J Inf Process Syst, Vol.13, No.1, pp.141~151, February 2017
https://doi.org/10.3745/JIPS.03.0063

ISSN 1976-913X (Print)
ISSN 2092-805X (Electronic)

JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

Wireless Channel Identification Algorithm Based on Feature Extraction and BP Neural Network

Dengao Li*, Gang Wu*, Jumin Zhao*, Wenhui Niu*, and Qi Liu*

Abstract

Effective identification of wireless channel in different scenarios or regions can solve the problems of multipath interference in process of wireless communication. In this paper, different characteristics of wireless channel are extracted based on the arrival time and received signal strength, such as the number of multipath, time delay and delay spread, to establish the feature vector set of wireless channel which is used to train backpropagation (BP) neural network to identify different wireless channels. Experimental results show that the proposed algorithm can accurately identify different wireless channels, and the accuracy can reach 97.59%.

Keywords

BP Neural Network, Channel Identification, Feature Extraction, Wireless Communication

1. Introduction

Wireless communication industry is an important communication mode, which has a huge impact on human life and society. In typical wireless channel, transmission of electromagnetic waves consists of different kinds of paths due to reflection and diffraction of signals. The multipath channel reduces the quality of received signal and causes multipath interference seriously, so an effective wireless channel identification algorithm is required to solve this problem.

The main innovative part of this paper is the feature extraction of wireless channel in different scenarios, and the establishment of character vector sets to represent the wireless channel. Actually, the number of multipath, time delay and received signals strength (RSS) of different wireless channels are likely to be different, and may exist certain rules. There have been some relevant studies, for example, several statistical features of RSS time series are exploited to identify and mitigate non-line-of-sight [1]. Li et al. [2] proposed a novel conception called demodulation-free protocol identification which only employed the features of physical layer samples, but the research about feature extraction of wireless channel is poor.

The second innovative part of this paper is using machine learning algorithm to identify wireless channel. Machine learning is an intercrossed subject which is widely used in many fields. For example,

Manuscript received April 1, 2016; accepted October 7, 2016.
Corresponding Author: Dengao Li (lidengao@tyut.edu.cn)
* College of Information Engineering, Taiyuan University of Technology, Taiyuan, China (lidengao@tyut.edu.cn, wugang0627@163.com, zhaojumin@tyut.edu.cn, {312417200, 576229258}@qq.com)

Pattanayak et al. [3] proposed an artificial neural network model for spectrum sensing in TV band specifically for detecting the presence of audio signals. Some machine-learning algorithms are used to identify wireless MAC technologies [4]. Although machine learning algorithms are widely used, there is not enough study about wireless communication.

In this paper, a wireless channel identification algorithm based on feature extraction and backpropagation (BP) neural network (NN) is proposed. Wireless channel is modeled by discrete linear system, and five indicators are selected to indicate wireless channel: the number of multipath, statistics of time delay, and delay spread, theoretical derivation. The calculation method of each characteristic parameter are presented respectively to construct a five-dimensional feature vector, which is used to train the BP NN until they are stable, so that different wireless channels can be identified effectively by the BP NN. This paper is organized as follows: in Section 2, the identification algorithm is introduced in details from two aspects, feature extraction and classification identification. In Section 3, the details of experiments and validity analysis of the algorithm based on measured data are described. Finally, some conclusions and remarks are given.

2. Wireless Channel Identification Algorithm

In this section, the algorithm is introduced from two parts: feature extraction and identification of wireless channel based on BP NN. In the first part, selection of five characteristic parameters and rigorous calculation of each parameter are given out respectively. In the second part, the identification process of wireless channel based on BP NN is described. The flow chart about the algorithm is shown in Fig. 1.

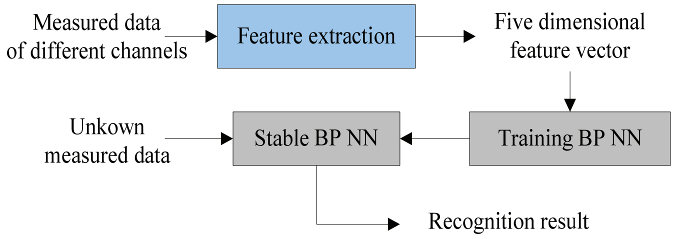

Fig. 1. Flow chart of the wireless channel identification algorithm.

