Channel Estimation Based on Extreme Learning Machine for High Speed Environments

Fang Dong, Junbiao Liu, Liang He, Xiaohui Hu and Hong Liu

Abstract Due to the complexity and extensive application of wireless systems, channel estimation has been a hot research issue, especially for high speed environments. High mobility challenges the speed of channel estimation and model optimization. Unlike conventional estimation implementations, this paper proposes a new channel estimation method based on extreme learning machine (ELM) algorithm. Simulation results of path loss estimation and channel type estimation show that the ability of ELM to provide extremely fast learning make it very suitable for estimating wireless channel for high speed environments. The results also show that channel estimation based on ELM can produce good generalization performance. Thus, ELM is an effective tool in channel estimation.

Keywords Channel estimation · Extreme learning machine · High speed · Path loss · COST 207

F. Dong (✉) · X. Hu · H. Liu
School of Information and Electrical Engineering, Zhejiang University City College,
HuZhou Street 50, Hangzhou, China
e-mail: dongf@zucc.edu.cn

X. Hu
e-mail: huxh@zucc.edu.cn

H. Liu
e-mail: liuhong@zucc.edu.cn

J. Liu
Institution of Information Science and Electrical Engineering, Zhejiang University,
Yugu Street 38, Hangzhou, China
e-mail: liujunbiao@zju.com

L. He
Department of Electronic Engineer, Tsinghua University,
Qinghua Street 1, Beijing, China
e-mail: heliang@mail.tsinghua.edu.cn

159

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 trans-
portation raises higher requirements for wireless communication systems, including
train ground communication (TGC) system [1], communication based train control
(CBTC) system, vehicle ad hoc network (VAN) [2], vehicle to vehicle (V2V) com-
munication, etc.

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 [3, 4].

Wireless channel play a key background role in transmission rate and quality
of mobile propagation. Only after channel characteristics in a communication sys-
tem are thoroughly researched, a variety of physical layer technologies are taken
or adapted, such as the best modulation and coding interleaving scheme, equalizer
design, or antenna configuration and subcarrier allocation in MIMO-OFDM system.
Propagation prediction or channel estimation has been extensively studied in three
areas: (1) to provide a theoretical performance bounds with information theory tool
for a new physical technology [5]; (2) to assess various candidate schemes in the
transmission system design [6]; (3) to estimate or predict channel parameters in the
deployment of a new wireless system, and then optimize deployment [7].

Based on theoretical analysis method in modeling, wireless channel model can be
divided into deterministic model, stochastic model, and semi-deterministic model
[8]. Among them, some famous models as COST 207, COST 231, WINNER ii
obtained by field measurements [9, 10] are wildly used in channel estimation. An
appropriate channel model can be selected according to a particular scenario, and
then its specific propagation parameters are set. Channel estimation in mobile prop-
agation usually has two types of technologies to obtain these parameters: blind and
pilot estimation [6, 11]. Pilot estimation is typically achieved by using pilot symbols
strategically placed at frame heads or subcarrier. In blind estimation, channel coeffi-
cients 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 paramet-
ric techniques [12]. Simsir et al. [13] 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,
[14] and [15] also proved FNN can be used in channel estimation for various wireless
environments. Unfortunately, the learning speed of FNN has been a major bottleneck

in many applications, and fast fading channel caused by high mobility makes this method unsuitable for channel estimation too. Unlike traditional FNN implementations, a simple learning algorithm called extreme learning machine (ELM) with good generalization performance [12, 16, 17] can learn thousands of times faster.

In this paper, we propose a channel estimation scheme based on ELM algorithm for high speed environments. Since researches in wireless channel have concentrate on large-scale and small-scale models [18], we choose path loss coefficient and fading classification as estimation objects. Compared with back-propagation (BP) algorithm, ELM shows its potential in channel estimation, especially for scenarios with high mobility.

The outline of the paper is as follows: In Sect. 2, ELM learning algorithm is present briefly. In Sect. 3, path loss estimation of wireless channel using ELM for high speed environments is proposed, and simulation results are analyzed. Section 4, fading classification estimation in COST 207 model based on ELM algorithm is provided. Conclusion is given in Sect. 5.

The performance of channel estimation based on ELM is in comparison with BP (LevenbergCMarquardt algorithm) which is a popular algorithm of FNN. All of the simulations are carried out in MATLAB 7.12.0. LevenbergCMarquardt algorithm is provided by MATLAB package, while ELM algorithm is downloaded from [19].

