

Wireless channel feature extraction via GMM and CNN in the tomographic channel model

Haihan Li^{1,2}, Yunzhou Li^{2*}, Shidong Zhou^{1,2}, Jing Wang²

1. Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

2. Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, Beijing 100084, China

* Corresponding author, email: liyunzhou@mail.tsinghua.edu.cn

Abstract: Wireless channel modeling has always been one of the most fundamental highlights of the wireless communication research. The performance of new advanced models and technologies heavily depends on the accuracy of the wireless CSI (Channel State Information). This study examined the randomness of the wireless channel parameters based on the characteristics of the radio propagation environment. The diversity of the statistical properties of wireless channel parameters inspired us to introduce the concept of the tomographic channel model. With this model, the static part of the CSI can be extracted from the huge amount of existing CSI data of previous measurements, which can be defined as the wireless channel feature. In the proposed scheme for obtaining CSI with the tomographic channel model, the GMM (Gaussian Mixture Model) is applied to acquire the distribution of the wireless channel parameters, and the CNN (Convolutional Neural Network) is applied to automatically distinguish different wireless channels. The wireless channel feature information can be stored offline to guide the design of pilot symbols and save pilot resources. The numerical results based on actual measurements demonstrated the clear diversity of the statistical properties of wireless channel parameters and that the proposed scheme can extract the wireless channel feature automatically with fewer pilot resources. Thus, computing and storage resources can be exchanged for the finite and precious spectrum resource.

Keywords: wireless channel modeling, tomographic channel model, Gaussian mixture model, convolutional neural network

Citation: H. H. Li, Y. Z. Li, S. D. Zhou, et al. Wireless channel feature extraction via GMM and CNN in the tomographic channel model [J]. Journal of communications and information networks, 2017, 2(1): 41-51.

1 Introduction

The wireless channel is the most important feature of wireless communication and the fundamental difference between wireless and wired communications. Wireless channel modeling plays a very important role in both theoretical and practical research on

modern wireless technology. Thus, channel modeling has always held a unique appeal for wireless communication researchers. Accurate and efficient wireless channel models help meet the demands of advanced technologies for CSI (Channel State Information) to achieve adequate performance in 5G and future wireless communication systems. In other words, the

Manuscript received Jan. 1, 2017; accepted Feb. 8, 2017

This work is supported by the National Natural Science Foundation of China (No. 61631013), National Key Basic Research Program of China (973 Program) (No. 2013CB329002), National Major Project (No. 2014ZX03003002-002), Program for New Century Excellent Talents in University (No. NCET-13-0321).

performance of new advanced technologies heavily depends on the accuracy of CSI.

The common way to acquire CSI is by using pilot symbols or a training sequence. A more complex wireless communication system means a higher frequency cost. Spectrum resources are limited and precious. Thus, the pilot overhead can become intolerable, especially for massive MIMO (Multiple-Input Multiple-Output) antenna systems and UDNs (Ultra-Dense Networks).

Different kinds of wireless channel models have been proposed to better understand the wireless channel. Proper wireless channel models are the basis of accurate CSI. Given the statistical properties of channel parameters, wireless channel models can generally be divided into deterministic and stochastic models.

Deterministic models are based on the belief that wireless channel parameters can be calculated from the propagation environment information and propagation law of the electromagnetic wave. The ray-tracing channel model and its extensions are representative deterministic models^[1]. The ray-tracing model can simulate almost all of the reflection, diffraction, refraction, and scattering by using well-known radio propagation environment information between the transmitter and receiver. Then, the information of almost every propagation path can be calculated theoretically to obtain the CSI of the channel.

GSCMs (Geometry-based Stochastic Channel Models) are one of the most popular stochastic models^[2]. GSCMs can reproduce the stochastic properties of different kinds of wireless channels over time, frequency, and space. Because of the scalability and reasonable complexity of GSCM, it has been widely adopted by many common wireless channel models, such as WINNER II/+, COST2100, and IEEE 802.11ad^[3-7].

The key for deterministic models to acquire sufficiently accurate CSI is sufficiently abundant information on the radio propagation environment. In other words, the accuracy of the CSI calculation result is closely related to the electromagnetic param-

eters that can be determined from the radio propagation environment, such as the reflection coefficient of different materials^[1]. As the required level of accuracy for CSI increases, the calculation process of a deterministic model becomes more complex. Thus, deterministic models are feasible but much too complex to be practical in real wireless communication systems. Stochastic models are more widely used for system design and performance testing because the reflection, diffraction, refraction, and scattering in the radio propagation environment are completely random^[8,9]. Thus, GSCMs are simpler but do not reveal information on the radio propagation environment. Such models cannot reflect the actual CSI in real communication systems.