2.1 Feature Extraction of Wireless Channel

In this paper, the characteristic parameters of wireless channel are selected and calculated according to the precision experiments results based on measured data, and the details about the experiments will be described in Section 3. In this paper, discrete linear system is used to model wireless channel, and the received signal is the unit sequence response of the system. In order to improve the quality of communication, filters are usually added to senders and receivers respectively, and the effect of filters can be represented by $g[m]$ equivalently, In addition, there is plural white Gaussian noise $u[k,n]$ in actual wireless channel, so the actual received signal at different time and the corresponding wireless channel model can be given respectively by

$$r[k,n] = \sum_{m=0}^{M-1} h[k-m,n] \cdot g[m] + u[m,n], k = 0,1,\cdots,K-1, n = 0,1,\cdots N-1 \quad (1)$$

$$h[k,n] = \sum_{l=0}^{L-1} h_l[n] \cdot \delta(k - \tau_l[n]) \quad (2)$$

where n represents the test moment, the sum is N. k represents the sample identification of discrete signal, the sum is K. L is the total number of path. $h_1[n]$ denotes the channel coefficients. $r_1[n]$ is the number of delayed sample points. $u[k,n]$ is plural white Gaussian noise. $r[k,n]$ is a binary function denoting the actual received signal. M is the length of filter $g[m]$. Fig. 2 shows the schematic diagram of wireless channel at different time which shows the correlation of sample n and sampling point k.

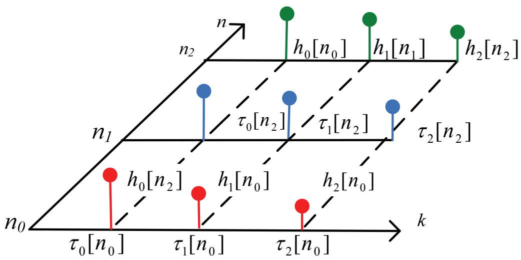

Fig. 2. Schematic diagram of wireless channel.

As shown in Fig. 2, the number of multipath, time delay and coefficient value are likely to change, and the rules can be summarized as follows.

During wireless communication, when the environment changes, the multipath number will change accordingly because of the transformation of propagation path, so the multipath number N can be selected as a characteristic parameter of wireless channel.

Multipath time delay of different channels is different because the propagation distance of electromagnetic waves along each path is different, which contains scenes characteristics, so the statistics of time delays (average $\bar{\tau}$, variance σ_r) can be selected as two characteristic parameters [5].

During wireless communication, the phenomenon that the width of received signal pulse is extended by multipath effect is called delay spread. It is caused by the reflection and scattering of propagation paths, which can represent the degree of multipath delay, so the delay spread (average delay v, mean square delay spread z) can be used as two characteristic parameters of wireless channel [6].

In conclusion, five characteristic parameters are selected to extract features of wireless channel, and the feature vector is shown as

$$K = [N, \bar{\tau}, \sigma_\tau, v, z]^T \quad (3)$$

In this paper, the multipath number N is obtained by

$$N = \sum_{k=1}^{K} N(k) = \sum_{k=1}^{K} \frac{E(k)}{P_r(k)} \quad (4)$$

where $N[k]$ is the number of arrived multipath signal. $P_r(k)$ is the energy of arrival signals of one single channel. $E(k)$ is the RSS of the multipath signals which is calculated by

$$E(k) = |r(n,k)|^2 \tag{5}$$

On the other hand, $P_r(k)$ is computed by free path propagation loss model [7] by

$$P_r(k) = P \times \left\{\frac{\sqrt{G_t\gamma}}{4\pi d(k)}\right\}^2 \tag{6}$$

where the transmitting power P, antenna gain G_t and signal wavelength γ are usually set to constant, and the length of multipath signal path $d(k)$ is decided by the arrival time t_k of the multipath signal arrived, as shown by

$$d(k) = ct_k \tag{7}$$

The computational formula of the multipath number N can be obtained according to (4)–(7), as shown by

$$N = \sum_{k=1}^{K} \alpha \cdot t_k^2 |r(n,k)|^2 \tag{8}$$

where α is a constant, $\alpha = \frac{16\pi^2 c^2}{PG_t\gamma^2}$.

