

Radio Frequency Fingerprint Identification Based on Multi-Intervals Differential Constellation Trace Figures

Yang Yang^{1,3}, Aiqun Hu², Jiabao Yu²

{17601544638@163.com, aqhu@seu.edu.cn, yujiabao@seu.edu.cn}

School of Cyber Science and Engineering, Southeast University, 210096 Nanjing, China¹,
School of Information Science and Engineering, Southeast University, 210096 Nanjing, China²,
School of Power Supply and Engineering, Nanjing Institute of Railway Technology, 210031
Nanjing, China³

Abstract. Differential constellation trace figure (DCTF) has been demonstrated good performance on radio frequency fingerprint (RFF) identification. However, DCTF easily blurs at low SNRs. This paper proposes two novel RFF identification methods for zigbee devices based on multi-intervals DCTFs. First, a low-complexity motion features extraction method is devised based on high-density regions. Besides, an improved 3D-2D CNN model is utilized to extract motion features and spatial features. We collected 54 different ZigBee devices for experiments and classified them by these two methods. The experimental results show that compared with using single DCTF, which is generated by a single differential interval, these two methods can effectively improve the identification accuracy at different SNR levels. The classification accuracy rate of the 3D-2D CNN classifier is over 92% even under the SNR level of 5 dB.

Keywords: radio frequency fingerprint, differential constellation trace figures, 3D-2D CNN, motion features, ZigBee.

1 Introduction

The Internet of Things (IoT) combines the networking of various information sensing devices to achieve the function of interconnection of people, machines and things at any place and at any time. The development of IoT has significantly improved the intelligence of various industries. ZigBee is a new type of wireless communication technology. As an important IoT device, it has the characteristics of low speed, low power consumption and low cost. It is suitable for communication between electronic devices with short transmission distance and low speed [1]. However, due to its cost limitation, ZigBee access authentication is generally based on MAC addresses, thus it is easily attacked by rogue devices [2, 3].

Radio frequency fingerprint (RFF) has been successfully used to identify ZigBee devices [4, 5, 6]. Compared with software or protocol layer-based authentication methods, RFF identification technology uses the inherent characteristics of the physical layer analog transmission circuit hardware, and has the same advantages of anti-spoofing and anti-cloning as biological fingerprints.

Based on data samples, the identification methods of RFF can be divided into transient and steady states [7, 8, 9]. DCTF is a new type of image-based identification method [10]. After differential processing, it distorts the constellation figure by expanding the characteristics of the

fingerprint. It can reflect the characteristics of the device fingerprint through statistical information such as the contour and point distribution of DCTF, which can more intuitively identify ZigBee devices.

There are many machine learning methods suitable for RFF identification. But when faced with a large number of devices, such as large-scale wireless sensor networks, machine learning methods usually cannot achieve a high identification accuracy. As a very effective end-to-end learning method, deep learning has also been widely used in the RFF field recently [11, 12, 13]. Convolutional neural networks (CNN) is translation invariant and can learn spatial hierarchies of patterns [14, 15]. Thanks to their effective learning of complex and abstract visual concepts, they are widely used in computer vision. DCTF is an effective method to reflect the RFF of the device on the figure in the form of statistical information. The fingerprint features of the device can be clearly and prominently displayed on the figures. It is appropriate to identify DCTFs with CNN models. A DCTF classification model based on 2D CNN deep learning was proposed in [16]. Besides, the effects of parameters such as the size and the number of sample points of DCTF on the classification results were explored. The results show that DCTF can be effectively used to identify device fingerprints at high SNRs.

However, DCTF scattered points are directly distributed on a limited image size. Scattered points at different timing intervals overlap due to noise, which has a great impact on fingerprint characteristics. Compared with the methods based on baseband signal data, single-frame DCTF identification method from each ZigBee device tends to perform poorly at low SNR scenarios. Hence, an efficient approach for identifying DCTFs at low SNRs is still urgently required.

