundancy; c) comparison of throughput.

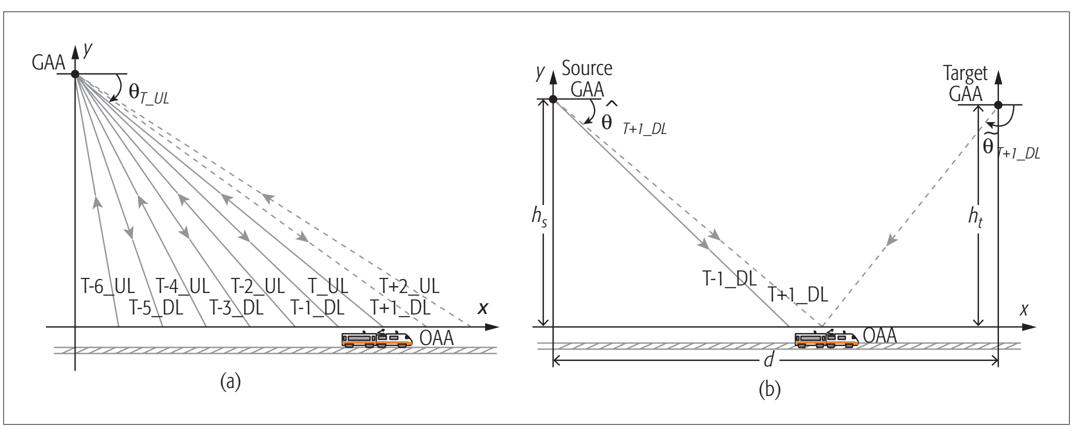

FIGURE 3. Solutions for BF issues in HSR; a) proposed BF algorithm; b) proposed BF training scheme.

Since most of the current BF based mmWave systems are used for point-to-point communications, the fewer research achievements about the beam synchronization in mobile systems have been released. The exhaustive training and multi-level training schemes have been adopted by 802.15.3c and 802.11ad specifications.

Another inevitable issue for BF in HSR is beam synchronization, also referred to as BF training, aiming to search for the best beam direction pointing toward the receiver. BF training is a crucial component for initial access and inter-beam handover. Since most of the current BF-based mmWave systems are used for point-to-point communications, fewer research achievements about beam synchronization in mobile systems have been released. Exhaustive training and multi-level training schemes have been adopted by 802.15.3c and 802.11ad specifications [15, 16]. However, those schemes, with long training delay, cannot adapt to the low delay requirements of HSR. In this article, we propose a fast and precise BF training scheme for inter-beam handover, which consists of three main steps. As shown in Fig. 3b, it is assumed that DL sub-frame in $T - 1$ time is the last sub-frame, like handover recommend signaling, received by OAA from source GAA before executing inter-beam handover. First, predict the DL angle in $T + 1$ time $\hat{\theta}_{T+1\_DL}$ according to the aforementioned schemes in the proposed BF algorithms. Second, calculate the predicted DL angle of the LOS between target GAA and OAA in $T + 1$ time by

$$\hat{\theta}_{T+1\_DL} = 180° - \arctan\frac{h_t}{d - h_s \cdot \tan\hat{\theta}_{T+1\_DL}},$$

where d, h_s, and h_t are the distances between source GAA and target GAA, between source GAA and railway, and between target GAA and railway, respectively. Finally, determine the optimal beam direction by the BF weight vector calculated by θ_{T+1_DL}.

The handover execution delay is directly related to the transmission termination length caused by hard handover. In a C-RAN based system, handover execution could be greatly simplified, because many handover procedures, such as admission control, SN status delivery and data forwarding, and so on, are operated by GBBU pool, and the corresponding signaling flow is eliminated.