In this paper, the arrival time of the received signal is used to obtain the time delay of each wireless channel, so the sample intervals of the experiment must be small enough which is set to 65 ns in the experiment to meet the requirement. We recorded the module values of received signal and the corresponding arrival time, as shown in Fig. 3.

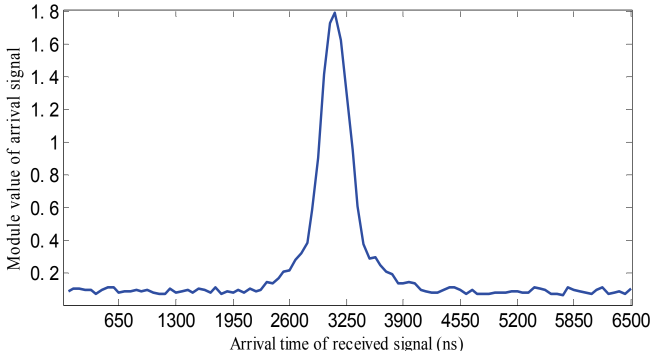

Fig. 3. Relationship between the module value of arrival signal and the arrival time.

As shown in Fig. 3, the high correlativity between the module value of arrival signal and the arrival time can be confirmed, so the mean and variance of multipath time delay can be calculated by the mean and variance of the received signal module value [8], as shown by

$$\bar{\tau} = \frac{1}{K}\sum_{k=1}^{K} |r(n,k)| \tag{9}$$

$$\sigma_\tau = \frac{1}{K}\sum_{k=1}^{K}(|r(n,k)-\bar{\tau}|)^2 \tag{10}$$

In wireless channel, delay spread is an important indicator of channel quality and is defined as the difference between maximum and minimum transmission delay. It is a statistical description about the delay characteristics of multipath channel [9]. Considering the experiment result, two parameters about delay spread are selected to indicate wireless channel (average delay v, mean square delay spread z), which are calculated by

$$v = \int_0^N tE(t)dt \tag{11}$$

$$z = \sqrt{\int_0^N t^2E(t)dt - v^2} \tag{12}$$

where $E(t)$ is normalization envelope curve of received signal, a spectrum of time delay constituted by signal with different time delays.

2.2 Identification of Wireless Channel Based on BP NN

BP NN is a classical machine learning algorithm for classification and identification. In this paper, a basic three layers BP NN with an implicit layer is constructed and trained firstly.

According to the process of feature extraction, five-dimensional feature vectors are the input of the BP NN, so the number of input neutral node is five, defined as $I_i(i = 1,2,3,4,5)$, and the hidden layer neuron node is six, defined as $h_i(j = 1,2,3,4,5,6)$. In our experiment, the measured data comes from three wireless channels, so the number of output neutral node is three, defined as $O_k(k = 1,2,3)$. The weight of connection from the input neurons to the hidden layer neurons is V_{ij}, and the weight of connection from the hidden layer neurons to the output neurons is W_{jk}. The topological structure of three layers BP NN is shown in Fig. 4.

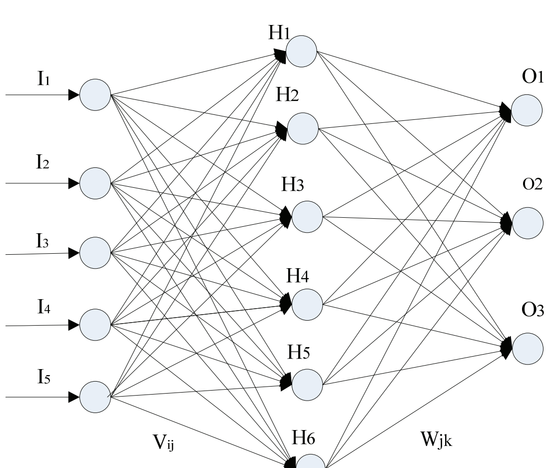

Fig. 4. The topological structure of three layers BP NN.

As shown in Fig. 4, the number of input neuron of the BP NN is five, the input activation vector is $I = [I_1, I_2, \dots, I_5]$, the number of hidden neuron is six, the hidden activation vector is $H = [H_1, H_2, \dots, H_6]$, and the activation function is f_1. The number of output neuron is three, so the output activation vector is $O = [O_1, O_2, O_3]$, and the activation function is f_2, the target vector is $T = [t_1, t_2, t_3]$. The training steps of the BP NN are as follows.