In this paper, we use multiple differential intervals to generate multi-frame DCTFs from each device, and extract motion features to identify different devices by two different methods. The main contributions of this paper are as follows:

- We propose two methods for extracting motion features from DCTFs with different differential intervals, which can be used for RFF identification. The multi-intervals DCTFs can be regarded as a continuous motion process of RFF. In order to verify our methods, we conducted a lot of experiments, and collected up to 54 zigbee devices with a total of 2430 frames of signals.
- We propose a unique method by extracting the high-density regions from DCTFs to generate trajectory figures. The generated trajectory figures can be directly used to classify the ZigBee devices. The experimental results show that this method have better classification performance under 5 dB and 10 dB than DCTF on 2D CNN.
- We also propose a 3D-2D fusion model for RFF identification by extracting motion features from 10 consecutive intervals DCTFs at the same time through 3D convolution kernels, and extracting the spatial features through deeper 2D convolution kernels. Compared with the classification method based on DCTF, this model can get a much higher accuracy rate of 92% even at 5 dB.

2 Differential Constellation Trace Figures

ZigBee uses the 2.4 GHz direct sequence spread spectrum (DSSS) O-QPSK modulation format defined by IEEE 802.15.4 [17]. Each byte is divided into two symbols, and each symbol is 4 bits. Each symbol is mapped to one of 16 pseudo-random sequences, and each sequence has 32 chips, thus completing the mapping from symbols to chips. In the end, In-phase and quadrature channels both transmit data at a speed of 1 M Chips/s. In order to complete the drawing of the

constellation trace, we use an oversampling rate of 10MHz to collect the data, that is, 10 points are used on each half-sine wave, and a total of 40K sampling points are collected as the DCTF sample points.

The signal transmitted by the ZigBee transmitter can be expressed by the following expression:

$$X(t) = (a_I x_I(t) + b_I + j(a_Q b_Q(t) + b_Q)) e^{-j2\pi f_{ctx} t} \quad (1)$$

Where a_I and a_Q are the I/Q gain imbalance factor, respectively. b_I and b_Q are the DC offset of I/Q, respectively. f_{ctx} is the carrier frequency of the transmitter.

After the signal is transmitted, assuming that the channel is ideal, the received signal on the receiver is

$$Y(t) = X(t) e^{j2\pi f_{crx} t} \quad (2)$$

Since the transmitter and receiver have different carrier frequencies $f_{ctx} \neq f_{crx}$, carrier frequency offset $\theta = f_{crx} - f_{ctx}$ can be used as a fingerprint feature. In order to obtain a stable differential constellation trace figure, we need to perform the following key steps:

- After the receiver receives the baseband signal and performs energy normalization, first perform the shift operation with shift value of φ on the I / Q channel for feature highlighting:

$$Y'(t) = (a_I x_I(t) + b_I + j(a_Q x_Q(t + \varphi) + b_Q)) e^{-j2\pi\theta t} \quad (3)$$

Since the signal of O-QPSK is a constant envelope, it is in the shape of a circle on the constellation figure. After this, it shows a distorted ring shape. This process can enlarge the morphological characteristics of the fingerprint, but it will not change the expected value of the cluster centers of the signal points.

- The rotation factor $e^{-j2\pi\theta t}$ is a function of time t , which will cause the constellation figure to change with time and appear as a rotation. By introducing a differential processing to compensate this rotation effect, the resulted constellation figure can be fixed.

$$S(t) = Y'(t) \cdot Y'^*(t + \lambda T_s) \quad (4)$$

where T_s is the sampling interval.

- After differential processing, the statistical characteristics of I and Q values can be reflected in the gray value of each pixel, that is, the higher the frequency of occurrence of the same position, the higher the brightness of the position. It is finally presented in the form of a heat figure or a gray scale figure.