Obviously, there is a tradeoff between the accuracy of CSI and the computational complexity or spectrum resources. For pilot-aided channel estimation, massive antennas and complex networks cost a huge amount of spectrum resources. Thus, new channel models and methods for acquiring CSI are necessary to reduce the pilot cost and improve the performance of the communication system. Typical propagation scenarios have their own unique radio propagation environment. There is a trend where radio propagation environment information of typical propagation scenarios is used to apply prior information to help model a wireless channel and reduce the cost used to determine the CSI. Some HST (High-Speed Train) channel models divide the HST propagation environment into different typical scenarios, such as open space, viaducts, cutting, hilly terrain, tunnels, and stations^[10-12]. The METIS (Mobile and wireless communications Enablers for the Twenty-twenty Information Society) has introduced the METIS channel models, which consist of a map-based model, a stochastic model, and a hybrid model combining the previous two^[13-15]. The map-based model is based on ray tracing and uses a simplified three-dimensional geometric description of the propagation environment. Thus, it inherently accounts for significant propagation mechanisms such as diffraction, specular reflection, diffuse scattering, and blocking. They reported that the map-based model enjoys ex-

cellent support for massive-MIMO, spherical waves, dual mobility, and other advanced technologies. A few studies based on the map-based model have produced large-scale measurement results. Detailed research on small-scale information or full CSI is needed. Refs. [16,17] presented a general framework for modeling and reproducing environment-specific channel properties in V2V (Vehicle-to-Vehicle) communication. In contrast to the traditional deterministic models or stochastic models, these new models try to introduce some deterministic pattern or feature into the stochastic models to obtain extra information with the purpose of acquiring CSI acquisition at a lower cost. Such advances inspired us to focus on the relationship between the radio propagation environment and CSI in order to incorporate radio propagation environment information of a wireless channel to acquire the CSI at lower cost and with greater accuracy. However, most of these new wireless channel models only focus on large-scale information^[18,19]. Detailed research on small-scale information is needed.

In this study, the relationship between the radio propagation environment and CSI and the statistical properties of wireless channel parameters were analyzed. The diversity of wireless channel parameters was used to help explain the concept of the tomographic channel model. The tomographic channel model consists of static, dynamic, and disturbing CSI^[20]. Based on the tomographic channel model, the GMM and EM (Expectation Maximization) were applied to extract the static CSI from the existing channel measurement campaign data. The statistical properties of the static CSI can guide the pilot design. An online-offline scheme to acquire CSI via the tomographic channel model is proposed. A CNN (Convolutional Neural Network) was applied to identify different wireless channels automatically. The numerical results of the channel measurement campaigns showed that the extracted and stored wireless channel features introduces extra information on the radio propagation environment and helps reduce the pilot cost. In addition, the CNN can automatically identify different wireless channels and help decide

which relevant wireless channel features should be used.

The paper is organized as follows. Section 2 presents the tomographic channel model and online-offline scheme for CSI acquisition. Section 3 describes the application of the GMM-EM algorithm to obtaining static CSI (wireless channel feature) from the huge amount of existing wireless channel data. Section 4 describes how the CNN was applied to identify different wireless channels automatically. Section 5 presents the results of repeated measurements and the performance of the method.

2 Tomographic channel model

The classic multi-path channel model can be expressed as^[21]

$$h(\tau; t) = \sum_{l=1}^L \beta_l(t) \delta(\tau - \tau_l), \quad (1)$$

$$\beta_l(t) = \alpha_l(t) e^{j(2\pi f_l t + \phi_l)} = \widetilde{\beta}_l(t) \cdot e^{j(2\pi f_l t)}, \quad (2)$$

$$\widetilde{\beta}_l(t) = \alpha_l(t) e^{j\phi_l}, \quad (3)$$

$$f_l = f_c \frac{v}{c} \cos \theta_l. \quad (4)$$

where L is the number of MPCs (MultiPath Components), $\beta_l(t)$ is the complex amplitude and τ_l is the time delay of the l th path. f_l is the Doppler frequency and ϕ_l is the phase of the l th path. $\widetilde{\beta}_l(t)$ is the complex amplitude upgraded with the phase of the l th path. f_c is the carrier frequency and θ_l is the DOA (Direction of Arrival) of the l th path. c is the speed of light and v is the velocity of the MS (Mobile Station). If the MS moves at a low speed ($v \ll c$), $f_l \approx 0$ and $\beta_l(t) = \alpha_l(t) e^{j\phi_l}$. Now, the $\beta_l(t)$ is the complex amplitude regardless of the change of DOA and velocity of the MS.

If the channel impulse response is assumed to be wide sense stationary over a small-scale time or distance, the $h(\tau; t)$ can be simplified as

$$h(\tau) = \sum_{l=1}^L \beta_l \delta(\tau - \tau_l). \quad (5)$$

Theoretically, the parameters of each path and the CSI are determined by the radio propagation environment. The radio propagation environment is physically defined by the position and motion of the BS (Base Station), MS, and surrounding scatterers. The scatterers can basically be divided into two kinds: static and moving. Static scatterers mainly include buildings, trees, and so on in the radio propagation environment. Moving scatterers mainly consist of moving people and cars. In a real radio propagation environment, the BS and most of the scatterers are usually static. When the MS moves a short distance, the static scatterers change slowly compared with the receivers motion. Previous research revealed that, because the BS and main scatterers (e.g., outdoor buildings and indoor walls) are static, some variables in the multi-path channel model mainly affected by static scatterers can be considered static^[22,23]. Thus, the number of static scatterers and the DOAs of rays from these scatterers, which are referred to as static parameters, stay relatively invariable. Thus, L and θ_l are usually static and τ_l and β_l can also be regarded as static. The velocity of the receiver ν is a dynamic parameter and needs to be obtained with GPS (Global Positioning System) or other speed testing instruments. Finally, the disturbing CSI refers to the reflection, diffraction, refraction and scattering caused by scatterers that appear randomly and may only affect part of the multiple paths. As mentioned above, the main static scatterers like buildings or walls are much taller than the MS. Thus, the power of the paths caused by the static scatterers is much stronger than that of the single-reflection signal from random scatterers. Therefore, Eq. (1) can be approximately expressed as