- The output vector of the hidden layer is calculated according to the input layer activation vector, which is used as the input vector of the output layer to calculate the output vector, to calculate the network error, the specific process is as follows:
  The output of neuron j of hidden layer is shown by

$$h_j = f_1 \cdot \sum_{i=1}^{n}(V_{ij} I_i + b_{1j}), j = 1,2,\cdots,6 \tag{13}$$

where b_{1j} denotes the error of neurons. The output of neuron k of output layer is shown by

$$O_k = f_2 \cdot \sum_{j=1}^{l}\left(W_{jk} h_j + b_{2k}\right), \mathrm{k} = 1,2,3 \tag{14}$$

The error function is defined by

$$E(VW, b) = \frac{1}{2}\sum_{k=1}^{3}(t_k - O_k)^2 \tag{15}$$

- If the error exceeds the expected error, the error backpropagation, and the weight is changed by gradient descent method: the change from the input of hidden neuron j to the output weight k is calculated as follows:

$$\Delta W_{jk} = -\mu \frac{\partial E}{\partial W_{jk}} = -\mu \frac{\partial E}{\partial O_k} \cdot \frac{\partial O_k}{\partial W_{jk}} = \mu(t_k - O_k) \cdot f_2' \cdot h_j = \mu \delta_{jk} \cdot h_j \tag{16}$$

where $\delta_{jk} = (t_k - O_k) \cdot f_2' \cdot W_{jk} \cdot f_1', \mu$ is learning rate.

$$\Delta b_{2k} = -\mu \frac{\partial E}{\partial b_{2k}} = -\mu \frac{\partial E}{\partial O_k} \cdot \frac{\partial O_k}{\partial b_{2k}} = \mu(t_k - O_k) \cdot f_2' = \mu \delta_{jk} \tag{17}$$

Similarly, the weight changes from the input neuron i to neuron j is calculated as follows:

$$\Delta V_{ij} = -\mu \frac{\partial E}{\partial V_{ij}} = -\mu \frac{\partial E}{\partial O_k} \cdot \frac{\partial O_k}{\partial h_j} \cdot \frac{\partial h_j}{\partial V_{ij}} = \mu(t_k - O_k) \cdot f_2' \cdot W_{jk} \cdot f_1' \cdot x_i = \mu \delta_{jk} \cdot I_i \tag{18}$$

Similarly, $\Delta b_{1j} = -\mu \frac{\partial E}{\partial b_{1j}} = \mu \delta_{ij}$.

- Weight adjustment:

$$V_{ij}(t+1) = V_{ij}(t) + \Delta V_{ij} \tag{19}$$

$$W_{jk}(t+1) = W_{jk}(t) + \Delta W_{jk} \tag{20}$$

- Threshold adjustment:

$$b_{1j}(t+1) = b_{1j}(t) + \Delta b_{1j} \tag{21}$$

$$b_{2k}(t+1) = b_{2k}(t) + \Delta b_{2k} \tag{22}$$

- The square sum of the error is calculated and adjusted again and again until all the samples are learned and trained.

Following above five steps, the BP NN is trained to be stable, and the input unknown scene could be identified according to Fig. 1.

3. Experiments and Results Analysis

3.1 Data Description and Processing Strategy

In our experiment, three different types of wireless channel from three different communication scenarios are selected. When gathering data, 7,500 samples are collected of each channel, the interval of adjacent sample is 2/15 ms, and a total of 22,500 samples are obtained, and the number of sampling points of each sample is 100, so the interval of adjacent sampling point is 65 ns, that is to say, the arrival time t_k can be calculated by k multiplied to 65 ns.

3.2 Validity Analysis

In order to verify the effectiveness of the proposed algorithm, the first 22,000 samples data in the 22,500 samples are used as training sets, and the remaining 500 are used as test sets. The whole process is implemented in MATLAB according to Fig. 1, and the results are shown in Fig. 5.