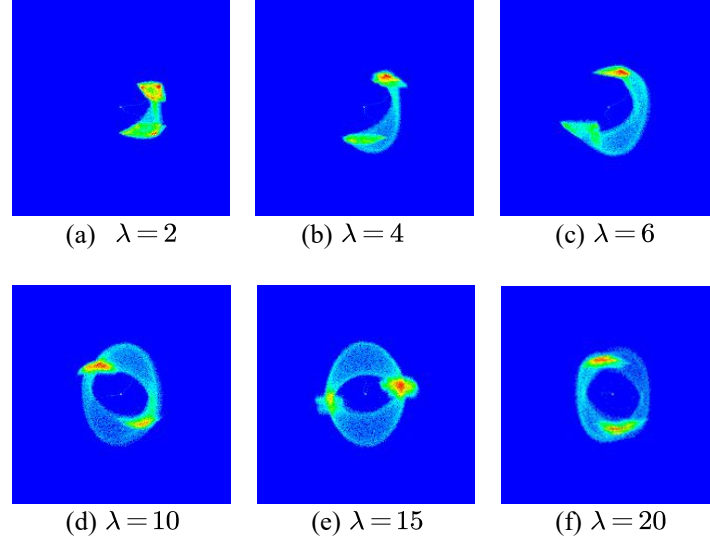


Fig. 1. DCTFs with different differential intervals λ .

Different differential time intervals λ will affect the rotation factor, so that the constellation figure has different point position distributions. When $\lambda = 0$, $S(t) = Y'(t) \cdot Y'^*(t)$, it is a positive real number, and the trajectory points are concentrated on the positive half axis of the I-axis. When λ gradually increases from 1, the aggregation points on the constellation figure gradually spread outward in the form of a distorted ring. Some DCTFs under different λ are shown in Fig. 1.

3 RFF Identification Based on motion features

DCTF under low SNRs will severely blur due to the influence of noise, resulting in a low recognition rate. RFF identification using multi-intervals DCTFs is a very feasible method.

In Section A, we propose a low-complexity motion features extraction method based on high-density region (HDR).

In Section B, we propose a 3D-2D CNN structure to identify different devices.

3.1 Proposed RFF Identification Based on HDR Method

For the different ZigBee devices l , the DCTF can generate 4 centers Γ_p^l , $p = 1, 2, 3, 4$. The mathematical expectation of the center position value is determined by the fingerprint characteristic values and the differential interval λ [10], which can be expressed as:

$$\Gamma_{p,\lambda}^l = RFF_{p,\lambda}^l \cdot e^{-j2\pi\theta\lambda} \quad (5)$$

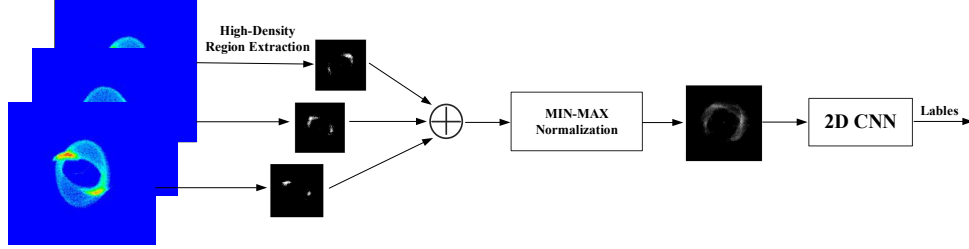


Fig. 2. A illustration of RFF identification with HDR method.

Where $RFF_{p,\lambda}^l$ are fingerprint feature values for different cluster center p under different devices l and λ . High-density regions will be formed around the centres. In low SNR scenarios, the contour features of DCTF will be destroyed. The high-density area has better anti-noise performance due to a large number of points distribution.