$$h(\tau; t) = \sum_{l=1}^{L_1} \beta_l(t) \delta(\tau - \tau_l) + \sum_{l=L_1+1}^{L_2+L_1} \beta_l(t) \delta(\tau - \tau_l), \quad (6)$$

where L_1 is the number of MPCs of the static scatterers and L_2 is the number of MPCs of the random scatterers. Static scatterers make up most of

the CSI. Luckily, they are static, so they can be extracted precisely via repeated prior measurements. This is the “feature” of the channel in each specific real propagation scenario. When the feature of the channel is obtained, the pilot can be designed at the main static time and frequency positions. This means that we already know the basic structure of the wireless channel and can put the pilot at a position where the wireless channel parameters have a large role. Tong et al.^[24] showed that a good pilot design can help obtain a more accurate CSI.

The tomographic channel model consists of static, dynamic, and disturbing channel information. Static channel information is CSI caused by static scatterers and can be extracted from a large amount of existing CSI data offline. The motion of the MS leads to dynamic channel information. The velocity of the MS can be determined with the help of GPS or other speed-testing instruments. Thus, the dynamic channel information can be calculated. Because paths caused by random scatterers seem rare and weak in some propagation scenarios, disturbing channel information can be ignored or estimated by a smaller pilot in practice.

Fig. 1 shows our online-offline scheme to acquire CSI. First, the static parameters can be extracted from the data of previous repeated experiments. The SAGE (Space Alternating Generalized Expectation-Maximization) algorithm is used to restore the multi-path parameters; then, the GMM is applied to obtain the distribution of the multi-path parameters based on the EM principle. The above process is carried out offline. With the distribution of the multi-path parameters, prior information can help in the pilot design. Second, we can obtain the velocity information of the receiver with GPS or other speed-testing instruments. With the velocity information, the dynamic part can be calculated. Finally, the disturbing CSI can be estimated with a well-designed pilot. Thus, the full CSI should be acquired based on the above three steps. Note that the static multi-path information is additive and the dynamic part is multiplicative. The disturbing CSI is always considered additive.

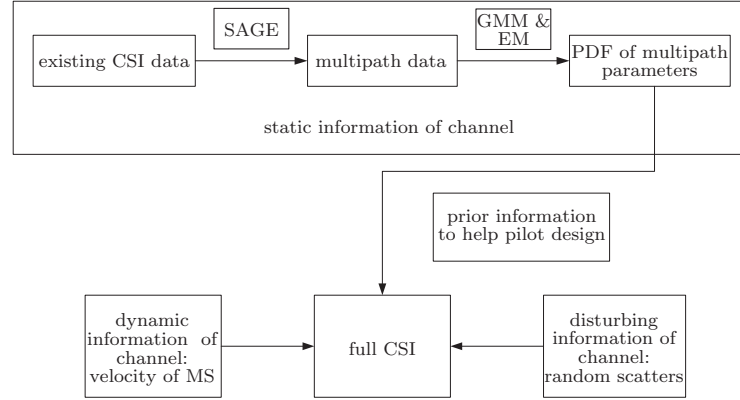


Figure 1 The scheme to acquire CSI in the tomographic channel model

3 Scheme to acquire CSI with the tomographic channel model

3.1 SAGE algorithm

The SAGE algorithm is often applied to parameter extraction from channel measurement data. As an extension of the EM algorithm, the SAGE algorithm classifies parameters (e.g., time delay, amplitude, Doppler frequency, AOA) that need to be estimated into subsets to improve the convergence speed compared with the EM algorithm^[25].

The vector $\Theta_l = (\beta_l, \theta_l, \tau_l, \varphi_l)^T$ is the set of wireless channel parameters. The SAGE algorithm classifies Θ_l into S subsets $\{\Theta_s, s = 1, \dots, S\}$. Thus, the log-likelihood function can be expressed as^[25]

$$\begin{aligned} A(\Theta_s; x_{ls}) &= \frac{1}{N_0} \left[2 \int_{D_0} \text{Real} \left\{ H(\Theta_s | \widehat{\Theta}_{\bar{s}}, t)^H \right\} x_{ls}(t) dt \right. \\ &\quad \left. - \int_{D_0} \left\| H(\Theta_s | \widehat{\Theta}_{\bar{s}}, t) \right\|^2 dt \right], \end{aligned} \quad (7)$$

$$(\widehat{\Theta}_s)_{ML}(x_{ls}) = \text{argmax} \{ A(\Theta_s; x_{ls}) \}, \quad (8)$$

x_l is the observable data. x_{ls} is the admissible hidden-data space with respect to Θ_s in the l th path. N_0 is the noise variance. D_0 is the interval of observation. $\widehat{\Theta}_{\bar{s}}$ is the given complementary set of Θ_s in Θ_l .

