CrossMark

Fading channel modelling using single-hidden layer feedforward neural networks

Junbiao Liu[1] · **Xinyu Jin[1]** · **Fang Dong[2]** ·
Liang He[3] · **Hong Liu[2]**

Abstract Due to the complexity and extensive application of wireless systems, fading channel modeling is of great importance for designing a mobile network, especially for high speed environments. High mobility challenges the speed of channel estimation and model optimization. In this study, we propose a single-hidden layer feedforward neural network (SLFN) approach to modelling fading channels, including large-scale attenuation and small-scale variation. The arrangements of SLFN in path loss (PL) prediction and fading channel estimation are provided, and the information in both of them is trained with extreme learning machine (ELM) algorithm and a faster back-propagation (BP) algorithm called Levenberg-Marquardt algorithm. Computer simulations show that our proposed SLFN estimators could obtain PL prediction and the instantaneous channel transfer function of sufficient accuracy. Furthermore, compared with BP algorithm, the ability of ELM to provide millisecond-level learning makes it very suitable for fading channel modelling in high speed scenarios.

✉ Junbiao Liu
liujunbiao@zju.edu.cn

Xinyu Jin
jinxy@zju.edu.cn

Fang Dong
dongf@zucc.edu.cn

Liang He
heliang@mail.tsinghua.edu.cn

Hong Liu
liuhong@zucc.edu.cn

[1] Institution of Information Science and Electrical Engineering, Zhejiang University,
Hangzhou 310037, Zhejiang, China

[2] School of Information and Electrical Engineering, Zhejiang University City College,
Hangzhou 310015, Zhejiang, China

[3] Department of Electronic Engineer, Tsinghua University, Beijing 100084, Beijing, China

1 Introduction

High-speed railways (HSR) and highway networks have developed rapidly for nearly ten years to meet people's travel needs. This explosive increase of high speed transportation raises higher requirements for wireless communication systems, including train ground communication (TGC) system (Shu et al. 2014), communication based train control (CBTC) system, vehicle ad hoc network (VAN) (Golestan et al. 2015), vehicle to vehicle (V2V) communication and so forth.

However, the speed of trains can reach 350 km/h and the speed of vehicles is up to 120 km/h, which make the users can not enjoy the smooth and high quality wireless services under low speed environment. In high mobility scenarios, large Doppler frequency shift, fast fading channel and fast handover issue seriously affect communication performances (Zhou and Ai 2014; Calle-Sanchez et al. 2013).

Wireless channel plays a key background role in transmission rate and the quality of mobile propagation. The modelling of radio wave propagation is essential for the planning of physical layer operations, such as the best modulation and coding interleaving scheme, equalizer design, antenna configuration and subcarrier allocation etc. Taking the continuous movements of the users into account, the profile of fading channel changes in real time and the parameters of the propagation process become random variables (Sotiroudis and Siakavara 2015). Fading channel is modelled in two ways, large-scale attenuation known as path loss (PL) prediction and small-scale variation known as channel transfer function estimation (Rappaport 2001). The former predicts the mean signal strength for an arbitrary large transmitter-receiver distance (several hundreds or thousands of meters; Phillips et al. 2013) while the latter characterizes the rapid fluctuations of received signal strength over very short distances (a few wavelengths) or short durations (on the order of seconds; Chang and Wu 2014).

PL prediction proposals in the literature can broadly be classified as (a) empirical and (b) measurement models. The models of type (a) (Friis 1946; Seidel and Rappaport 1991; ITU-R 2002; Herring et al. 2012) make use of available prior knowledge collected from a given environment, so that it is suitable in the situations impossible or difficult to obtain measurements. These empirical models are widely adopted into network simulators or served to derive a simplified version of more complex models since they are easy to implement, although they are questionably accurate. The models of type (b) (Lee 1985) define a sampling method and a means of PL predicting where PL could not be measured, based on the assumption that the burden of measurements and prediction is unavoidable and acceptable. The typical methods as explicit mapping (Hills 2001), partition-based models (Durgin 2009), iterative heuristic refinement (Robinson et al. 2008) and active learning and geostatistics (Marchant and Lark 2007) choose the measurements set to improve the accuracy of modelling.

Fading channel estimation is crucial for receiver design in coherent communication systems. Channel estimation models can be grouped into three categories: (a) estimation based on channel frequency response (Beek et al. 1995), (b) estimation with parametric model (Yang et al. 2001; Qing and Gang 2014) and (c) iterative channel estimation (Valenti and Woerner 2001; Kim and Hansen 2015). Approach (a) is a basic method with the minimal model complexity. It estimates the channel frequency response of pilot symbols first, and then

obtains the responses of data symbols via interpolation. Approach (b) and (c) can provide better performance. For approach (b), a set of parameters are estimated in the reconstruction of channel response, which is suitable for sparse channels with 2–6 dominant paths (Chen et al. 1999). Approach (c) takes advantage of error-control coding techniques such as turbo code (Li et al. 2010) and LDPC code (Fang 2010) to estimate channel in an iterative manner.

Both PL prediction and fading channel estimation have two ways to model channels: blind and pilot channel modelling (Vieira et al. 2013; Larsen et al. 2009). Pilot model (Pun et al. 2011) is typically constructed by using pilot symbols strategically placed at the frame head or subcarrier. In blind modelling (Mielczarek and Svensson 2005), channel coefficients are predicted by using statistical features of received signals. Once a channel is estimated its time-frequency characteristics, relevant parameters are used to update the pre-set model. As in any estimation application, wireless channel estimation aims to quantify the best performance of wireless systems. However, due to the unlimited number of received signal, it is a challenge to extract optimal channel coefficients.

Feedforward neural networks (FNN) is extensively used to provide models for a natural or artificial phenomena that are difficult to handle using classical parametric techniques (Cao et al. 2012). A single-hidden layer FNN (SLFN) is found to be more computationally efficient for the work than 2- or 3-hidden layered forms (Sarma and Mitra 2013). Simsir and Taspinar (2015) and Seyman and Taspinar (2013) demonstrated that channel estimation based on neural network ensures better performance than conventional Least Squares (LS) algorithm without any requirements for channel statistics and noise. In the meantime, Sotiroudis and Siakavara (2015) and Cheng et al. (2015) also proved SLFNs can be used in channel estimation for various wireless environments. Most of the literatures (Sotiroudis and Siakavara 2015; Seyman and Taspinar 2013; Panda et al. 2015; Taspinar and Cicek 2013) proved that SLFNs are effective for static and slowly varying channels. Unfortunately, the learning speed of SLFNs has been a major bottleneck in many applications, and fast fading channel caused by high mobility makes this method unsuitable for channel modelling too. Unlike traditional SLFN implementations, a simple learning algorithm called extreme learning machine (ELM) with good generalization performance (Cao et al. 2012; Cao and Lin 2015; Cao and Lin 2014) can learn thousands of times faster, which randomly chooses hidden nodes and analytically determined the output weights of SLFNs. The extremely fast learning speed of ELM makes the algorithm viable in fading channel modelling.