The matrix Φ^l of DCTF is defined as the statistical results of the signal points in each subzone ($N \times N$). In order to obtain motion trajectory figures under $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$, we set thresholds α to extract high-density regions and set the values of matrix to 255 or 0.

$$\Phi_{\alpha}^l(m, n) = \begin{cases} 255 & \text{when } \Phi^l(m, n) \geq \alpha \\ 0 & \text{when } \Phi^l(m, n) < \alpha \end{cases} \quad (0 \leq m, n < N) \quad (6)$$

Then we generate k matrices $\{\Phi_{\lambda_1, \alpha}^l, \Phi_{\lambda_2, \alpha}^l, \dots, \Phi_{\lambda_k, \alpha}^l\}$ under λ . Add these high density matrices to get the trajectory matrix $\Phi_{\lambda, \alpha}^l$.

$$\Phi_{\lambda, \alpha}^l = \sum_{i=1}^k \Phi_{\lambda_i, \alpha}^l \quad (7)$$

Normalize the maximum and minimum values of the trajectory matrix $\Phi_{\lambda, \alpha}^l$ to (0, 255) and generate a grayscale image of the normalized matrix $\bar{\Phi}_{\lambda, \alpha}^l$.

$$\bar{\Phi}_{\lambda, \alpha}^l = \frac{\Phi_{\lambda, \alpha}^l - \min(\Phi_{\lambda, \alpha}^l)}{\max(\Phi_{\lambda, \alpha}^l) - \min(\Phi_{\lambda, \alpha}^l)} \times 255 \quad (8)$$

For different devices, the generated trajectory figures can be used to identify different devices. The process can be found in Fig. 2. The trajectory figures (in RGB) generated by different devices are shown in the Fig. 3.

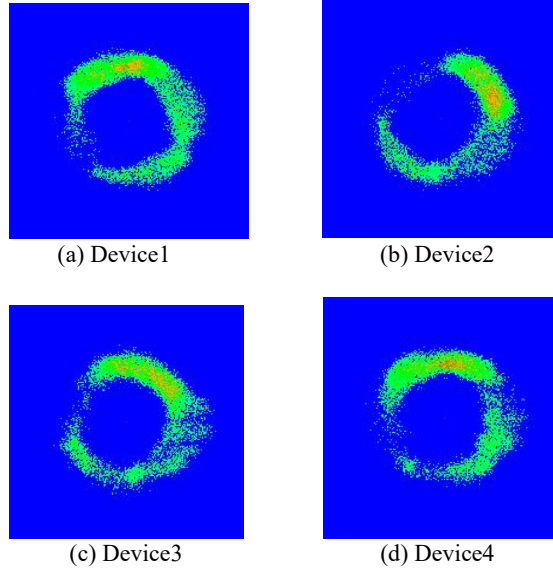


Fig.3. Generated trajectory figures from different devices ($\lambda = 15 - 20$).

3.2 3D-2D CNN model for RFF Identification

3D-CNN model can extract motion features from the multi-frame images, it is widely used in the field of human action behavior recognition. 3D convnet is well-suited for temporal feature learning, while 2D convnet only learn spatial information [18, 19, 20,21].

In order to obtain more accurate classification performance at low SNRs, we use a 2-layer 3D convolution kernels to extract the motion features between DCTFs of 10 consecutive differential time intervals $\{\lambda_0, \lambda_0 + 1, \dots, \lambda_0 + 9\}$, and add a 2-layer 2D convolution kernels after the 3D convolution kernels, which is used to extract more abstract spatial features. The 3D-2D CNN structure in Fig. 4 is as follows:

- **Input layer:** In order to reduce the complexity of the model, we use a 61x61 grayscale image instead of a color heat map at the input end to classify, that is, the size of the input end is 10x61x61x1.
- **3D convolution layer:** We use 2 consecutive 3D convolution layers to extract the motion track information between multiple frames. The convolution kernel sizes are 8x3x3x5 (8 3x3x5 3D kernels) and 16x3x3x3 (16 3x3x3 3D kernels).
- **2D convolution layer:** After the output of the second 3D convolution layer, the dimensions of the feature maps are reshaped, transformed from (57,57,4,16) to (57,57,64), and then sent to the 2D CNN to extract more abstract spatial features. The 2D kernels are 32x5x5 (32 5x5 2D kernels) and 64x5x5 (64 5x5 2D kernels). In order to reduce the dimension and size of the features, we add a layer of Max Pooling (2x2) behind each 2D kernel.