The ML estimate of the subsets Θ_s instead of Θ_l is needed to improve the convergence speed. L is determined by the MDL (Minimum Description Length) criterion. With L is identified, $\beta_l, \theta_l, \tau_l, \varphi_l$

can be estimated off-line via SAGE algorithm. This algorithm estimates and updates Θ_l iteratively until it achieves convergence.

3.2 Gaussian mixture model

The GMM is a popular parametric density model for approximating arbitrary continuous multivariate PDFs (Probability Density Functions) in unsupervised learning or clustering procedures^[26,27]. The GMM is made up of K multivariate Gaussian distributions known as mixture components. Each component has its own $\{\mu_k, \Sigma_k\}$, $k = 1, \dots, K$, where μ_k is the mean value vector and Σ_k is the covariance matrix of the multivariate Gaussian component. Each mixture component also has an associated non-negative mixing weight π_k . The PDF (Probability Density Function) for GMM can be expressed as

$$p(x) = \sum_{k=1}^K \pi_k \cdot \mathcal{N}(x | \mu_k, \Sigma_k). \quad (9)$$

$\Theta = (\beta, \theta, \tau, \phi)^T$ can be regarded as four dimensional variables. $K \approx L$. The PDF of Θ can be expressed as

$$p(\Theta) = \sum_{l=1}^L \pi_l \cdot \mathcal{N}(\Theta | \mu_l, \Sigma_l). \quad (10)$$

3.3 EM algorithm

The next step is to estimate $\{\pi_l, \mu_l, \Sigma_l\}$ to maximize the log-likelihood function $\sum_{i=1}^N p(\Theta_i)$. The EM algorithm can be carried out as follows^[28]:

Expectation step: estimate the probability of Θ_i generated by the l th mixture Gaussian component.

$$\gamma(i, l) = \frac{\pi_l \cdot \mathcal{N}(\Theta_i | \mu_l, \Sigma_l)}{\sum_{j=1}^L \pi_j \cdot \mathcal{N}(\Theta_i | \mu_j, \Sigma_j)}, \quad (11)$$

$\pi_j, \mu_j, \Sigma_j, j = 1, \dots, L$ need to be initialized.

Maximization Step: Estimate the π_l, μ_l, Σ_l .

$$\mu_l = \frac{1}{N_l} \sum_{i=1}^N \gamma(i, l) \Theta_i, \quad (12)$$

$$\Sigma_l = \frac{1}{N_l} \sum_{i=1}^N \gamma(i, l) (\Theta_i - \mu_l)(\Theta_i - \mu_l)^T, \quad (13)$$

$$\pi_l = \frac{N_l}{N}, \quad (14)$$

N is the number of training multi-path data sets. $N_l = \sum_{i=1}^N \gamma(i, l)$. $\{\pi_l, \mu_l, \Sigma_l\}$ can be estimated iteratively.

4 Channel recognition via CNN

Deep learning enjoys excellent performance in automatic speech recognition, face recognition, and text recognition. Deep learning enables an artificial neural network to automatically extract features from the input training data and reduce the incompleteness brought by the artificial settings^[29]. Fig. 2 shows a DNN.

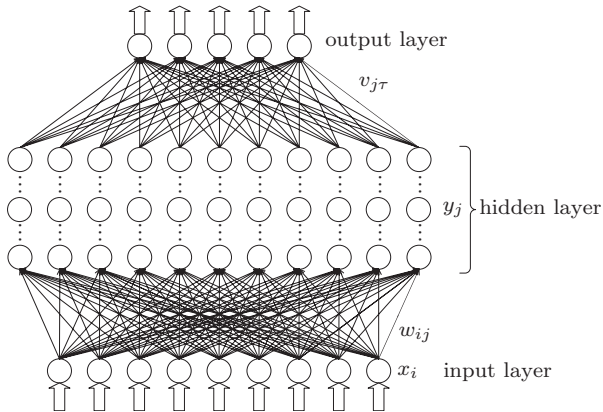


Figure 2 Deep neural network

A CNN is one kind of widely used deep learning structure. A CNN is characterized by consisting of one or more pairs of convolution and pooling layers. The pair of a convolution layer and pooling layer is used for feature extraction^[30]. In the convolution