In this paper, we propose SLFN approaches based on ELM to modelling fading channel. Since the modelling has concentrate on large-scale attenuation and small-scale variation (Rappaport 2001), we choose path loss coefficient and channel transfer function to estimate. The performance of ELM estimator with SLFN architecture is compared to that of BP estimator. We use training and testing time, estimation error, bit error rate (BER) as the performance criteria.

The outline of the paper is as follows: Sect. 2 describes the basic considerations related to fading channel modelling using SLFNs. Sect. 3 proposes a PL prediction method based on ELM for large-scale attenuation and analyzes the simulation results. Sect. 4 develops an ELM estimator for channel transfer function and derives the results from the simulations. Conclusion is given in Sect. 5.

The performance of fading channel modelling based on ELM is in comparison with back-propagation (BP) algorithm (Tariq 1991) (Levenberg-Marquardt algorithm; Bogdan et al. 2010) which is a popular algorithm of SLFNs. All of the simulations are carried out in MATLAB 7.12.0. Levenberg-Marquardt algorithm is provided by MATLAB package, while ELM algorithm is from (Cao et al. 2015).
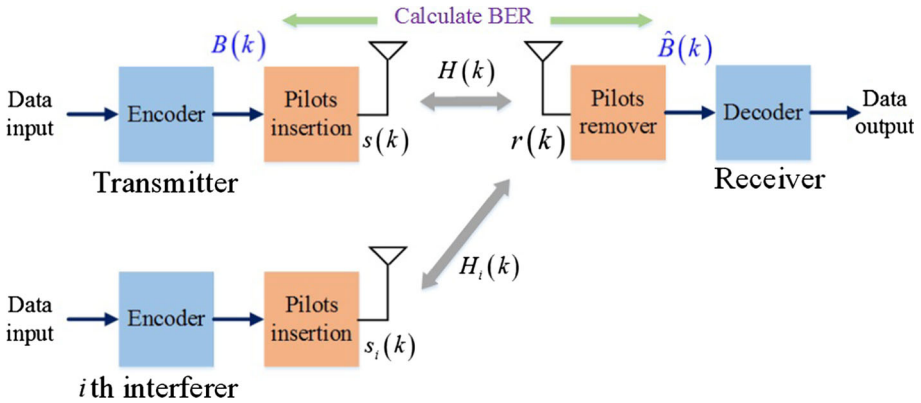
Fig. 1 Fading channel model with transmitter and receiver's block diagrams

2 Fading channel and its modelling using SLFNs

In this section, fading channel model and its radio propagation expressions are developed. The SLFN architecture of ELM and BP algorithms are discussed in the arrangement for PL prediction and fading channel estimation.

2.1 Radio propagation

A linear fading channel in Fig. 1 is with output r and input s, where r can be expressed as $r = f(s)$. $f(\cdot)$ denotes a transform function at the receiver end, representing the propagation related variations due to fading, shadowing or phase shifting.

The input-output relationship is given by

$$r(k) = H(k)s(k) + w(k) \tag{1}$$

where $s(k)$ is the transmit symbol at kth time slot, $r(k)$ is the receive signal, $w(k)$ is an additive noise and $H(k)$ denotes channel transfer function, respectively. Considering the interference of co-channel, the relationship in (1) is written as

$$r(k) = H(k)s(k) + \sum_{i=1}^{L} H_i(k)s_i(k) + w(k) \tag{2}$$

where L is the number of interferers, $H_i(k)$ is channel transfer function of the ith interferer.

If the transmitter has transmit power P_t dBm and antenna gain G_t dBi, the total effective isotropic radiated power (EIRP) in the log domain is $P_t + G_t$. The entire channel in Eq. (1) in the form of power is expressed as

$$P_r = P_t + G_t + G_r - PL(k) - N_0 \tag{3}$$

where P_r and G_r are receive power and antenna gain. PL encompasses all attenuation due to path loss and N is the power of thermal noise. Therefore, the signal to noise ratio (SNR) in log domain is $SNR = P_t - N_0$. Considering interference in Fig. 1, the signal to interference and noise ratio (SINR) is defined as

$$SINR = P_{\mathrm{t}} - \left(N_0 + \sum_{i=1}^{L} I_i \right) \tag{4}$$

where I_i is the receive power from ith interferer.

Commonly, fading channel modelling is aim to estimate the value of PL

$$PL(k) = L_{\mathrm{p}}(k) + L_{\mathrm{s}}(k) + L_{\mathrm{f}}(k) \tag{5}$$

where L_{p} is free-space path loss, L_{s} is the loss due to shadowing (slow fading) from large obstacles. $L_{\mathrm{p}}(k) + L_{\mathrm{s}}(k)$ is the large-scale fading whereas L_{f} is the small-scale fading.

2.2 SLFN architecture and learning algorithms

The idea to employ SLFN algorithm for the solution of fading channel modelling is based on its learning ability to find out the relationship between channel parameters and the receive signals, even when the channel is complicated. SLFN basically consists of input, hidden and output layers where neurons connected to each other by parallel synaptic weights.

For N independent training samples $(\mathbf{x}_i, \mathbf{t}_i)$ where $\mathbf{x}_i \in \mathbf{R}^n, \mathbf{t}_i \in \mathbf{R}^m, i = 1, 2, \cdots, N$, SLFNs with $\tilde{N}$ hidden nodes are modelled as

$$\sum_{j=1}^{\tilde{N}} \beta_j g_j(x_i) = \sum_{j=1}^{\tilde{N}} \beta_j g\left(\mathbf{w}_j \cdot \mathbf{x}_i + \mathbf{b}_j \right) = \mathbf{o}_i \tag{6}$$

where $\mathbf{w}_j$ is the weight vector between the input node and jth hidden node , β_j is the weight vector between jth hidden node and the output nodes, and $\mathbf{b}_j$ is the threshold of jth hidden nodes. The operational symbol '·' represents inner product.

The objective function of SLFNs E that will be minimized is written as

$$E\left(\mathbf{w}_j, \beta_j, \mathbf{b}_j \right) = \sum_{i=1}^{N} \| \mathbf{o}_i - \mathbf{t}_i \| \tag{7}$$

where vector $\mathbf{w}$ is the set of weight $\left(\mathbf{w}_j, \beta_j \right)$ and biases $\left(\mathbf{b}_j \right)$ parameters . $\mathbf{o}_i$ is ith actual output and $\mathbf{t}_i$ is ith desired output.

2.2.1 Levenberg-Marquardt algorithm

BP algorithm iteratively adjusts all of its parameters to minimize the objective function by using gradient-based algorithms (Bogdan et al. 2010). In the minimization procedure, vector $\mathbf{w}$ is iteratively adjusted as follows:

$$\mathbf{w}_{l+1} = \mathbf{w}_l - \eta \frac{\partial E(\mathbf{w})}{\partial \mathbf{w}} \tag{8}$$

where μ is the learning speed.