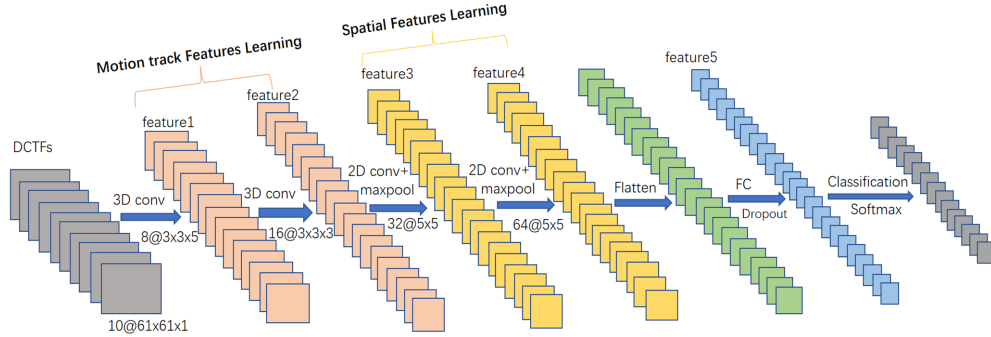


Fig. 4. 3D -2D CNN model structure.

- Fully connected layer:** The feature maps finally input to the fully connected layer. After flattening, an one-dimensional tensor is formed. After vertical connection, it is sent to the fully connected layers consisting of weights and biases. To prevent overfitting, we use the dropout strategy and set its value to 0.5. Finally, it is sent to the softmax classifier to get the predicted device number.

We divide all data sets into a training set and a test set according to 4:1. We choose K-fold cross-validation to divide all training sets into 5 partitions, one of which is the validation set and the remaining four partitions are used for training. When the classification result of the validation set is the highest, the network model is selected as the best model. Our training chooses cross entropy as the classification loss function, the SGD optimizer as the optimizer. The learning rate is 0.01, the batch size is 128. The epoch is 200, and the model parameters are total 4,005,890. The process of identification with 3D-2D CNN can be seen in Fig. 5.

4 EXPERIMENTS AND DISCUSSION

4.1 Experimental System and Processing

To verify our proposed methods (HDR and 3D-2D CNN), we collected 54 Ti CC2530 ZigBee devices for experiments. USRP N210 is used as the receiving device, and each ZigBee collects 45 frames of signals at close range for a total of 2430 frames of signals. Due to the good signal acquisition environment, the SNR of the scene is about 30 dB, and the influence of multipath is ignored. The collected signal is considered to be an ideal signal in the LOS scene.

On the PC side, we use Matlab 2016a to perform energy normalization preprocessing on the collected data. In order to evaluate the classification performance of the model under different SNRs, tested from 0-30 dB (AWGN) every 5 dB. Subsequent I/Q shift and differential processing operations are performed to finally generate multi-intervals DCTFs.