layer, a set of filters is applied to process small local parts of the input data. These filters are replicated along the whole input space. Neurons in the convolution layer multiply. In the pooling layer, a lower-resolution version of the convolution layer activation is generated by taking the maximum filter activation from different positions within a specified window. The number of neurons in the pooling layer stays the same. The fully connected layer combines inputs from all positions to classify the overall inputs. A CNN has three unique concepts: local filters, max-pooling, and weight sharing. The paths in wireless channel are always clustered. Thus, the nearby paths can share a similar nature and influence each other. As a result, the unique characteristics of a CNN allow it to extract the wireless channel features better than other deep learning algorithms. Different propagation environments lead to different wireless channels. Based on the tomographic channel model, the static part of the CSI data can be regarded as the wireless channel feature. The CSI data we measured under different measurement conditions may contain different wireless channel features. The datasets of multi-path parameters (e.g., time delay, amplitude, Doppler frequency) $\Theta = \{\Theta_1, \dots, \Theta_L, \dots, \Theta_L\}$, $\Theta_l = (\beta_l, \theta_l, \tau_l, \phi_l)^T$, $l = 1, \dots, L$ are the input data in the CNN for wireless channel feature extraction, as shown in Fig. 3. The input data sets comprise many multi-path parameter samples. One sample is a matrix whose row vectors can be the time delay, amplitude, or Doppler frequency. The length of each row vector is the number of paths L . If $L = 10$ and the data of the time delay, amplitude, and Doppler frequency are given, one input data sample can be a 10×3 matrix. The output is the number of wireless channels to which the input dataset belongs. For example, the input dataset may consist of three different wireless channels, which means that the data of the wireless parameters are obtained in three different positions or under three different measurement conditions. The output can be three bits: 001, 010, or 100.

For the feedforward pass, the convolution layer activations can be computed as a convolution-like oper-

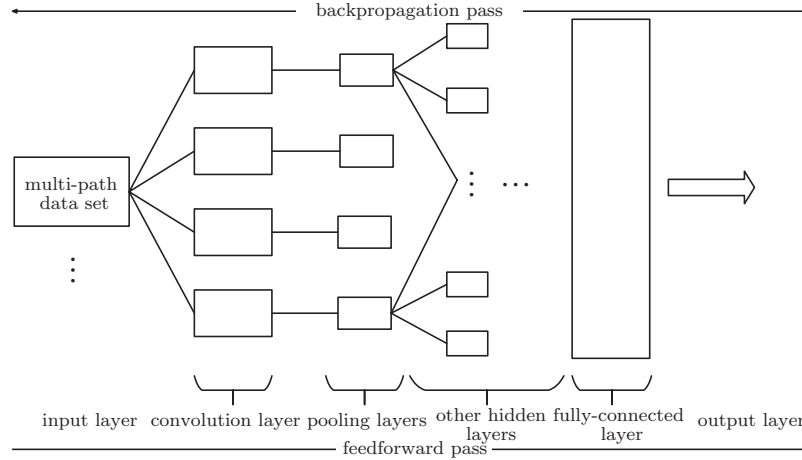


Figure 3 CNN for multi-path data

ation of each filter on the lower layer bands followed by a nonlinear activation function $f(\Theta)^{[31]}$,

$$net = \sum_{i=1}^n w_i x_i + b, \quad (15)$$

where x_i can be any datum in Θ .

$$sigmoid(t) = \frac{1}{1 + e^{-t}} \quad (16)$$

$$f(\Theta) = sigmoid(net) = \frac{1}{1 + \exp[-(\sum_{i=1}^n w_i x_i + b)]} \quad (17)$$

For backpropagation, the loss function is defined to estimate the prediction error. There are many kinds of loss functions. In our CNN, the loss function is defined $E(\Theta)$ as

$$E(\Theta) = \frac{1}{2} \|y - f(\Theta)\|_2^2, \quad (18)$$

where y is the correct channel number and $f(\Theta)$ is the output result of CNN.

The derivative of $E(\Theta)$ passes along the CNN from the output layer to the input layer.

$$\frac{\partial E}{\partial w_i} = \frac{\partial E}{\partial f} \cdot \frac{\partial f}{\partial net} \cdot \frac{\partial net}{\partial w_i}. \quad (19)$$

The parameters of CNN update in the negative gradient direction and gradually approximate the optimal solution with iterations. A backpropagation algorithm with an appropriate learning rate can guarantee convergence to the local minimum but not convergence to the global minimum.

5 Experimental evaluation

To verify the performance of the proposed method, indoor and outdoor channel measurement campaigns were carried out in and around the FIT building in Tsinghua University, Beijing. Tab. 1 presents the system configuration.

Table 1 System configuration of measurement campaigns

	indoor	outdoor
transmitter power	20 dBm	20 dBm
carrier frequency	900 MHz	900 MHz/2.4 GHz
bandwidth	20 MHz	20 MHz
transmitter Antenna Height	1 m	10 m
receiver Antenna Height	1 m	1 m
LOS/NLOS	LOS	LOS

The indoor channel measurement campaigns were carried out in a meeting room. The room was 8.57 m long and 3.91 m wide, as shown in Fig. 4. Tx and Rx were placed on the desk in the center of the room. The horizontal distance between Tx and Rx was 4 m. The desk was about 1 m high. The outdoor channel measurement campaigns were carried out in the small yard in the middle of the building. Tx was placed outside the window on the fourth floor at a height of about 10 m. Rx was placed in the center of the yard in the FIT building at a height of about 1 m. The horizontal distance between Tx and Rx was about 39 m. Figs. 4 and 5 show the indoor

and outdoor layouts, respectively. Throughout the measurements, Tx and Rx stayed still.