2 Review of ELM

Traditional FNN solution iteratively adjusts all of its parameters to minimize the cost function by using gradient-based algorithms. Although BP's gradients can be computed efficiently, an inappropriate learning rate might raise several issues, such as slow convergence, divergence, or stopping at a local minima.

ELM algorithm steps are as follow:

1. Assign a training set $\aleph = \left\{ \left(\mathbf{x}_i, \mathbf{t}_i \right) \middle| \mathbf{x}_i \in \mathbf{R}^n, \mathbf{t}_i \in \mathbf{R}^m, i = 1, 2, \ldots, N \right\}$, active function $g(x)$ and the number of hidden neurons $\tilde{N}$,
2. Randomly assign input weight vector $\mathbf{w}_i, i = 1, 2, \ldots, \tilde{N}$ and bias value $b_i, i = 1, 2, \ldots, \tilde{N}$,
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 $\mathbf{T} = \left[\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_N \right]^{\mathrm{T}}$.

In a word, for a linear system $\mathbf{H}\beta = \mathbf{T}$, ELM algorithm finds a least-squares solution $\hat{\beta}$ rather than iterative adjustment. 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 [12, 16] shows that ELM can produce good generalization performance in most cases and can learn more than hundreds of times faster than BP.

3 Large-Scale/Path Loss Channel Estimation

3.1 Large-Scale Channel Model

Large-scale/path loss channel models predict the mean signal strength for an arbitrary large transmitter-receiver distance (several hundreds or thousands of meters) in order to estimate the radio coverage area of a transmitter. Since the estimation of large-scale channel coefficients use statistical features of received signals, a blind estimation solution might work.

Both theoretical and measurement-based propagation channel models (such as free-space model, two-ray model, Okumura model, Hata model and etc.) [20] indicate that average received signal power P_r decreases logarithmically with distance [18]. Considering shadowing effects component ψ obeys a log-normal distribution, a statistical path loss model [21] is

$$
\begin{aligned}
P_r \text{dBm} &= P_t \text{dBm} + K \text{dBm} - 10\gamma \log_{10}\left[\frac{d}{d_0}\right] - \psi \\
&= P_t \text{dBm} + 20 \log_{10} \frac{\lambda}{4\pi d_0} - 10\gamma \log_{10}\left[\frac{d}{d_0}\right] - \psi
\end{aligned}
\tag{1}
$$

where P_t is the transmit power, γ is the path loss exponent indicating the rate at which path loss increases with distance, reference distance d_0 for practical systems is typically chosen to be 1 m, d is the transmitter-receiver distance, and shadowing effect exponent ψ is a zero-mean Gaussian distributed random variable with standard deviation σ_ψ (also in dBm).

γ is obtained by fitting the minimum mean square error (MMSE) of measurements

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

where $M = P_t/P_r$, in dBm. And the variance σ_ψ^2 is given by

$$
\sigma_\psi^2 = \frac{1}{n} \sum_{i=1}^{n} \left[M_{\text{measured}}(d_i) - M_{\text{model}}(d_i) \right]^2
\tag{3}
$$

3.2 Approximation of Path Loss Exponent

In Eq. (1), path loss exponent γ and shadowing effect exponent ψ are determined by carrier frequency and propagation terrain. Typical value of γ is between 1 and 4. The smaller γ is, the less energy loss of wireless signal due to transceiver-receiver distance is. 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.

If γ calculated distance is 1500 m and the vehicle's velocity is 120 km/h, γ needs to be calculated every 45 s; if the velocity is up to 350 km/h, γ needs to be calculated every 15.4 s. According to Eqs. (2) and (3), the calculation of γ requires hundreds or thousands of receive signal measurements, the introduction of learning algorithm into γ estimation might be effective in simplifying the data processing.

We use ELM and BP algorithms to approximate the path loss exponent γ. Without loss of generality, we set velocity $v = 120$ km/h, carrier frequency $f_c = 2.35$ GHz, transmit power $P_t = 39.5$ dBm and distance d is obtained by means of GPS [22]. 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 [23]. Shadowing effect exponent ψ has been added to all training samples while testing data are shadowing-free.