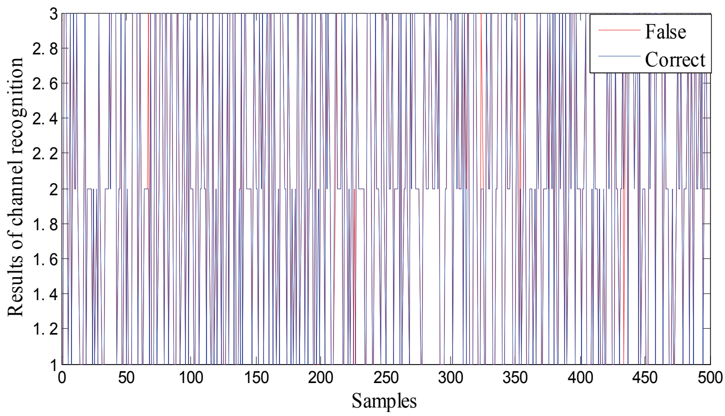

Fig. 5. Results of channel identification.

As shown in Fig. 5, the horizontal axis is 500 samples, and the vertical axis is results of channel samples identification which is the output of the BP NN from MATLAB. A small amount of test samples are falsely identified which is red part of Fig. 5.

In order to show the validity of the algorithm deeply, three evaluating indicators are used: accuracy (AC), sensitivity (SE), and specificity (SP), they are defined by using true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) of the experiment results [10]. In this paper, AC is the ability of the algorithm to identify both known and unknown types of wireless channel correctly, so the computing method of AC is shown as

$$AC = \frac{TP}{TP+FP} \times 100\% \tag{23}$$

SE measures the ability of the algorithm to identify different wireless channel correctly, and the computing method of SE is shown as

$$SE = \frac{TP}{TP+FN} \times 100\% \tag{24}$$

SP measures the ability of the algorithm to reject channel of unknown type correctly, and the computing method of SP is shown as

$$SP = \frac{TN}{TN+FP} \times 100\% \tag{25}$$

According to the test results of 500 samples in Fig. 5, TP, TN, FP, FN, AC, SE and SP of three different channel are calculated in Table 1.

Table 1. Performance of the channel identification algorithm

Channel	Results				AC (%)	SE (%)	SP (%)
	TP	TN	FP	FN			
Channel1	157	331	4	5	97.52	96.91	98.81
Channel2	163	325	5	3	97.02	98.19	98.48
Channel3	168	320	3	4	98.24	97.67	99.07

In order to analyze the effectiveness of the algorithm further, two kinds of machine learning algorithms, namely KNN and Naive Bayesian, are selected to replace the BP NN to do the contrast experiment [11], and the identification accuracy based on different times of cross validation about data set is compared, to prove that the BP NN algorithm can accomplish the task better, and the experiment results is shown in Fig. 6.

As shown in Fig. 6, the KNN and Naive Bayesian algorithm can also get high accuracy, which is lower than BP NN, and the identification accuracy of BP NN and KNN tends to be equal with the times of cross validation about data set increases, this is due to while the times of cross validation about data set is bigger, the training effects of the machine learning algorithm is better, thus the identification accuracy is higher, but due to the existence of errors, the accuracy could not reach 100%.

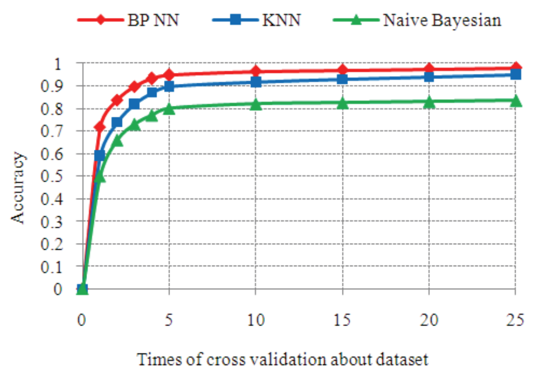

Fig. 6. Identification accuracy based on different machine learning algorithm.

4. Conclusions

In this paper, according to the real data experiment, a wireless channel identification algorithm based on feature extraction and BP NN is proposed, which has been implemented in MATLAB to verify the validity.

The contributions include: 1) The feature extraction of wireless channel is completed and a feature vector is constructed which is concise and effective; 2) BP NN is applied in the identification of wireless channel, providing a research foundation for the effective separation of multipath channel, which is an existing difficult problem due to its complex statistical properties.