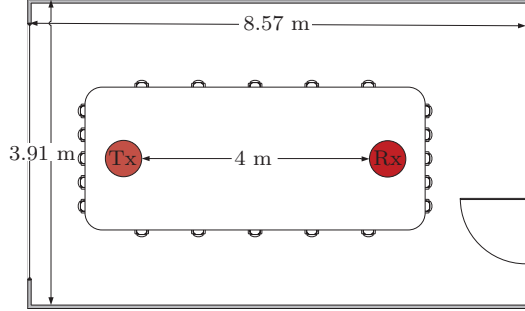


Figure 4 Layout of the indoor experiment environment

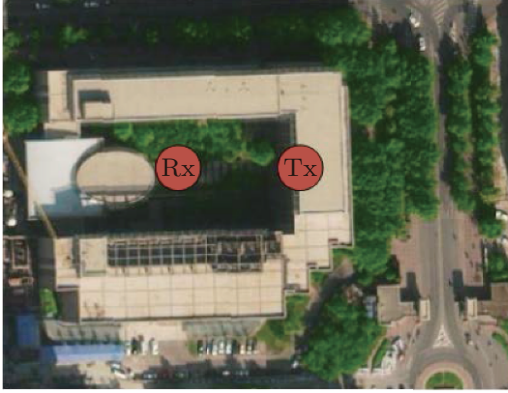


Figure 5 Layout of the outdoor experiment environment

GMM-EM was applied to obtain the PDF of wireless channel parameters. Based on the channel feature, fewer but redesigned pilots were used for channel estimation. There were 64 pilot symbols in the original system. Based on the distribution of the paths, 48/32/24/16 pilot symbols were used. The normalized mean square error (NMSE) was used as an evaluation standard to compare channel estimation results with the original 64 pilot symbols and with the redesigned 48/32/24/16 pilot symbols.

$$NMSE = 10 \lg \frac{\sum_{k=1}^K |h(\hat{k}) - h(k)|^2}{\sum_{k=1}^K |h(k)|^2}, \quad (20)$$

$h(k)$ is the real channel response estimated with original 64 pilot symbols. $h(\hat{k})$ is the channel response estimated with redesigned 48/32/24/16 pilot symbols. K is the number of sample data.

Figs. 6 and 7 show the CDF of NMSE with the redesigned 48/32/24/16 pilot symbols in the indoor and outdoor measurements. The goal of channel estimation with prior information extracted from the huge amount of existing CSI data was achieved. Fewer pilot symbols reduced the performance. Thus, there is a tradeoff between the accuracy of the channel estimation and the pilot cost. The outdoor propagation environment changed in a more complex manner than the indoor propagation environment. The proposed scheme performed better in the indoor propagation environment.

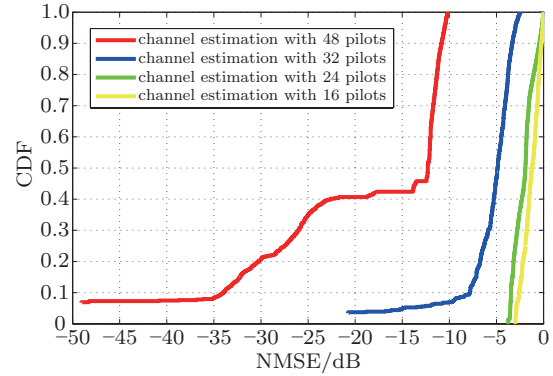


Figure 6 CDF of NMSE in the outdoor measurement

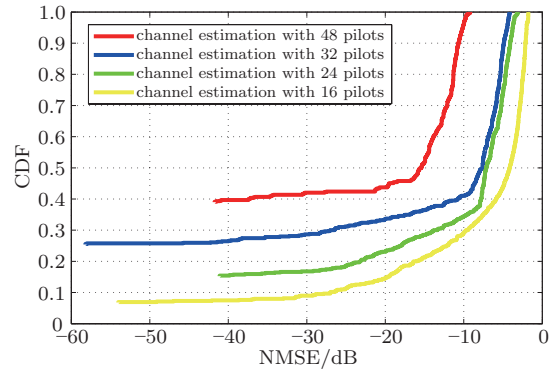


Figure 7 CDF of NMSE in the indoor measurement

The channel recognition performance via the CNN was astonishing. Datasets of three different wireless channels were input to the input layer: the outdoor channel (channel number: 100) with a carrier frequency of 900 MHz, the outdoor channel (channel number: 010) with a carrier frequency of 2.4 GHz, and the indoor channel (channel number: 001) with a carrier frequency of 900 MHz. The rest of the sys-

tem configuration was the same as that presented in Tab. 1. Tab. 2 presents the channel recognition results.