The training steps of Levenberg-Marquardt algorithm are illustrated as:

(1) Initialize the weight vector $\mathbf{w}$ and learning speed μ.
(2) Compute the objective function in terms of mean square error (MSE), i.e. $E = \frac{1}{N} \sum_{i=1}^{N} |\mathbf{o}_i - \mathbf{t}_i|^2$.
(3) Update weight vectors with formula $\Delta\mathbf{w} = -\left[\mathbf{J}^{\mathrm{T}}\mathbf{J} + \mu\mathbf{I} \right]^{-1} \mathbf{J}^{\mathrm{T}}E$ where $\mathbf{J}$ is the Jacobian matrix.

Fig. 2 The application of SLFN
for fading channel modelling
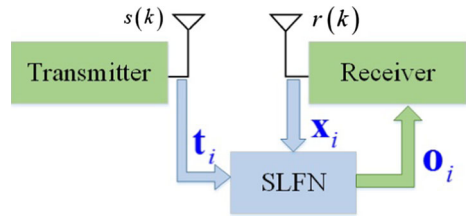
(4) Recompute the objective function with $\mathbf{w} + \Delta\mathbf{w}$. α is a constant that regulates μ and $\alpha < 1$. If error is smaller than the MSE value in step 2, reduce μ α times and go to step 3. If error is not reduced, increase μ α^{-1} times and go to step 3.

(5) Finish the training when error is less than its preset value or iteration epoches reach the maximum.

2.2.2 ELM algorithm

Although BP's gradient can be computed efficiently, an inappropriate learning speed might raise several issues, such as slow convergence, divergence, or stopping at a local minima (Cao et al. 2015). ELM algorithm randomly assigns the hidden node learning parameters and analytically determines the network output weights without learning iterations. Therefore, this method is able to make learning extremely fast.

ELM algorithm steps are as follow:

(1) Assign a training set $\aleph = \{(\mathbf{x}_i, \mathbf{t}_i) | i = 1, 2, \cdots, N\}$, active function $g(x)$ and the number of hidden neurons $\tilde{N}$,

(2) Randomly assign input weight vector $\mathbf{w}_j$ and bias value b_j, $j = 1, 2, \cdots, \tilde{N}$,

(3) Calculate the hidden layer output matrix $\mathbf{H}$ and its $Moore-Penrose$ generalized inverse matrix $\mathbf{H}^\dagger$,

(4) Calculate the output weight $\hat{\beta} = \mathbf{H}^\dagger \mathbf{T}$ with the least squares, where $\mathbf{T} = [\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_N]^{\mathrm{T}}$.

For a linear system $\mathbf{H}\beta = \mathbf{T}$, ELM algorithm finds a least-squares solution $\hat{\beta}$ rather than iterative adjustment (Cao and Xiong 2014). Seen from the steps, the learning time of ELM is mainly spent on calculating $\mathbf{H}^\dagger$. Therefore, ELM saves a lot of time in most applications. The performance evaluation in Cao et al. (2012); Cao and Lin (2015) shows that ELM can produce good generalization performance in most cases and can learn more than hundreds of times faster than BP.

2.3 Configuring SLFN for fading channel modelling

The application of SLFN for fading channel modelling is shown in Fig. 2. In our work, by adjusting the combination of training data, the corresponding training set is applied to the SLFN estimator with learning ability. The learning procedure is carried out with no parametric dependence.

Figure 3 shows the arrangement of SLFN to perform PL prediction in subgraph (a) and (b) and fading channel estimation in subgraph (c) and (d). For PL prediction, the SLFN estimator is trained by receive signal $r(k)$, $k = 1, 2, \cdots, N$ or train data set $(r(k), \tilde{\gamma}(k))$ where $\tilde{\gamma}(k)$ is the reference PL exponent obtained by pilot symbols. The output of SLFN estimator is $\tilde{\gamma}(k)$ at kth time slot. Similarly, the SLFN estimator in fading channel estimation is trained by $r(k)$ or $\left(r(k), \tilde{H}(k)\right)$ and its output is channel transfer function $\hat{H}(k)$.

$$\tilde{\gamma}(1) \quad \tilde{\gamma}(N)$$

$r(1) \rightarrow$ SLFN $\rightarrow \hat{\gamma}(1)$ $r(1) \rightarrow$ SLFN $\rightarrow \hat{\gamma}(1)$

$r(N) \rightarrow$ $\rightarrow \hat{\gamma}(N)$ $r(N) \rightarrow$ $\rightarrow \hat{\gamma}(N)$

(a)   **(b)**

$$\tilde{H}(1) \quad \tilde{H}(N)$$

$r(1) \rightarrow$ SLFN $\rightarrow \hat{H}(1)$ $r(1) \rightarrow$ SLFN $\rightarrow \hat{H}(1)$

$r(N) \rightarrow$ $\rightarrow \hat{H}(N)$ $r(N) \rightarrow$ $\rightarrow \hat{H}(N)$
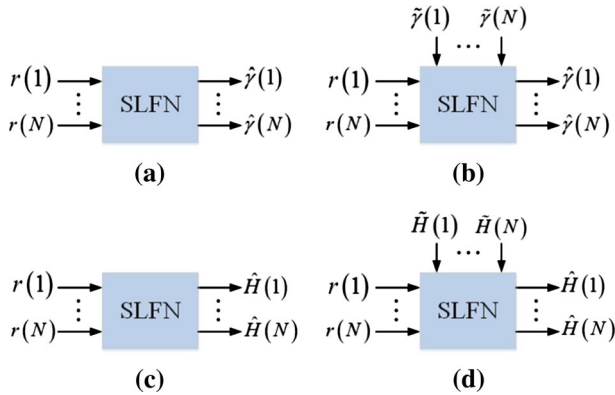
(c)   **(d)**

Fig. 3 Arrangement of SLFN in PL prediction and fading channel estimation

Table 1 Parameters used in fading channel modelling

Parameters	Values
Carrier frequency	2.35 GHz
Transmission rate	1 Mbps
Oversampling factor	4
Modulation type	BPSK
Velocity of terminals	300 km/h (83.3 m/s)
	120 km/h (33.3 m/s)
Doppler category	Jakes
Transmit power	39.5 dBm
Transmitter-receiver distance	500 m
Generation of pilot symbols	Hadamard matrix
Frame length	100
Maximum number of errors	500
Maximum number of packets	3000
Number of neurons	10, 20

The performance of our proposed SLFN estimator for fading channels will be demonstrated by computer simulations. The performance of our proposed estimator is evaluated using MSE or BER versus SNR criterion. In addition, the mobility of terminals greatly limits the estimation speed, so we also consider the consuming time of the estimators. The parameters of fading channel we used in the simulations are shown in Table 1.

3 PL prediction modelling

In this section, the structure of ELM estimator in the prediction of PL exponent is proposed and computer simulations show its performance compared with BP estimator.

3.1 Model analysis

PL exponent γ indicates the rate at which PL increases with distance. The smaller γ is, the less energy loss of wireless signal due to transceiver-receiver distance is. Typical value of γ is determined by carrier frequency and propagation terrain. For example, in HSR environment, γ is slightly larger than 2 in rural area (within 250–3200 m) with narrow band communication system while it is near to 4 in hilly terrain (within 800–2500 m) with broadband system.