3.3   Simulation Results

The number of hidden neurons of ELM is initially set at 20 and active function is sigmoidal. Simulation result is shown in Fig. 1. The train accuracy measured in terms of root mean square error (RMSE) is 0.27734 due to shadowing effect, whereas the test accuracy is 0.012445. Figure 1 confirms that the estimation results of γ are accurate, and there is a visible margin of error only when $P_r > -30$ dBm.

Average 200 trails of simulation have been conducted for both ELM and BP algorithm, whose results are shown in Table 1. ELM learning algorithm spents 6.6 ms CPU time on training and 6.8 ms on testing, however, it takes 53.6 s for BP algorithm on training and 67.1 ms on testing. ELM runs 8000 times faster than BP. In high speed environment, when a vehicle's velocity is 120 km/h in cells with radius 1 km, it will carry out a handover procedure per 60 s; when a train's velocity is 350 km/h in same

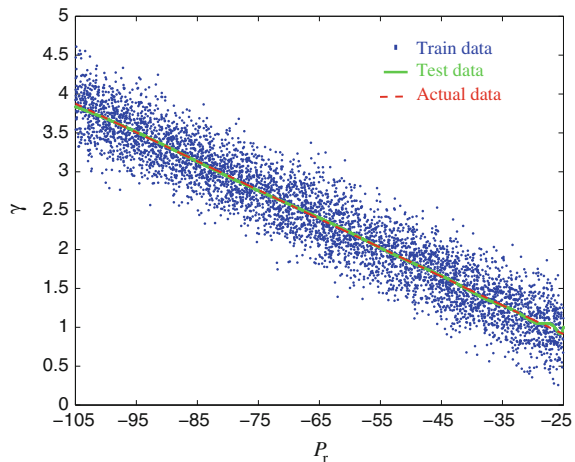**Fig. 1** The estimation of path loss exponent γ by ELM learning algorithm

Table 1 Performance comparison for learning algoritms in large-scale channel estimation

Algorithms	Time (s)		Training (s)		Testing (s)		Hidden neurons
	Training	Testing	RMS	Dev	RMS	Dev	
ELM	0.0066	0.0068	0.2828	0.0219	0.0475	0.0500	20
BP	53.6168	0.0671	0.0745	0.0028	0.0031	0.0012	20

Fig. 2 The generalization performance of ELM in estimation of path loss exponent γ
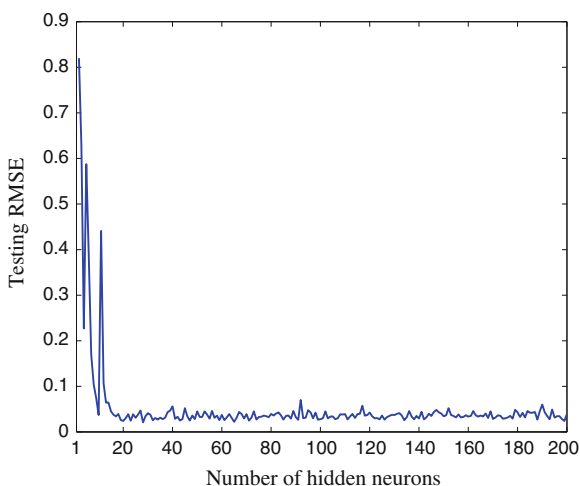
cell, handover will occur per 20.6 s. Therefore, BP is too time-consuming to be used in wireless system with high mobility. Although ELM has a much higher testing error 0.0475 compared with 0.0028 in BP, this estimation error can be acceptable in our environments. In addition, assuming that network transmission rate is 1Mbps, the collection of 1000 test data takes only 1 ms, so that a packet of 125 bytes can estimate the path loss exponent γ based on ELM with accuracy rate 95 % within a time interval of less than 8 ms.

Figure 2 shows the relationship between the generalization performance of ELM and the number of hidden neurons n for γ estimation. Every n simulates 50 times. Obviously, the generalization performance of ELM is stable when $n \geq 12$. Thus, the simulation result in Fig. 1 is reasonable when n is set to 20.