Table 2 Channel recognition result via CNN

	<i>hiddenlayers</i>	<i>width</i>	<i>epochs</i>	<i>alpha</i>	<i>error</i>
CNN1	6	6,12,24	1 000	0.1	0.039 2
CNN2	4	6,12	1 000	0.1	0.044 2
CNN3	4	6,12	100	1	0.078 3
CNN4	4	6,12	100	0.1	0.045 2
CNN5	4	6,12	100	0.001	0.259 0
CNN6	4	8,16	100	1	0.058 2

hiddenlayers represents the number of hidden layers in the CNN. As noted above, the hidden layers in a CNN come in pairs (*C* layer and *S* layer). Thus, *hiddenlayers* = 4 means that there are two pairs of convolution and sub-pooling layers. *width* means the number of connections in the convolution layer. *epochs* can be regarded as the training times. *alpha* is the learning rate with the range of (0, 1). The comparison between CNN1 and CNN2 showed that the deeper CNN performed better. The comparison between CNN2 and CNN4 shows that training a CNN more improves its performance. The comparison between CNN3 and CNN6 shows that the wider CNN performed better. CNN3, CNN4, and CNN5 showed that *alpha* cannot be too big or too small.

6 Conclusions and future work

This study focused on wireless channel feature extraction and recognition via a machine learning algorithm. A GMM was applied to acquire the distribution of wireless channel parameters, and a CNN was applied to distinguish different wireless channels automatically. The results based on indoor and outdoor channel measurement data showed that these machine learning algorithms can enjoy an astonishing level of performance in wireless communication. The contribution of this study is the application of a CNN to dealing with multi-path data. The numerical results help prove the concept of the wireless channel feature and the idea of the tomographic

channel model. The recognition of different wireless channels can help with applying prior information to guide the pilot design and save pilot resources.

Research and development of wireless big data is actively ongoing. Most wireless big data research is on the huge amount of service data, and few studies have focused on channel data. Most research has focused on the challenges of fast-growing big data in wireless networks, or “wireless for big data”. In this study, we attempted to apply machine learning to wireless communication, which we term “big data for wireless”.

In the future, our main focus will be carrying out more channel measurement campaigns in different typical propagation scenarios to build a sufficiently complete CSI database. More deep learning algorithms can be applied to deal with wireless communication problems with the database, such as the construction of the magnetic map.

References

- [1] G. E. Athanasiadou, A. R. Nix. A novel 3-D indoor ray-tracing propagation model: the path generator and evaluation of narrow-band and wide-band predictions [J]. IEEE transactions on vehicular technology , 2000, 49(4): 1152-1168.
- [2] J. Poutanen. Geometry-based radio channel modeling: propagation analysis and concept development [D]. Helsinki: Aalto University, 2011.
- [3] R. Verdone, A. Zanella. Pervasive mobile and ambient wireless communications: COST action 2100 [M]. Springer Science & Business Media, 2012.
- [4] J. Meinilä, P. Kyösti, T. Jämsä, et al. WINNER II channel models [C]//Radio Technologies and Concepts for IMT-Advanced, Wiley, 2009: 39-92.
- [5] J. Meinila, P. Kyosti, L. Hentila, et al. D5.3: WINNER+ final channel models. Wireless World Initiative New Radio WINNER, 2010.
- [6] S. Jaeckel, L. Raschkowski, K. Börner, et al. QuaDRiGa-Quasi deterministic radio channel generator, user manual and documentation [R]. Fraunhofer Heinrich Hertz Institute, 2014.
- [7] A. Maltsev, V. Erceg, E. Perahia, et al. Channel models for 60 GHz WLAN systems [EB/OL]. IEEE doc, 2010: 802.11-09.
- [8] J. Karedal, F. Tufvesson, N. Czink, et al. A geometry-based stochastic MIMO model for vehicle-to-vehicle