In Friis (1946), a basic formula for free-space transmission loss with ideal isotropic antennas is simplified to

$$\frac{P_t}{P_r} = \left(\frac{\lambda}{\pi d}\right)^2 \tag{9}$$

where λ is the carrier's wavelength and d is the transmitter-receiver distance in meters. Friis assumed that $\gamma = 2$ so that signals degraded as a function of d^2 (Friis 1946). However, a common extension to non-line-of-sight (NLOS) environments is to use a larger γ. Both theoretical and measurement-based channel models (such as Egli model, two-ray model, Okumura model, Hata model etc) indicate that γ is in the range 1–4 (Rappaport 2001). Considering shadowing effects component ψ obeys a log-normal distribution, the statistical PL model (Erceg et al. 1999; Kastell 2011) is given by

$$P_r = P_t + 20 \log_{10} \frac{\lambda}{4\pi} - 10\gamma \log_{10} d - \psi \tag{10}$$

where ψ is a zero-mean Gaussian distributed random variable with standard deviation σ_ψ (also in dBm).

PL exponent γ could be obtained by fitting the minimum mean square error (MMSE) of measurements (Liu et al. 2012)

$$\gamma^* = F_{\text{MMSE}}(\gamma) = \min_\gamma \sum_{i=1}^{N} [M_{\text{measured}}(d_i) - M_{\text{model}}(d_i)]^2 \tag{11}$$

where $M = P_t/P_r$, $M_{\text{measured}}(d_i)$ is the measured values at distance d_i, and $M_{\text{model}}(d_i) = 20 \log_{10}(\lambda/(4\pi)) - 10\gamma \log_{10} d_i$. And the variance σ_ψ^2 is given by

$$\sigma_\psi^2 = \frac{1}{N} \sum_{i=1}^{N} [M_{\text{measured}}(d_i) - M_{\text{model}}(d_i)]^2 \Big|_{\gamma^*} \tag{12}$$

If the radius of the transmitter's cell is 1 km and the vehicle's velocity is 120 km/h, the receiver will carry out a handover procedure every 60 s, so that γ needs to be calculated at the receiver at least once per minute. if the velocity is up to 350 km/h, γ needs to be calculated every 20.6 s. According to Eqs. (11) and (12), the calculation of γ requires hundreds or thousands of receive signal measurements, the introduction of learning algorithm into PL prediction would be effective in simplifying data processing.

3.2 PL prediction based on ELM

As Fig. 3a, our proposed ELM estimators's structure at receiver end is shown in Fig. 4 and it periodically adjust PL prediction according to the latest received frame to trace the variation of large-scale propagation.

After the receiver receiving a frame, the procedure of ELM estimator is stated as follows:

(1) Initialize the standard deviation $\hat{\sigma}_\psi^2(0)$.

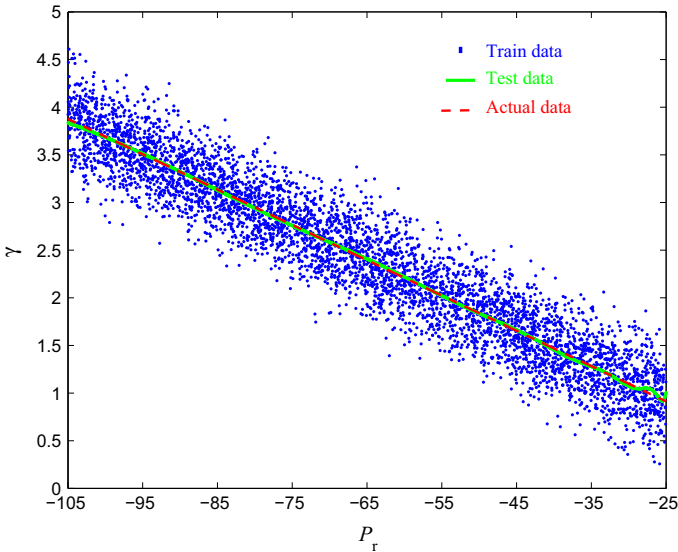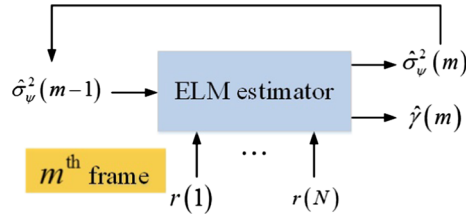**Fig. 4** The structure of ELM estimator in PL prediction

Fig. 5 The prediction of PL exponent γ by ELM learning algorithm

(2) After receiving the mth frame, assign a testing set $\{(P_{ri}, \gamma_i) \mid i = 1, 2, \cdots, N\}$ when N is the frame length.

(3) Run ELM algorithm and obtain the prediction $\hat{\gamma}(m) = \overline{\hat{\gamma}_i}$ and $\hat{\sigma}_{\psi}^2(m)$.

(4) Update the value of the standard deviation for the next prediction.

Obviously, the whole procedure doesnt have any parameters to be tuned iteratively.

3.3 Simulation results

We use ELM and BP algorithms to predict PL exponent γ. The velocity is set at 120 km/h. A training set (P_{ri}, γ_i) and testing set (P_{ri}, γ_i) with 1000 data, respectively are created where P_{ri} is uniformly randomly distributed on the interval $(-105, -25)$ dBm (Liu et al. 2012). Shadowing effect exponent ψ has been added to all training samples while testing data are shadowing-free.

The number of hidden neurons of ELM is initially set at 20 and active function is sigmoidal. Simulation result is shown in Fig. 5. The train accuracy measured in terms of root mean square error (RMSE) is 0.27734 due to the shadowing effect, whereas the test accuracy is 0.012445. Figure 5 confirms that the prediction results of γ are accurate, and there is a visible margin of error only when $P_r > -30$ dBm.

Table 2 Performance comparison for learning algorithms in PL prediction

Algorithms	Time (s)		Training		Testing		Hidden neurons
	Training	Testing	RMS	Dev	RMS	Dev	
ELM	0.0025	0.0008	0.5971	0.0318	0.4955	0.1615	5
BP	30.6555	0.0599	0.0759	0.0027	0.0014	0.0008	5
ELM	0.0040	0.0014	0.3347	0.0238	0.1486	0.0884	10
BP	39.8888	0.0658	0.0748	0.0028	0.0023	0.0010	10
ELM	0.0068	0.0029	0.2828	0.0219	0.0475	0.0500	20
BP	53.6168	0.0671	0.0745	0.0028	0.0031	0.0012	20

Average 200 trails of simulation have been conducted for both ELM and BP algorithm, whose results are shown in Table 2. In BP, the input and output layers have log-sigmoidal and the hidden layer have tan-sigmoidal activation functions. When the number of hidden neurons is 20, ELM estimator spends 6.8 ms CPU time on training and 2.9 ms on testing. However, it takes 53.6 s for BP estimator on training and 67.1 ms on testing, which makes it unsuitable for high speed scenarios since the consuming time is larger than 20.6 s (for 350 km/s) or slightly lower than 60 s (for 120 km/h). ELM runs 7800 times faster than BP. Compared with 60 s and 20.6 s in Sect. 3.1, BP is too time-consuming to be used in PL prediction in a scenario with high mobility. Although ELM has a much higher testing error 0.0475 compared with 0.0031 in BP, this prediction error can be acceptable as shown in Fig. 5.