Future research can address the problems of low efficiency of the algorithm due to the training process of the machine learning algorithm. In addition, to improve the performance of wireless channel identification algorithm, the process of the feature extraction could be improved, some more effective and concise characteristic parameters should be found and the calculation of the parameters should be simplified.

Acknowledgement

This paper is supported in part by the National High Technology Research and Development Program of China (Grant No. 2015AA016901), International Cooperation Project of Shanxi Province (Grant No. 201603D421012), the General Object of National Natural Science Foundation (Grant No. 61371062), the General Object of National Natural Science Foundation (Grant No. 61572346), and the Youth Science Foundation Project of National Natural Science Foundation (Grant No. 61303207).

References

[1] Z. Xiao, H. Wen, A. Markham, N. Trigoni, P. Blunsom, and J. Frolik, "Non-line-of-sight identification and mitigation using received signal strength," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1689-1702, 2014.

[2] A. Li, C. Dong, S. Tang, F. Wu, C. Tian, B. Tao, and H. Wang, "Demodulation-free protocol identification in heterogeneous wireless networks," *Computer Communications*, vol. 55, pp. 102-111, 2015.
[3] S. Pattanayak, R. Nandi, and P. Venkateswaran, "Identification of spectrum holes using ANN model in TV bands with AWGN," in *Proceedings of 2014 IEEE Asia Pacific Conference on Wireless and Mobile*, Bali, Indonesia, 2014, pp. 188-193.
[4] S. A. Rajab, W. Balid, M. O. Al Kalaa, and H. H. Refai, "Energy detection and machine learning for the identification of wireless MAC technologies," in *Proceedings of 2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*, Dubrovnik, Croatia, 2015, pp. 1440-1446.
[5] S. Buzzi, L. Venturino, and A. Zappone, "Multipath delay acquisition in asynchronous doubly-selective DS/CDMA fading channels," *IEEE Communications Letters*, vol. 14, no. 4, pp. 276-278, 2010.
[6] H. Arslan and T. Yucek, "Delay spread estimation for wireless communication systems," in *Proceedings of 8th IEEE International Symposium on Computers and Communication*, Kemer-Antalya, Turkey, 2003, pp. 282-287.
[7] Y. J. Xu and W. B. Li, "Propagation path loss prediction model of multi-sensor network in forest," *Procedia Engineering*, vol. 15, pp. 2206-2210, 2011.
[8] X. Ma and C. L. Nikias, "Joint estimation of time delay and frequency delay in impulsive noise using fractional lower order statistics," *IEEE Transactions on Signal Processing*, vol. 44, no. 11, pp. 2669-2687, 1996.
[9] K. P. Ho and J. M. Kahn, "Delay-spread distribution for multimode fiber with strong mode coupling," *IEEE Photonics Technology Letters*, vol. 24, no. 21, pp. 1906-1909, 2012.
[10] D. Li, Y. Bai, and J. Zhao, "A method for automated J wave detection and characterisation based on feature extraction," in *Proceedings of International Conference on Big Data Computing and Communications*, Taiyuan, China, 2015, pp. 421-433.
[11] R. G. Devi and P. Sumanjani, "Improved classification techniques by combining KNN and Random Forest with Naive Bayesian classifier," in *Proceedings of 2015 IEEE International Conference on Engineering and Technology (ICETECH)*, Coimbatore, India, 2015, pp. 1-4.

Dengao Li

He received B.E., M.E., and Ph.D. degree from Taiyuan University of Technology Shanxi, China. He is a Professor at College of Information Engineering in Taiyuan University of Technology. His current research fields include signal and information processing and wireless sensor network. He is focused on research of compass navigation and positioning satellite signal processing, mobile communication system, and the signal restoration, etc.

Gang Wu

He is currently pursuing M.S. degree in Taiyuan University of Technology, his research interests are navigation signal processing.

Jumin Zhao

She received B.E., M.E., and Ph.D. degree in Taiyuan University of Technology. She is currently a Professor at Taiyuan University of Technology. Her research interests include wireless sensor networks and communication signal processing.

Wenhui Niu

He received B.E. and M.E. degrees from Taiyuan University of Technology Shanxi, China. His research interests are navigation signal processing.

Qi Liu

He is currently pursuing M.S. degree in Taiyuan University of Technology, and major in Electronic and Communication Engineering. His research is GNSS signal processing.