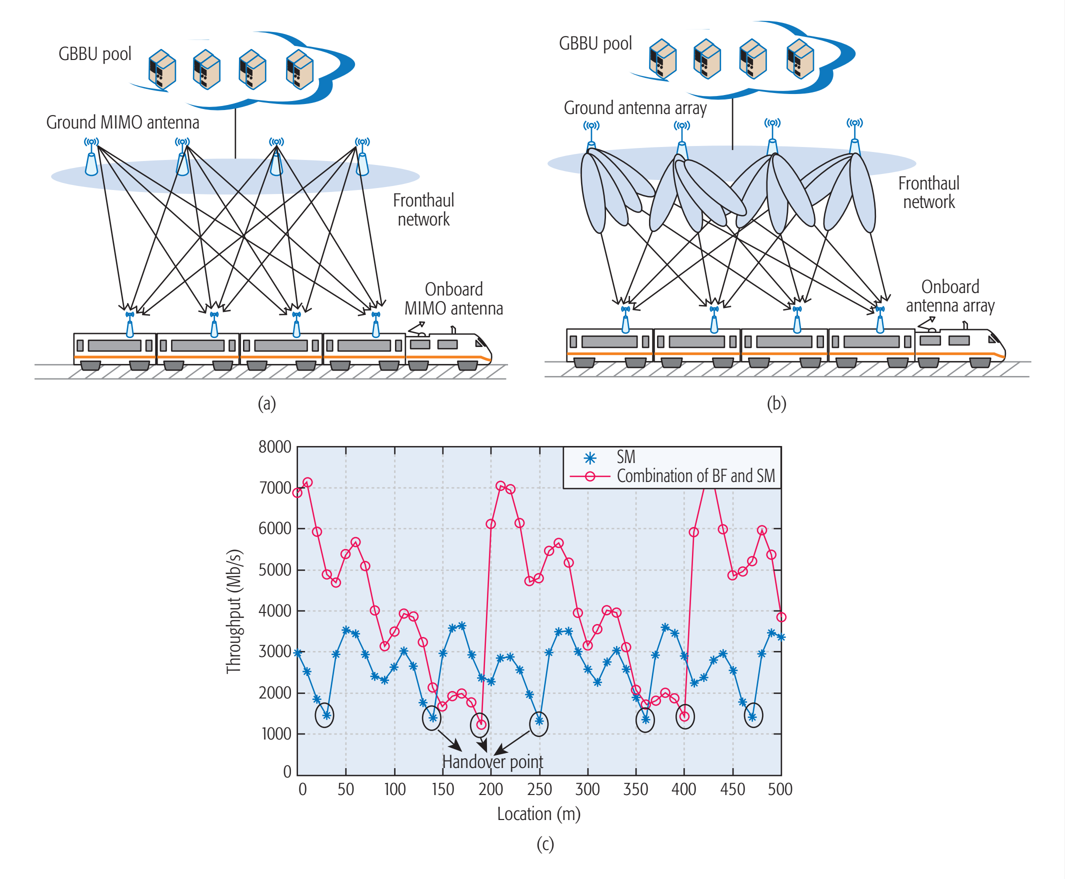

FIGURE 4. The potential SM-based architectures: a) without BF; b) combined with BF; c) comparison of throughput.

SPATIAL-MULTIPLEXING-BASED ARCHITECTURE

SM can greatly enhance spectral efficiency by multi-stream transmissions and has been seen as a potential technology to support outdoor mmWave communications. We propose a distributed MIMO (D-MIMO)-based architecture, shown in Fig. 4a, which has been shown to have better performance than collocated MIMO (C-MIMO) because of the weak channel spatial correlation and high spatial degrees of freedom [17]. It is noteworthy that conventional MIMO is configured with omnidirectional antennas, the coverage of which may be far smaller than BF under the same condition of mmWave frequency and power limitation. For current antenna selection algorithms, aimed at spectral efficiency and energy efficiency, the ground MIMO antenna reselection, referred to as MIMO handover in this article, would be very frequent, especially under high mobility. Setting a reasonable MIMO handover threshold could alleviate this problem; for example, handover is triggered only if the system throughput is lower than a threshold. With frequent MIMO handovers, the time of onboard MIMO antennas transmitting with the same ground MIMO antennas assembly is short. In such a situation, the MIMO handover execution delay and handover trigger hysteresis are critical for transmission efficiency. To be more specific, the handover trigger hysteresis would postpone handover execution and extend the transmission time under deteriorating channel conditions. Besides, the handover execution delay is directly related to the transmission termination length caused by hard handover. In a C-RAN-based system, handover execution could be greatly simplified, because many handover procedures, including admission control, sequence number (SN) status delivery, data forwarding, and so on, are operated by the GBBU pool, and the corresponding signaling flow is eliminated. Moreover, some procedures, such as contention access and radio resource control (RRC) reconfiguration, are not needed anymore. As for shortening the handover trigger hysteresis, there are three main methods:

- Lower the complexity of antenna selection algorithms to decrease calculation time.
- Shorten the time of measurement reports.
- Carry out handover pre-triggering and pre-decision.

Lowering algorithm complexity will degrade performance, and the calculation time can be short enough with strong calculation ability of the GBBU pool. The report time is always determined by the air interface, which is hard to control. Therefore, we adopt the third approach and propose the scheme as follows. First, by taking advantage of the strong regularity of the channel quality changes in HSR, we predict the channel quality of each link between D-MIMO antennas by inputting the previous actual measurement values into the prediction algorithm. Second, we perform the handover triggering and antenna selection algorithms according to the prediction values, and execute the handover in advance. With our proposed scheme, when the MIMO handover is triggered, the channel quality of the source MIMO assembly has not become too bad to trigger actual handover. Hence, the system always transmits under good channel conditions. Unfortunately, the aforementioned schemes can mitigate the negative impact of frequent MIMO handovers, but the improvement is limited. For further improvement, we propose a combination of BF- and SM-based architecture as shown in Fig. 3b. The main idea is to utilize the BF to replace the omni-antenna as the D-MIMO antenna. For example, in Fig. 4b, each GAA sends four beams that point toward four different OAAs, and transmits the same data stream. In this way, the BF gain for a MIMO channel brings higher SM capacity, which can be computed as $B \cdot \log[\det(I + (P_t)/(N_0)Q]$, where $Q = HH^H$ or H^HH, and H is the channel matrix [3]. With the combination of BF, the element of H can be expressed as

$$h_{ij} = c_{ij} \cdot \left[\alpha_{ij}(\theta_{ij}) \cdot \beta_{ij}(\varphi_{ij})\right] + \sum_{p \neq i} c_{pj} \cdot \left[\alpha_{pj}(\theta_{pj}) \cdot \beta_{pj}(\varphi_{pj})\right].$$

Since SM is achieved by singular value decomposition (SVD), the received energy of receiver i from the side lobe of the beams, which does not point toward i, is not interference but gain. Accordingly, $c_{ij} \cdot [a_{ij}(\theta_{ij}) \cdot \beta_{ij}(\varphi_{ij})]$ and $c_{pj} \cdot [a_{pj}(\theta_{pj}) \cdot \beta_{pj}(\varphi_{pj})]$ represent the main lobe gain and side lobe gain, respectively. From the simulation results shown in Fig. 4c, we can see that the throughput when the combination of BF and SM architecture is employed considerably exceeds that for SM-based architecture. Furthermore, to estimate the MIMO handover frequency, we set the handover threshold to be the one in which the system throughput is lower than or equal to 1.5 Gb/s. Obviously, the handover frequency for the combination of BF and SM (about one handover every 210 m) is significantly lower than that for only SM (about one handover every 120 m), which indicates that the former can further improve the frequent handover problem.