In addition, considering network transmission rate is 1 Mbps, the reception of 1000 test data takes only 1 ms, so that ELM estimator can predict PL exponent γ with RMSE less than 0.05 within a time interval less than 8 ms. If the velocity is 120 km/h, the vehicle has only moved 0.27 m during PL prediction interval. If the velocity is 350 km/h, the moving distance is less than 0.78 m. For large-scale fading, γ and σ_ψ^2 barely have a sharp change in the range, so that it is reasonable to use σ_ψ^2 in the next prediction in Fig. 4.

Figure 6 shows the relationship between the generalization performance of ELM and the number of hidden neurons $\tilde{N}$ in PL prediction. Every $\tilde{N}$ simulates 50 times. Obviously, the generalization performance of ELM is stable when $\tilde{N} \geq 12$. Thus, the simulation result in Fig. 5 is reasonable when $\tilde{N}$ is set at 20.

Figure 7a shows the relationship between RMSE of ELM and the number of train/test data N when $\tilde{N} = 20$, and Fig. 7b shows the impact of N on consuming time. Training RMSE is almost a constant (slightly less than 0.3) because ELM use Moore-Penrose inverse matrix calculation to solve the problem of finding the smallest norm least-squares output weight. Unlike it, testing RMSE decreases with an increase in the number of test data. The simulation confirms the conclusion in Cao et al. (2012) that ELM has no over-trained phenomenon. Both train and test consuming time increases with N, however, the increase of test time is less than that of train time. It should also be noted that, even N is up to 10^4, the time consumed by ELM estimator is still acceptable, which is less than 70 ms.

4 Fading channel estimation modelling

The channel transfer function $H(k)$ of fading channels in Eq. (1) is assumed to maintain a constant in a frame (Sarma and Mitra 2013), i.e. $H(k) = H$. A frame's input-output relationship is
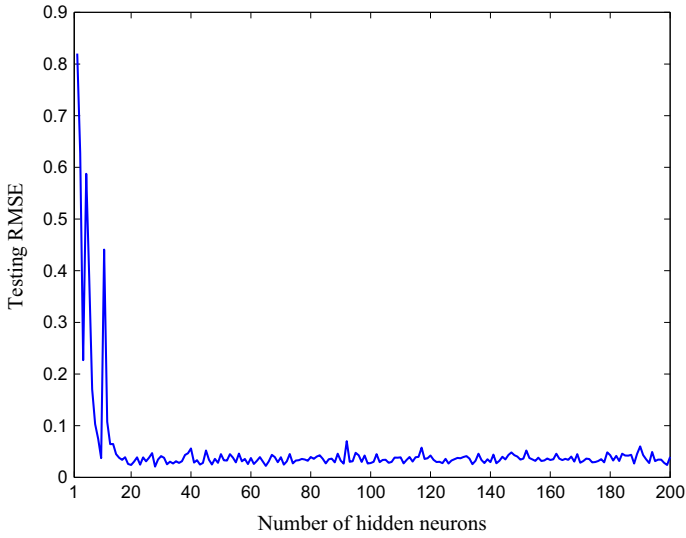
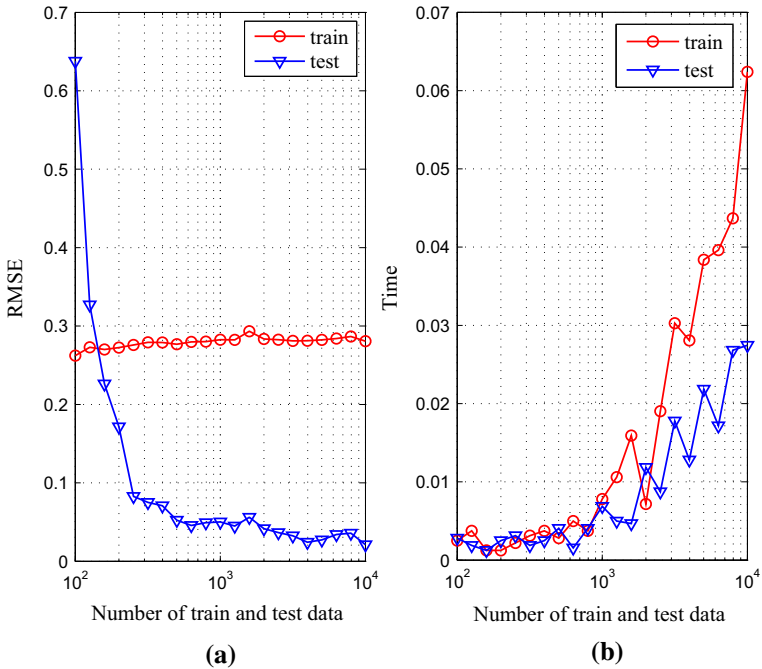**Fig. 6** The generalization performance of ELM in PL prediction

Fig. 7 Number of train/test data of ELM in PL prediction

$$r(k) = Hs(k) + w(k), k = 1, 2, \dots, N \tag{13}$$

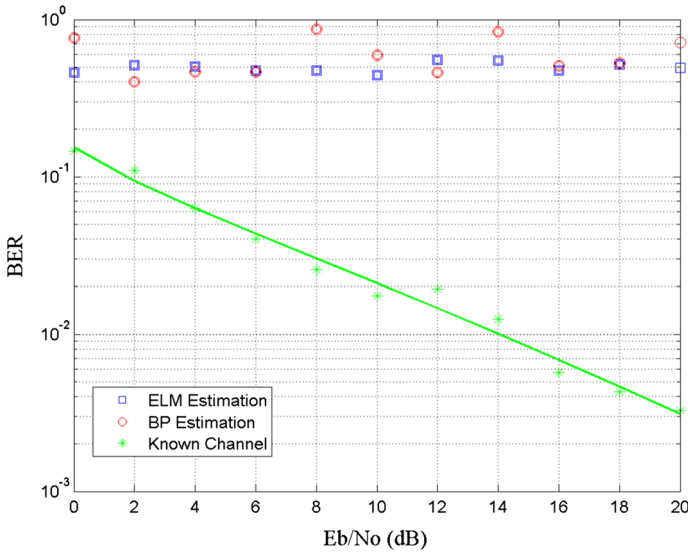The actual H and $w(k)$ is unknown to the receiver.

Fig. 8 The BER of Rayleigh fading channel with $\hat{H}$ estimated by ELM and BP algorithm

The Rayleigh multi-path fading channel assumes that the delay power profile and the Doppler spectrum of channel are separate. The channel with normalized power is given by

$$H(k) = \frac{1}{\sqrt{2}} \left[n_c(k) + j n_s(k) \right] \tag{14}$$

where n_c and n_s are the phase and quadrature branches of the baseband signal in Gauss process.