For the collected 2430 frames, we choose continuous differential intervals to generate trajectory figures and multi-frame DCTFs, then send them to 2D CNN and 3D-2D CNN for classification to verify the proposed methods separately. These models were built

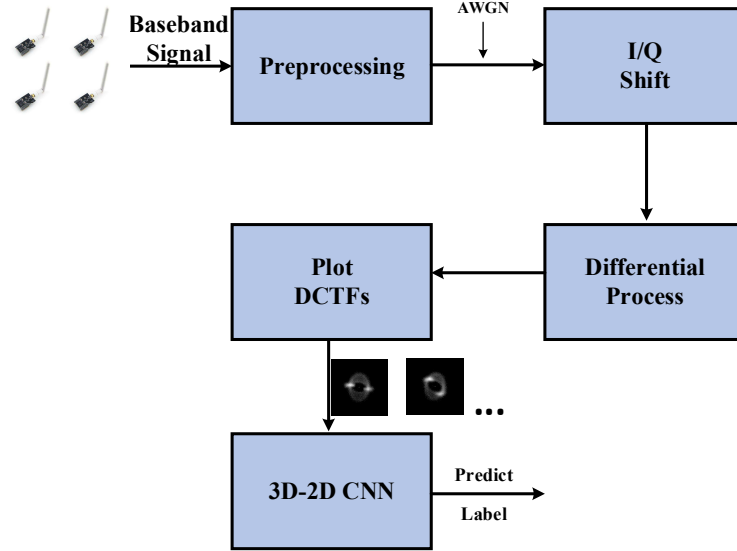


Fig. 5. RFF identification based on 3D-2D CNN.

under the Tensorflow 1.14 framework with the python 3 environment. The generation and classification process after signal acquisition are all completed on the same computer, configured as CPU I7-9700F, 32GB of RAM, GPU RTX 2060s.

4.2 Identification Accuracy with Different λ

To test the accuracy of RFF recognition based on single-frame DCTF method under different SNRs, we collected 54 devices and tested the RFF recognition performance at 5-20 dB for λ is 1-15. 2430 DCTFs are generated for each differential interval. We use the generated DCTFs to train and test on 2D CNN model. For SNRs above 20 dB, a classification accuracy of more than 98% can be obtained. When SNR is at 10 dB, the performance of RFF recognition decreases rapidly.

It can be seen from Fig. 6 that the classification performance is low and unstable when λ is 1-4. This is because when λ is small, the point distribution is more clustered, and the difference of DCTFs is small. At the same time, when λ is higher than 5, the classification performance is of slight differences and remains stable above 92%. Below 10 dB, there is a large difference in classification performance under different λ . This is because the noise destroys the characteristics of DCTF seriously, and the 2D CNN model classification cannot obtain stable classification results.

4.3 Comparison of Different Classification Methods

In order to evaluate the performance of different classification methods (3D-2D CNN versus HDR versus DCTF) under different SNRs, we divide the SNRs into two levels,

{0dB, 5dB, 10dB} as the low SNRs, and {15dB, 20dB, 25dB, 30dB} as the medium-high SNRs.

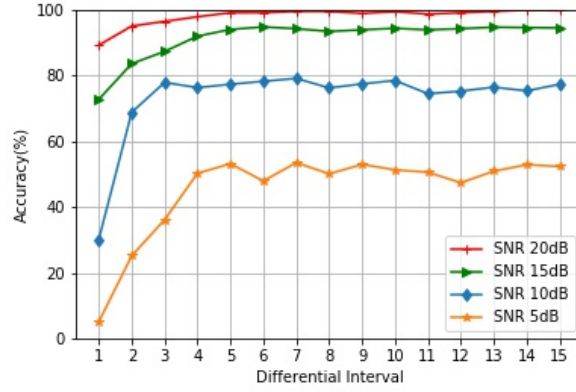


Fig. 6. Accuracy for DCTF on 2D CNN with different λ under different SNRs.

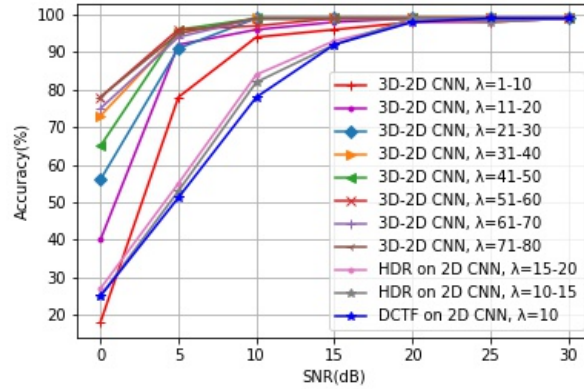


Fig. 7. Accuracy for different methods with different λ under different SNRs.