Figure 3a shows the relationship between RMSE of ELM and the number of train/test data, and Fig. 3b shows the impact of this number on consuming time. Training RMSE number of train/test data 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 training RMSE, testing RMSE decreases with increasing number of test data. The simulation confirms the conclusion in [12] that ELM has no over-trained phenomenon. Both train and
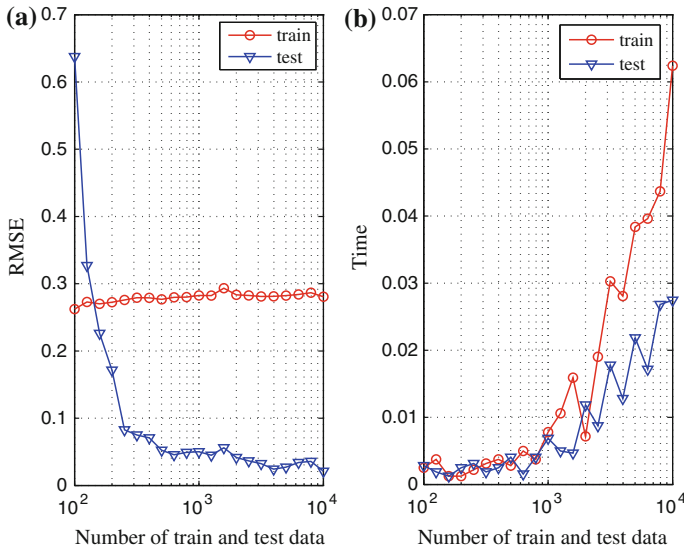
Fig. 3 Number of train/test data of ELM in estimation of path loss exponent γ

test consuming time increases with data number, however, the increase of test time is less than train time. It should also be noted that, even the number of data is up to 10^4, the time consuming of ELM is still acceptable, which is less than 70 ms.

4 Small-Scale/Fading Estimation

4.1 Small-Scale Channel Model

Small-scale/fading models characterize the rapid fluctuations of the received signal strength over very short distances (a few wavelengths) or short durations (on the order of seconds) in order to estimate the influence of multi-path propagation and the speed of a mobile terminal.

COST 207 model [9] for mobile radio specifies the power delay profiles and Doppler spread for four typical environments, i.e. rural area (RA), typical urban area (TU), bad urban area (BU) and hilly terrain (HT). The RA case consists of two distinct channel models, while the other cases each comprises four channel models. Thus, COST 207 has a total of 14 channel models: RAx4, RAx6, TUx6, TUx6alt, TUx12, TUx12alt, BUx6, BUx6alt, BUx12, BUx12alt, HTx6, HTx6alt, HTx12 and HTx12alt.

Due to radio waves' reflection and refraction, the propagation between transceiver and receiver has several paths, hence each channel model has multiple taps. For example, RAx4 is short for rural area environment with 4 taps, and HTx6alt stands

Table 2 Performance comparison for learning algorithms in small-scale channel estimation

Algorithms	Time (s)		Success rate (%)		Hidden neurons
	Training	Testing	Training	Testing	
ELM	0.0135	0.0085	86.46	72.73	20
BP	53.77	0.0694	90.60	31.08	20

for hilly terrain with 6 alternative taps. Each tap is characterized by a relative delay (with respect to the first path delay), a relative power and a Doppler spectrum category.

ELM and BP algorithms are used to estimate COST 207 channel models C_T based on modulated receive signals P_M. In order to facilitate channel estimation, each channel model is assigned an integer value [24], such as $C_T = 1$ for RAx4, $C_T = 2$ for RAx6, and etc. We still set $v = 120$ km/h, $f_c = 2.35$ GHz. Transmission rate is 1Mbps and sampling factor is 4, so that the simulation sampling rate is 4×10^6 samples per second. PSK modulation and bi-Gaussian Doppler are used in this simulation. A training set $\left(C_{Ti}, P_{Mi}\right)$ has 1000 data whereas testing set with 300 data.

4.2   Simulation Results

The hidden neurons of ELM is initially set at 20 and active function is sigmoidal. Average 50 trails of simulation have been conducted for both ELM and BP algorithm, whose results are shown in Table 2. Similarly, ELM learns up to hundreds of times faster than BP. Although BP can reach the learning rate 90.60 %, its testing rate drops to 31.08 %. On the contrary, ELM learning rate 86.46 % is slightly lower than BP, but it can achieve average testing rate 72.73 %. This is mainly because Matlab BP function *newff* doesn't support complex data. Modulated receive signals P_M must be turned into real.

5   Conclusion

In this paper, channel estimation based on ELM is proposed for high speed environments. In large-scale model, the estimation performance 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 model, fading classification estimation is provided, which shows ELM is an effective tool to classify channel type. Compared with BP, ELM still works when the elements in training set or testing set are complex. Therefore, ELM is an effective tool in channel estimation.