Analysis and Performance Evaluation

Table 2 presents the system performance, including average throughput and handover severity (with our proposed solution schemes), and the system budget, including system and handover complexity, for our proposed architectures based on different MIMO technologies. Apparently, under the same simulation parameters shown in Table 1, the average throughputs for all our proposed architectures can achieve high-level throughput (the average throughput for SM and that for the combination of BF and SM are limited by MIMO handover thresholds, which would be larger if the MIMO handover threshold is set to be higher). The average throughput for SM is far below the others. However, this high throughput builds on the high system complexity, particularly for the combination of BF and SM, which is realized based on two kinds of complicated MIMO technologies. Moreover, the complexity of BF is higher than SM because high mobility aggravates the operational difficulty for BF, due to accurate BF weights being difficult to obtain, which may be assisted by a more complex BF algorithm, such as our proposed BF algorithm based on a Grey Markov chain. As for handover complexity, except for the synchronization between ORRHs and target handover GRRHs, BF and SM also need to perform BF training and MIMO antennas reselection, respectively, causing higher complexity. Undoubtedly, the combination of BF and SM possesses the highest handover complexity, which needs to perform both BF training and reselection. Frequent handovers are a fundamental problem for HSR, which becomes even worse in mmWave systems with small cell coverage. By our proposed BF with redundancy scheme, seamless handover can be achieved, and the simulation results demonstrate that this scheme can solve the handover problem effectively, and achieve the slightest level of severity. For SM, our proposed pre-triggering and pre-decision can alleviate the negative effect of the MIMO handover problem by maintaining communications under superior channel condition, but the MIMO handover still happens frequently. Besides, the MIMO handover frequency can be brought down significantly by the combination of BF and SM, mostly half of that with only the SM scheme used, as shown in Fig. 4c.

	BF (redundancy)	SM	Combination of BF and SM
Average throughput (Mb/s)	5210	2680	4460
System complexity	High	Medium	Highest
Handover severity	Slight	Serious	Medium
Handover complexity	High	High	Highest

Table 2. Comparison of performance and budget.

Conclusion

3–300 GHz mmWave bands have the potential to provide multi-gigabit rate radio access for future HSR mobile communication systems. However, due to inherent channel hostility of mmWave propagation and the particularity of HSR scenarios, many challenges will be encountered in actual applications. In this article, we have developed multiple access techniques and frame structures based on OFDM and SC. To solve the high Doppler shift problem and relieve the feedback delay sensitivity problem, we redesign the technical details of ODFM frame structure, such as subcarrier spacing, symbol and sub-frame length, and feedback mechanism. In terms of SC, we design a feasible SC-FDE frame structure, in which ordinary user traffic and train control traffic are supported by different bands following different QoS requirements. Moreover, building on C-RAN, we design several train–trackside mmWave network architectures, in which MIMO technologies and onboard mobile stations are adopted to overcome high propagation loss as well as group handovers, respectively. In addition, for inherent disadvantages of our proposed architectures, we have developed corresponding solutions, such as improving BF algorithms by a prediction algorithm, solving frequent inter-beam handover by redundancy, alleviating frequent MIMO handover by pre-triggering and pre-decision, and combining BF with SM. Finally, we compare the performance and implementation of our proposed architectures by analysis and simulation, and we conclude that the high system throughput achieved by BF- and SM-based architectures

The design, improvements, and simulation results provided in this article demonstrate that the potential for mmWave used in future HSR mobile communication systems with multi-gigabit level data rate is tremendous. However, it is noteworthy that there are still significant challenges ahead of us.

comes at the cost of high system complexity and a severe handover problem. However, the redundant transmission scheme can solve the handover problem of BF based architecture effectively, while a pre-triggering and pre-decision scheme can only alleviate the MIMO handover problem, which can be further solved by the combination of BF and SM schemes.

Clearly, the design, improvements, and simulation results provided in this article demonstrate that the potential for mmWave used in future HSR mobile communication systems with multi-gigabit-level data rate is tremendous. However, it is noteworthy that there are still significant challenges ahead of us, including the BF algorithm for high mobility, MIMO handover, and so on. Hence, solving these remaining challenges is the target of our future research. Furthermore, a critical problem should be considered: how to guarantee the reliability of control signaling information in an mmWave mobile system. A control and data decoupled architecture of HSR could be a potential solution to this issue, in which control signaling information is supported by Marco eNB at the reliable and lower spectra, while Phantom eNB utilizes higher spectra, such as mmWave, to support data transmission and enhance the system capacity [18]. In particular, when GSM for Railway (GSM-R) is replaced by next generation HSR mobile networks, the excellent 4 MHz lower frequency band (876–880 MHz for uplink, 921–925 MHz for downlink in Europe, and 885–889 MHz for uplink, 930–934 MHz for downlink in China) previously allocated to GSM-R can be used to provide reliable control signaling transmission.