We select multi-intervals DCTFs with different λ as the input of 3D-2D CNN for classification under continuous intervals λ within 80. For the HDR method, we choose the continuous intervals λ of 10-15 and 15-20 for verification. The classification results are shown in Fig. 8 and the fusion matrices of 3D-2D CNN, HDR, DCTF are shown in Fig.8-10.

At low SNRs, the 3D-2D CNN classification performance is very sensitive to λ , and the classification variation range exceeds 60%. When λ is 1-10, the classification performance is poor, and the classification accuracy is less than 20% at 0 dB, which is

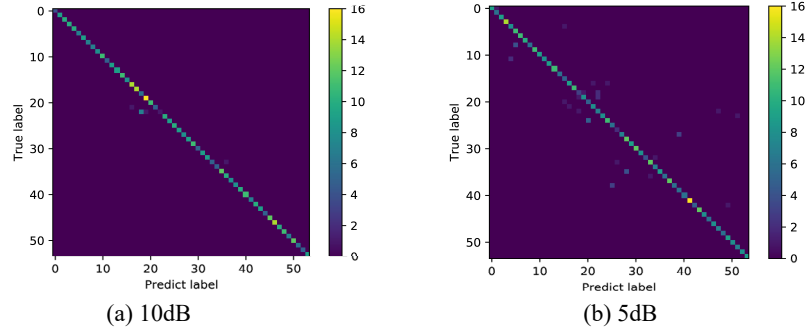


Fig.8. 3D-2D CNN fusion matrices ($\lambda = 31 - 40$).

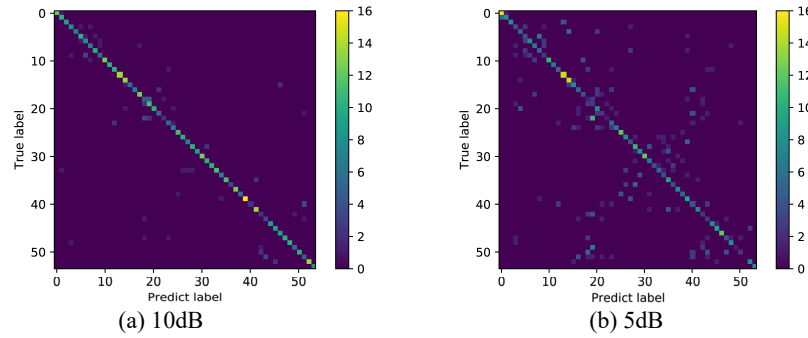


Fig.9. HDR on 2D CNN fusion matrices ($\lambda = 15 - 20$).

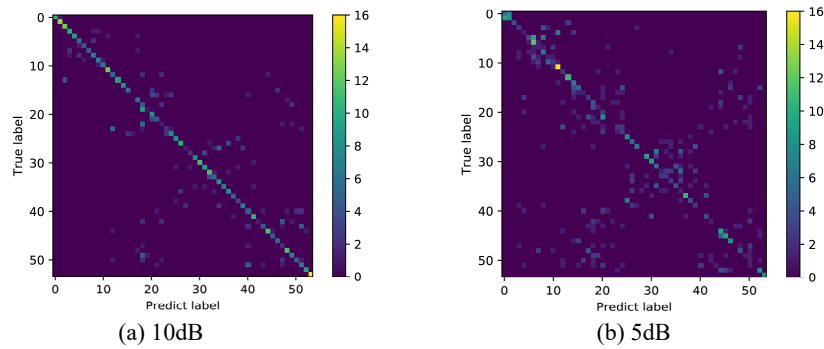


Fig.10. DCTF on 2D CNN fusion matrices ($\lambda = 10$).

lower than HDR and DCTF classification methods. When λ is small, the points of DCTF are too clustered.