Since BPSK modulation is employed in the model in Fig. 1, the transmit signal $s(k)$ is real. Rayleigh channel H is a complex variable so that the receive signal $r(k)$ is complex too. Considering the data sets in ELM and BP algorithms $\mathbf{x}_i \in \mathbf{R}^n$, $\mathbf{t}_i \in \mathbf{R}^m$, the estimation of H should be divided into real and imaginary parts.

If we set up the data sets in SLFN as $\aleph 1 = \{(r(k), \text{Re}(H)) | k = 1, 2, \cdots, N\}$ and $\aleph 2 = \{(r(k), \text{Im}(H)) | k = 1, 2, \cdots, N\}$. As shown in Fig. 3c, the estimated channel transfer function is

$$\hat{H} = \frac{1}{N} \sum_{k=1}^{N} \hat{H}(k) = \frac{1}{N} \sum_{k=1}^{N} \left[\text{Re}\left(\hat{H}(k)\right) + \text{Im}\left(\hat{H}(k)\right) \right] \tag{15}$$

Figure 8 shows the ability of SLFN to capture the time-varying characteristic of fading channel. When Rayleigh pattern is applied, SLFN shows an inability to deal directly with the time-varying nature. The BER performance of ELM estimation is around 0.45, which is similar to random decision, whereas the BER performance of BP estimation is even worse. This is due to the receive signal $r(k)$ affected by three Gaussian random variables, $n_c(k)$, $n_s(k)$ and $w(k)$. Thus, we draw a conclusion that blind modelling is not suit for fading channel estimation base on SLFN architectures.

Fig. 9 The structure of ELM estimator in fading channel estimation
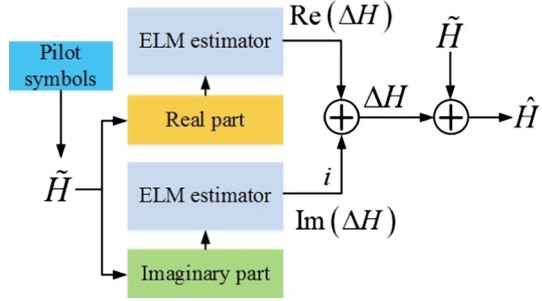
In order to narrow the estimation down, pilot signals are used in fading channel estimation to provide reference channel transfer function $\tilde{H}$, and it can be calculated by

$$\tilde{H} = \frac{1}{N_p} \sum_{k=-N_p}^{-1} r(k) s(k) \tag{16}$$

4.1 Fading channel estimation based on ELM

Our proposed ELM estimator's structure at receiver end is shown in Fig. 9. As Fig. 3, the transmitter in fading channel inserts pilot symbols in frames' headers to trace the time-varying characteristic of channel.

In Fig. 9, ELM estimator of channel transfer function H is divided into real and imaginary parts. The reference channel transfer function $\tilde{H}$ is used to estimate the error of channel transfer function ΔH, i.e. the data sets are $\aleph 1 = \left(\mathrm{Re}\left(\tilde{H}\right), \mathrm{Re}\left(\Delta H\right) \right)$ and $\aleph 2 = \left(\mathrm{Im}\left(\tilde{H}\right), \mathrm{Im}\left(\Delta H\right) \right)$. Thus, the estimation value is given by

$$\hat{H} = \tilde{H} + \Delta H \tag{17}$$

After the receiver receiving a frame, the procedure of ELM estimator is stated as follows:

(1) Calculate the reference channel transfer function $\tilde{H}$ in Eq. (16).
(2) Assign real testing set $\left\{ \left(\mathrm{Re}\left(\tilde{H}\right), \mathrm{Re}\left(\Delta H_i\right) \right) \middle| i = 1, 2, \cdots, N \right\}$ and imaginary testing set $\left\{ \left(\mathrm{Im}\left(\tilde{H}\right), \mathrm{Im}\left(\Delta H_i\right) \right) \middle| i = 1, 2, \cdots, N \right\}$.
(3) Run ELM algorithm and obtain the estimation of channel transfer function $\hat{H} = \tilde{H} + \Delta \overline{H_i}$.
(4) Count BER and decide whether or not to adjust the number of pilot symbols.

4.2 Simulation results

We use ELM and BP algorithms to estimate channel transfer function H. The number of hidden neurons of SLFN is set at 10 and average 3000 packets have been conducted for two algorithm for comparison. The simulation results are shown in Table 3. ELM estimator spends less than 8 ms on training and 6 ms on testing. In contrast, BP estimator needs at least 42 s to complete the training process and 30 ms to perform the testing procedure. For the same reason in PL prediction, BP algorithm also could not be applied to fading channel estimation for high speed scenarios. In addition, both SNR and the number of pilot symbols determine estimation accuracy. ELM shows a higher accuracy than BP when SNR is 0 dBm.

Table 3 Performance comparison for learning algorithms in fading channel estimation

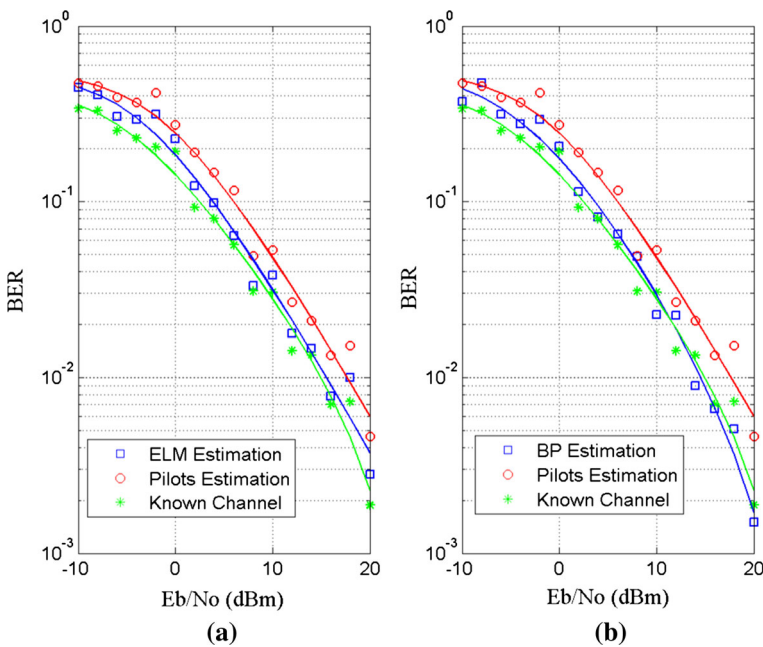| Algorithms	Training		Testing		Pilot symbols	SNR
	Time (s)	RMS	Time (s)	RMS		
ELM	0.0078	0.0508	0.0043	0.0615	1	0
BP	43.4310	0.1272	0.0396	0.1192	1	0
ELM	0.0044	0.0074	0.0034	0.0080	1	20
BP	43.8050	0.0514	0.0372	0.0008	1	20
ELM	0.0054	0.0506	0.0061	0.0429	2	0
BP	42.6190	0.1294	0.0312	0.0807	2	0
ELM	0.0070	0.0073	0.0039	0.0055	2	20
BP	58.5160	0.0485	0.0369	0.0006	2	20

Fig. 10 BER performance of estimators versus SNR with 1 pilot symbol

The result is quite opposite when SNR is 20 dBm. It implies that ELM estimator is more suitable for fading modelling with deteriorated channel condition.