Acknowledgment

The work of H. Song and X. Fang was partially supported by the 973 Program of China under Grant 2012CB316100, NSFC under Grant 61471303, and EU FP7 QUICK project under Grant PIRS-ES-GA-2013-612652, Program for Development of Science and Technology of China Railway Corporation under Grant 2015X007-B. The work of Y. Fang was partially supported by the U.S. National Science Foundation under grant CNS-1343356.

References

[1] CISCO Whitepaper, "CISCO Visual Networks Index: Global Mobile Data Traffic Forecast Update 2014–2019," 2015.

[2] Z. Pi and F. Khan, "An Introduction to Millimeter-Wave Mobile Broadband Systems," *IEEE Commun. Mag.*, vol. 49, no. 6, June 2011, pp. 101–07.

[3] S. Sun *et al.*, "MIMO for Millimeter-Wave Wireless Communications: Beamforming, Spatial Multiplexing, or Both?" *IEEE Commun. Mag.*, vol. 52, no. 12, Dec. 2014, pp. 110–21.

[4] S. Sesia, I. Toufik, and M. Baker, *LTE-The UMTS Long Term Evolution: From Theory to Practice*, Wiley, 2012.

[5] 3GPP TR 25.996, "Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations, v6.1.0," Sept. 2003.

[6] 3GPP TS 36.214, "Physical Layer Measurements, v11.1.0," Dec 2012.

[7] R. Wang, H. Hu, and X. Yang, "Potentials and Challenges of C-RAN Supporting Multi-RATs Toward 5G Mobile Networks," *IEEE Access*, vol. 2, Oct. 2014, pp. 1187–95.

[8] P. T. Dat *et al.*, "WDM RoF-MMW and Linearly Located Distributed Antenna System for Future High-Speed Railway Communications," *IEEE Commun. Mag.*, vol. 53, no. 10, Oct. 2015, pp. 86–94.

[9] P. Soma *et al.*, "Propagation Measurements and Modeling of LMDS Radio Channel in Singapore," *IEEE Trans. Vehic. Tech.*, vol. 52, no. 3, May 2003, pp. 595–606.

[10] 3GPP TR 36.814, "Further Advancements for E-UTRA Physical Layer Aspects, v9.0.0," Mar 2010.

[11] S. Oh *et al.*, "An Enhanced Handover Scheme to Provide the Robust and Efficient Inter-Beam Mobility," *IEEE Commun. Lett.*, vol. 19, no. 5, Feb. 2015, pp. 739–42.

[12] R. Maiberger, D. Ezri, and M. Erlihson, "Location Based Beamforming," *Proc. 2010 IEEE 26th Convention of IEEEI*, Nov. 2010, pp. 184–87.

[13] C. L. Fan *et al.*, "Grey Markov Chain and Its Application in Drift Prediction Model of FOGs," *J. Sys. Eng. Electron.*, vol. 16, no. 2, 2005, pp. 388–93.

[14] M. M. Abdalla, M. B. Abuitbel, and M. A. Hassan, "Performance Evaluation of Direction of Arrival Estimation Using MUSIC and ESPRIT Algorithms for Mobile Communication Systems," *Proc. 2013 6th Joint IFIP WMNC*, Apr. 2013, pp. 1–7.

[15] T. Baykas *et al.*, "IEEE 802.15.3c: The First IEEE Wireless Standard for Data Rates over 1 Gb/s," *IEEE Commun. Mag.*, vol. 49, no. 7, July 2011, pp. 114–21.