- communications [J]. *IEEE transactions on wireless communications*, 2009, 8(7): 3646-3657.
- [9] J. Medbo, K. Börner, K. Haneda, et al. Channel modelling for the fifth generation mobile communications [C]//The 8th European Conference on Antennas and Propagation (EuCAP), The Hague, Netherlands, 2014: 219-223.
- [10] C. X. Wang, A. Ghazal, B. Ai, et al. Channel measurements and models for high-speed train communication systems: a survey [J]. *IEEE communications surveys & tutorials*, 2016, 18(2): 974-987.
- [11] J. Z. Wang, H. L. Zhu, N. J. Gomes. Distributed antenna systems for mobile communications in high speed trains [J]. *IEEE journal on selected areas in communications*, 2012, 30(4): 675-683.
- [12] B. H. Chen, Z. D. Zhong, B. Ai, et al. Channel characteristics in high-speed railway: a survey of channel propagation properties [J]. *IEEE vehicular technology magazine*, 2015, 10(2): 67-78.
- [13] M. Metis. Wireless communications enablers for the Twentytwenty Information Society. EU 7th Framework Programme project [R]. ICT-317669-METIS.
- [14] M. Fallgren, B. Timus. Future radio access scenarios, requirements and KPIs [J]. METIS deliverable D, 2013, 1.
- [15] A. Osseiran, F. Boccardi, V. Braun, et al. Scenarios for 5G mobile and wireless communications: the vision of the METIS project [J]. *IEEE communications magazine*, 2014, 52(5): 26-35.
- [16] X. H. Wang, E. Anderson, P. Steenkiste, et al. Improving the accuracy of environment-specific channel modeling [J]. *IEEE transactions on mobile computing*, 2016, 15(4): 868-882.
- [17] X. H. Wang, E. Anderson, P. Steenkiste, et al. Improving the accuracy of environment-specific vehicular channel modeling [C]//The 7th ACM International Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization, Istanbul, Turkey, 2012: 43-50.
- [18] M. Boban, J. Barros, O. K. Tonguz. Geometry-based vehicle-to-vehicle channel modeling for large-scale simulation [J]. *IEEE transactions on vehicular technology*, 2014, 63(9): 4146-4164.
- [19] J. Medbo, H. Asplund, J. E. Berg, et al. Directional channel characteristics in elevation and azimuth at an urban macrocell base station [C]//The 6th European Conference on Antennas and Propagation (EuCAP), Prague, Czech Republic, 2012: 428-432.
- [20] H. H. Li, Y. Z. Li, S. D. Zhou, et al. A novel method to obtain CSI based on Gaussian mixture model and expectation maximization [C]//The 8th International Conference on Wireless Communications & Signal Processing (WCSP), Yangzhou, China, 2016: 1-5.
- [21] D. Tse, P. Viswanath. *Fundamentals of wireless communication* [M]. Cambridge: Cambridge University Press, 2005.
- [22] Y. Zhang, S. J. Liu, Y. H. Rui, et al. Wideband wireless channel predictor relying on environment information [C]//The 8th International Symposium on Antennas, Propagation and EM Theory (ISAPE), Kunming, China, 2008: 945-948.
- [23] Y. Zhang, S. J. Liu, Y. H. Rui, et al. Channel prediction assisted by radio propagation environments information [C]//The 4th IEEE International Conference on Circuits and Systems for Communications (ICCSC), Shanghai, China, 2008: 733-736.
- [24] L. Tong, B. M. Sadler, M. Dong. Pilot-assisted wireless transmissions: general model, design criteria, and signal processing [J]. *IEEE signal processing magazine*, 2004, 21(6): 12-25.
- [25] B. H. Fleury, P. Jourdan, A. Stucki. High-resolution channel parameter estimation for MIMO applications using the SAGE algorithm [C]//International Zurich Seminar on Broadband Communications Access-Transmission-Networking, Zurich, Switzerland, 2002: 30-30.
- [26] Z. Zivkovic. Improved adaptive Gaussian mixture model for background subtraction [C]//The 17th International Conference on Pattern Recognition (ICPR), Cambridge, UK, 2004, 2: 28-31.
- [27] T. Toda, A. W. Black, K. Tokuda. Statistical mapping between articulatory movements and acoustic spectrum using a Gaussian mixture model [J]. *Speech communication*, 2008, 50(3): 215-227.
- [28] T. K. Moon. The expectation-maximization algorithm [J]. *IEEE signal processing magazine*, 1996, 13(6): 47-60.
- [29] Y. Bengio. Learning deep architectures for AI [J]. *Foundations and trends in Machine Learning*, 2009, 2(1): 1-127.
- [30] J. Bouvrie. Notes on convolutional neural networks [EB/OL]. http://cogprints.org/5869/1/cnn_tutorial.pdf, 2006.
- [31] Y. LeCun, Y. Bengio. Convolutional networks for images, speech, and time series [C]//The handbook of brain theory and neural networks, MIT Press, Cambridge, USA, 1998: 255-258.

About the authors



Haihan Li received the B.S. and M.S. degree from Tsinghua University, Beijing China in 2011 and 2014 respectively. He is currently working toward the Ph.D. degree. His research now focuses on channel modeling and radio channel measurements, as well as the data mining of statistical CSI. (Email:

lih14@mails.tsinghua.edu.cn)



Yunzhou Li [corresponding author] received his Ph.D. degree from Tsinghua University in 2004. currently, he is an associate professor of Tsinghua University. Dr. Li mainly focus on signal processing technologies in wireless and mobile communications, including spatial-time signal processing, channel estimation, multi-user detection, and synchronization algorithms for CDMA/OFDM system. Also he is interested in analysis, optimization design and enhancement of cellular system and WLAN. (Email: liyunzhou@mail.tsinghua.edu.cn)



Shidong Zhou received the B.S. and M.S. degrees in wireless communications from Southeast University, Nanjing, China, in 1991 and 1994, respectively, and the Ph.D. degree in communication and information systems from Tsinghua University, Beijing, China, in 1998. From 1999 to 2001, he was in charge of several projects in the China 3G Mobile Communication R&D Project. He is currently a professor at Tsinghua University. He is currently a member of the China Fu-TURE Project. His research interests are in the area of wireless and mobile communications.



Jing Wang received the B.S. and M.S. degree in electronic engineering from Tsinghua University, Beijing China in 1983 and 1986 respectively. He has been on the faculty at the Tsinghua University since 1986. He is currently a professor of the School of Information Science and Technology. He serves as the vice director of Tsinghua National Lab for Information Science and Technology. His research interests are in the area of wireless communications, including transmission and networking technologies of 4G/B4G. He has published more than 150 conference and journal papers.