Figures 10 and 11 show the BER performance of estimators versus SNR, which is counted 3000 BPSK frames at every SNR value. Both figures indicate that the performance of ELM with SLFN architecture in Fig. 9 is as good as that of BP estimator, although their RMS is different in Table 3. In the range of SNR, the performance of SLFN estimators is better than pilot-based algorithm. The BER gap between ELM/BP algorithm and the known channel that the receiver uses ideal channel estimation decreases as the increase of the number of pilot symbols. In Fig. 10, the performance of BP estimator is even better than that of the known
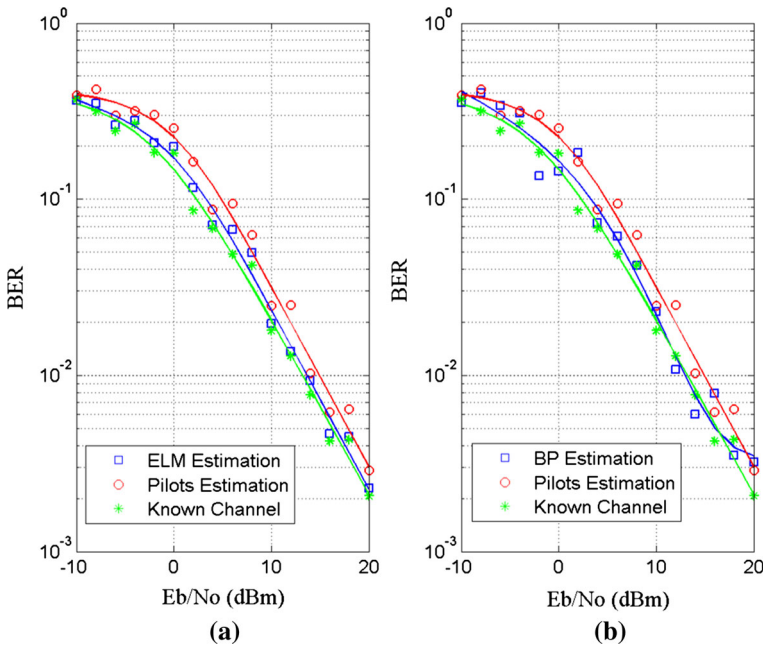
Fig. 11 BER performance of estimators versus SNR with 2 pilot symbols

channel when SNR is larger than 15 dBm. As shown in Fig. 11, although the implementation of ELM estimator is quite than BP, the improvement of ELM performance is much stable.

Therefore, both ELM and BP estimators show the effectiveness of SLFN in fading channel modelling. Due to learning speed, ELM algorithm is more practical in the real scenarios with high mobility.

5 Conclusion

In this paper, fading channel modelling based on single-hidden layer feedforward neural networks is proposed for the scenarios with high mobility. In large-scale attenuation, the ELM estimator for the prediction of path loss exponent is developed, whose experimental results show that ELM run 8000 times fast than BP learning algorithm and its testing error is acceptable. In small-scale variation, the SLFN architecture based on ELM algorithm for the estimation of fading channel transfer function is provided, and simulation results shows ELM with fast learning speed has the same effectiveness as BP algorithm.

References

Beek, J. J., Edfors, O., Sandell, M., Wilson, S. K., & Borjesson, P. (1995). On channel estimation in OFDM systems. In *Proceedings of IEEE Vehicular Technology Conference*, 2, 815C819. New York: IEEE.

Bogdan, M., Wilamowski, B., & Yu, H. (2010). Improved computation for Levenberg-Marquardt training. *IEEE Transactions on Neural Networks*, *21*, 930–937.

Calle-Sanchez, J., Molina-Garcia, M., Alonso, J., & Fernandez-Duran, A. (2013). Long term evolution in high speed railway environments: Feasibility and challenges. *Bell Labs Technical Journal*, *18*, 237–253.

Cao, J., & Lin, Z. (2015). Extreme learning machines on high dimensional and large data applications: A survey. *Mathematical Problems in Engineering, 2015*. doi:10.1155/2015/103796.

Cao, J., Chen, T., & Fan, J. (2015). Landmark recognition with compact BoW histogram and ensemble ELM. *Multimedia Tools and Applications*. doi:10.1007/s11042-014-2424-1.

Cao, J., & Lin, Z. (2014). Bayesian signal detection with compressed measurements. *Information Sciences*, *289*, 241–253.

Cao, J., Lin, Z., Huang, G., & Liu, N. (2012). Voting based extreme learning machine. *Information Sciences*, *185*, 66–77.

Cao, J., & Xiong, L. (2014). Protein sequence classification with improved extreme learning machine algorithms. *BioMed Research International*, *2014*, 103054.

Cao, J., Zhao, Y., Lai, X., Ong, M., Yin, C., Koh, Z., et al. (2015). Landmark recognition with sparse representation classification and extreme learning machine. *Journal of The Franklin Institute*, *352*, 4528–4545.

Chang, D., & Wu, D. (2014). Despread-ahead cyclic-prefix code division multiple access receiver with compressive sensing channel impulse response estimation. *IET Communications*, *8*, 1425–1435.

Cheng, C., Huang, Y., Chen, H., & Yao, T. (2015). Neural network-based estimation for OFDM channels. In *Proceeding of International Conference on Advanced Information Networking and Applications*, 29, pp. 600–604, New York: IEEE.

Chen, J. T., Paulraj, A., & Reddy, U. (1999). Multichannel maximum-likelihood sequence estimation (MLSE) equalizer for GSM using a parametric channel model. *IEEE Transactions on Communications*, *47*, 53–63.

Durgin, G. D. (2009). The practical behavior of various edge-diffraction formulas. *IEEE Antennas and Propagatation Magazine*, *51*, 24–35.

Erceg, V., Greenstein, L. J., Tjandra, S. Y., Parkoff, S. R., Gupta, A., & Kulic, B. (1999). An empirically based path loss model for wireless channels in suburban environments. *IEEE Journal on Selected Areas in Communications*, *17*, 1205–1211.

Fang, Y. (2010). Joint source-channel estimation using accumulated LDPC syndrome. *IEEE Communications Letters*, *14*, 1044–1046.

Friis, H. T. (1946). A note on a simple transmission formula. *Proceedings of the IRE*, *34*, 254–256.

Golestan, K., Sattar, F., Karray, F., Kamel, M., & Seifzadeh, S. (2015). Localization in vehicular ad hoc networks using data fusion and V2V communication. Computer Communications, In Press.