16] T. Nitsche *et al.*, "IEEE 802.11ad: Directional 60 GHz Communication for Multi-Gigabit-per-Second WiFi," *IEEE Commun. Mag.*, vol. 52, no. 12, Dec. 2014, pp. 132–41.

[17] D. M. Wang *et al.*, "Spectral Efficiency of Distributed MIMO Systems," *IEEE JSAC*, vol. 31, no. 10, Oct. 2013, pp. 2112–27.

[18] L. Yan, X. Fang, and Y. Fang, "Control and Data Signaling Decoupled Architecture for Railway Wireless Networks," *IEEE Wireless Commun.*, vol. 22, no. 1, Mar. 2015, pp. 103–11.

Biographies

Hao Song [S'14] (songhao992013@gmail.com) received his B.E. degree in electrical information engineering in 2011 from Zhengzhou University, China. He is currently a Ph.D student with Key Laboratory of Information Coding and Transmission, Southwest Jiaotong University, Chengdu, China. His research interests focus on mobile network optimization, unlicensed spectra cognitive systems, next generation mobile networks for high-speed railway, and 5G mobile networks.

Xuming Fang [SM'16] (xmfang@home.swjtu.edu.cn) received his B.E. degree in electrical engineering in 1984, M.E. degree in computer engineering in 1989, and Ph.D. degree in communication engineering in 1999, all from Southwest Jiaotong University. In September 1984, he was a faculty member with the Department of Electrical Engineering, Tongji University, Shanghai, China. He then joined the Key Laboratory of Information Coding and Transmission, School of Information Science and Technology, Southwest Jiaotong University, where he has been a professor since 2001 and chair of the Department of Communication Engineering since 2006. He held visiting positions with the Institute of Railway Technology, Technical University of Berlin, Germany, in 1998 and 1999, and with the Center for Advanced Telecommunication Systems and Services, University of Texas at Dallas, in 2000 and 2001. He has to his credit around 200 high-quality research papers in journals and conference publications. He is the author or co-author of five books and textbooks. His research interests include wireless broadband access control, radio resource management, multihop relay networks, and broadband wireless access for high-speed railways. He is an Editor of *IEEE Transaction on Vehicular Technology* and the *Journal of Electronics and Information*, and the Chair of the IEEE Vehicular Technology Society Chengdu Chapter.

Yuguang Fang [F'08] (fang@ece.ufl.edu) received an M.S. degree from Qufu Normal University, Shandong, China, in 1987, a Ph.D. degree from Case Western Reserve University in 1994, and a Ph.D. degree from Boston University in 1997. He joined the Department of Electrical and Computer Engineering at the University of Florida in 2000 and has been a full professor there since 2005. He held a University of Florida Research Foundation (UFRF) Professorship from 2006 to 2009, a Changjiang Scholar Chair Professorship with Xidian University, China, from 2008 to 2011 and with Dalian Maritime University, China, and a Guest Chair Professorship with Tsinghua University, China, from 2009 to 2012. He received the U.S. National Science Foundation Career Award in 2001 and the Office of Naval Research Young Investigator Award in 2002, and was the recipient of the Best Paper Award from IEEE ICNP (2006). He also received the 2010–2011 UF Doctoral Dissertation Advisor/Mentoring Award, 2011 Florida Blue Key/UF Homecoming Distinguished Faculty Award, and 2009 UF College of Engineering Faculty Mentoring Award. He is the Editor-in-Chief of *IEEE Transactions on Vehicular Technology*, was the Editor-in-Chief of *IEEE Wireless Communications* (2009–2012), and serves/has served on several Editorial Boards, including *IEEE Transactions on Mobile Computing* (2003–2008, 2011–present), *IEEE Transactions on Communications* (2000–2011), and *IEEE Transactions on Wireless Communications* (2002–2009). He has been actively participating in conference organizations such as serving as the Technical Program Co-Chair for IEEE INOFOCOM 2014 and Technical Program Vice-Chair for IEEE INFOCOM 2005.