References

1. Shu, X., Zhu, L., Zhao, H., Tang, T.: Design and performance tests in an Integrated TD-LTE based train ground communication system. In: 17th IEEE International Conference on Intelligent Transportation Systems, pp. 747–750. IEEE Press, New York (2014)
2. Golestan K., Sattar F., Karray F., Kamel M., Seifzadeh S.: Localization in vehicular ad hoc networks using data fusion and V2V communication. Comput. Comm. (2015) (in press)
3. Zhou, Y., Ai, B.: Handover schemes and algorithms of high-speed mobile environment: a survey. Comput. Comm. **7**, 1–15 (2014)
4. Calle-Sanchez, J., Molina-Garcia, M., Alonso, J., Fernandez-Duran, A.: Long term evolution in high speed railway environments: feasibility and challenges. Bell Labs Tech. J. **18**, 237–253 (2013)
5. Singh, C., Subadar, R.: Capacity analysis of M-SC receivers over TWDP fading channels. Int. J. Electr. Comm. **68**, 166–171 (2014)
6. Vieira, M., Taylor, M., Tandon, P., Jain, M.: Mitigating multi-path fading in a mobile mesh network. Ad Hoc Netw. **11**, 1510–1521 (2013)
7. Shu, H., Simon, E., Ros, L.: On the use of tracking loops for low-complexity multi-path channel estimation in OFDM systems. Signal Process. **117**, 174–187 (2015)
8. Molisch, F., Kuchan, A., Laurila, J., Hugl, K., Schmalenberger, R.: Geometry-based directional model for mobile radio channels-principles and implementation. Eur. Trans. Telecomm. **14**, 351–359 (2003)
9. COST 207: Digital Land Mobile Radio Communications. Technical report, Office for Official Publications of the European Communities (1989)
10. 3GPP TS 45.005 V7.9.0: Radio transmission and reception (Release 7). Technical report, 3GPP GSM/EDGE Radio Access Network (2007)
11. Larsen, M., Swindlehurst, A., Svantesson, F.: Performance bounds for MIMO-OFDM channel estimation. IEEE Trans. Signal Process. **57**, 1901–1916 (2009)
12. Huang, G., Zhu, Q., Siew, C.: Extreme learning machine: theory and applications. Neuro Comput. **70**, 489–501 (2006)
13. Simsir, S., Taspinar, N.: Channel Estimation Using Neural Betwork in OFDM-IDMA System. In: 2014 Telecommunications Symposium, pp. 1–5. IEEE Press, New York (2014)
14. Cheng, C., Huang, Y., Chen, H., Yao, T.: Neural Network-Based Estimation for OFDM Channels. In: 29th IEEE International Conference Advanced Information Networking and Applications, pp. 600–604. IEEE Press, New York (2015)
15. Sotiroudis S., Siakavara K.: Mobile Radio Propagation Path Loss Prediction Using Artificial Neural Networks with Optimal Input Information for Urban Environments. In: International Journal of Electronics and Communications (2015) (in press)
16. Cao, J., Lin, Z., Fan, J.: Extreme learning machine on high dimensional and large data applications: a survey. Math. Probl. Eng. **2015**, 1–12 (2015)
17. Cao, J., Lin, Z., Fan, J.: Bayesian signal detection with compressed measurements. Inf. Sci. **289**, 241–253 (2014)
18. Rappaport, T.: Wireless communications: principles and practice. Prentice Hall Press, Cranbury (2009)
19. Huang Guang-bin. http://weibo.com/elm201x
20. Erceg, V.: An empirically based path loss model for wireless channels in suburban environments. IEEE J. Select. Areas Comm. **17**, 1205–1211 (1999)
21. Kastell, K.: Challenges And Improvements in Communication with Vehicles and Devices Moving with High-speed. In: 13th IEEE International Conference Transparent Optical Networks, pp. 26–30. IEEE Press, New York (2011)
22. Luan, F., Zhang, Y., Xiao, L., Zhou, C., Zhou, S.: Fading characteristics of wireless channel on high-speed railway in hilly terrain scenario. Int. J. Antennas Propag. **12**, 188–192 (2013)
23. Liu, J., Jin, X., Dong, F.: Wireless fading analysis in communication system in high-speed train. J. Zhejiang Univ. **46**, 1580–1584 (2012)
24. Cao J., Xiong K.: Protein sequence classification with improved extreme learning machine algorithms. BioMed Res. Int. (2014). Article ID 103054