Herring, K. T., Holloway, J. W., Staelin, D. H., & Bliss, D. W. (2012). Pathloss characteristics of urban wireless channels. *IEEE Transactions on Antennas and Propagation*, *58*, 171–177.

Hills, A. (2001). Large-scale wireless LAN design. *IEEE Communications Magazine*, *39*, 98–107.

ITU-R. (2002). *Terrestrial land mobile radiowave propagation in the VHF/UHF bands*. Geneva, Switzerland: ITU Publications.

Kastell, K. (2011). Challenges and improvements in communication with vehicles and devices moving with high-speed. In *Proceeding of International Conference Transparent Optical Networks*, pp. 26–30. New York: IEEE.

Kim, W., & Hansen, J. H. (2015). Advanced parallel combined Gaussian mixture model based feature compensation integrated with iterative channel estimation. *Speech Communication*, *73*, 81–93.

Larsen, M. D., Swindlehurst, A., & Svantesson, F. (2009). Performance bounds for MIMO-OFDM channel estimation. *IEEE Transactions on signal processing*, *57*, 1901–1916.

Lee, W. C. (1985). Estimate of local average power of a mobile radio signal. *IEEE Transations on Vehicular Technology*, *34*, 22–27.

Li, S., Le-Ngoc, L., & Kwon, S. C. (2010). Turbo joint channel estimation, synchronization, and decoding for coded OFDM systems. *International Journal of Electronics and Communications*, *64*, 663–670.

Liu, J., Jin, X., & Dong, F. (2012). Wireless fading analysis in communication system in high-speed Train. *Journal of Zhejiang University*, *46*, 1580–1584.

Marchant, B. P., & Lark, R. M. (2007). Optimized sample schemes for geostatistical surveys. *Mathematical Geology*, *39*, 113–134.

Mielczarek, B., & Svensson, A. (2005). Modeling fading channel-estimation errors in pilot-symbol-assisted systems with application to Turbo codes. *IEEE Transactions on Communications*, *53*, 1822–1832.

Panda, S., Mohapatra, P. K., & Panigrahi, S. P. (2015). A new training scheme for neural networks and application in non-linear channel equalization. *Applied Soft Computing*, *27*, 47–52.

Phillips, C., Sicker, D., & Grunwald, D. (2013). A survey of wireless path loss prediction and coverage mapping methods. *IEEE Communications Surveys and Tutorials*, *15*, 255–270.

Pun, M. O., Molisch, A. F., Orlik, P., & Okazaki, A. (2011). Super-resolution blind channel modeling. In *Proceedings of IEEE International Conference on Communications*, 47, 1–5. New York: IEEE.

Qing, H., & Gang, Y. (2014). Parametric channel modeling based OFDM channel estimation. *The Journal of China Universities of Posts and Telecommunications*, 21, 1–8.

Rappaport, T. S. (2001). *Wireless communications: Principles and practice*. New Jersey: Prentice-Hall.

Robinson, J., Swaminathan, R., & Knightly, E. W. (2008). Assessment of urban-scale wirelesss networks with a small number of measurements. *Proceedings of ACM Mobicom*, 14, 187–198.

Sarma, K. K., & Mitra, A. (2013). MIMO channel modelling using multi-layer perceptron with finite impulse response and infinite impulse response synapses. *IET Communications*, 7, 1540–1549.

Seidel, S., & Rappaport, T. S. (1991). Path loss prediction in multifloored buildings at 914 MHz. *Electronics Letters*, 27, 1384–1387.

Seyman, M., & Taspinar, N. (2013). Channel estimation based on neural network in space time block coded MIMOCOFDM system. *Digital Signal Processing*, 23, 275–280.

Shu, X., Zhu, L., Zhao, H. & Tang, T. (2014). Design and performance tests in an integrated TD-LTE based train ground communication system. In *Proceedings of IEEE International Conference on Intelligent Transportation Systems*. New York: IEEE, pp. 747–750.

Simsir, S., & Taspinar, N. (2015). Channel estimation using radial basis function neural network in OFDM-IDMA system. *Wireless Personal Communications*, 82, 1–11.

Sotiroudis, S. P., & Siakavara, K. (2015). Mobile radio propagation path loss prediction using artificial neural networks with optimal input information for urban environments. *International Journal of Electronics and Communications*, 69, 1453–1463.

Tariq, S. (1991). Back propagation with expected source values. *Neural Networks*, 4, 615–618.

Taspinar, N., & Cicek, M. (2013). Neural network based receiver for multiuser detection in MC-CDMA systems. *Wireless Personal Communications*, 68, 463–472.

Valenti, M. C., & Woerner, B. D. (2001). Iterative channel estimation and decoding of pilot symbol aAssisted Turbo codes over flat-fading channels. *IEEE Journal on Selected Areas in Communications*, 19, 1697–1705.

Vieira, M. A., Taylor, M. E., Tandon, P., & Jain, M. (2013). Mitigating multi-path fading in a mobile mesh network. *Ad Hoc Networks*, 11, 1510–1521.

Yang, B., Letaief, K., Cheng, R. S., & Cao, Z. (2001). Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling. *IEEE Transactions on Communications*, 49, 467–479.

Zhou, Y., & Ai, B. (2014). Handover schemes and algorithms of high-speed mobile environment: A survey. *Computer Communications*, 47, 1–15.

Junbiao Liu received the B.S. degree in Circuit and System from Zhejiang University, Hanghzou, China, in 2003, and the M.S. degree in Control Science and Engineering from Hangzhou Dianzi University, Hangzhou, China, in 2008, respectively. He is studying for Ph.D. degree in the department of Information Science and Electrical Engineering, Zhejiang University, Hangzhou. His research interests include wireless communication, neuro networks and machine learning.

Xinyu Jin is a professor in the faculty of Information Science and Electrical Engineering since 1999, Zhejiang University, China. His research interests are network communication and intelligent electronic systems.

Fang Dong received the B. S. degree in Communication Engineering from Hangzhou Dianzi University, Hangzhou, China, in 2002, and the M.S. and Ph.D. degree in Information and Communication Engineering from Zhejiang University, Hangzhou, China, in 2005 and 2008, respectively. In 2008, she joined the department of Information and Electronic Engineering, Zhejiang University City College, where she is an assistant professor. Her research interests include cooperative communication, green communicaiton and machine learning.

Liang He received the B.M. degree in Communication Engineering from Civil Aviation University of China, Dalian, China, in 2004, the M.S. degree in Information and Communication Engineering from Zhejiang University, China in 2006 and the Ph.D. degree in Information and Communication Engineering from Tsinghua University, China in 2011. In 2011, he joined the department of Electronic Engineering, Tsinghua University, where he is an assistant professor. His research interests include speaker, language and speech recognition, pattern recognition, and machine learning.

Hong Liu is an Associate Professor in the Department of Information and Electrical Engineering, City College of Zhejiang University. His research focuses on network-wide traffic monitoring, modeling and control, which includes traffic flow modeling and simulation, traffic control and operations, and internet of vehicles. He also focuses on embedded system and control science and engineering. On these topics, he has published over 20 articles.