When SNR is between 5 dB and 10 dB, HDR is slightly improved over the DCTF method, which proves that motion features have better anti-noise performance at low SNRs.

The classification results of the 3D-2D CNN model is rapidly improved over 5 dB, and the performance of most differential intervals is above 90%, and the highest even reaches

95.49%. It is much higher than HDR and DCTF. When SNR reaches 10 dB, all the 3D-2D models with different λ are above 94%, and the highest is 98.03%.

In the mid-to-high SNRs, when the SNR level is 15 dB, all the models of 3D-2D CNNs reach more than 98% except λ is 1-10. However, HDR and DCTF also achieved a classification accuracy of 92.1% ($\lambda = 15 - 20$) and 91.79%. When SNR is higher than 20 dB, all methods achieve a classification accuracy close to 100%.

4 CONCLUSION

This paper proposes two RFF identification methods based on continuous multi-intervals DCTFs. We first proposed a low-complexity motion features extraction method, and successfully used these features for ZigBee device identification. In order to improve the low identification rate under a low SNR level, a multi-intervals DCTFs based on DL RFF identification method is proposed. The 3D convolution kernels are used to extract more complex multi-frame motion features. Using 2D convolution kernels to extract static spatial features, classification accuracy of 92% and 98% were obtained at 5 dB and 10 dB. The classification results at low SNRs are significantly improved. In the future, we will consider ways to improve the accuracy of RFF identification based on DCTF in a time-varying multipath environment.

5 ACKNOWLEDGMENT

This work was supported in part by the Research Fund of Jiangsu Provincial Key Research and Development Program under Grant BE2019109, National Mobile Communications Research Laboratory 2020B05, Southeast University.

References

- [1] J. Han, C. Choi, and I. Lee, "More efficient home energy management system based on zigbee communication and infrared remote controls," *IEEE Trans. Consumer Electron.*, vol. 57, no. 1, pp. 85–89, February 2011
- [2] J. Niu, B. Wang, L. Shu, T. Q. Duong, and Y. Chen, "Zil: An energy-efficient indoor localization system using zigbee radio to detect wifi fingerprints," *IEEE J. Select. Areas Commun.*, vol. 33, no. 7, pp. 1431–1442, July 2015
- [3] P. Radmand, M. Domingo, J. Singh, J. Arnedo, A. Talevski, S. Petersen, and S. Carlsen, "Zigbee/zigbee pro security assessment based on compromised cryptographic keys," in 2010 International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, Nov 2010, pp. 465–470
- [4] W. Razouk, "Zigbee security within the framework of iot," in 2014 IEEE 7th International Conference on Service-Oriented Computing and Applications, Nov 2014, pp. 265–265
- [5] S. Jana and S. K. Kasera, "On fast and accurate detection of unauthorized wireless access points using clock skews," *IEEE Trans. Mobile Comput.*, vol. 9, no. 3, pp. 449–462, March 2010

- [6] Q. Tian, Y. Lin, X. Guo, J. Wen, Y. Fang, J. Rodriguez, and S. Mumtaz, "New security mechanisms of high-reliability iot communication based on radio frequency fingerprint," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7980–7987, Oct 2019
- [7] Z. Shi, M. Liu, and L. Huang, "Transient-based identification of 802.11b wireless device," in 2011 International Conference on Wireless Communications and Signal Processing (WCSP), Nov 2011, pp. 1–5
- [8] J. Toonstra and W. Kinsner, "Transient analysis and genetic algorithms for classification," in WESCANEX 95. Communications, Power, and Computing. Conference Proceedings. IEEE, 1995
- [9] O. Ureten and N. Serinken, "Detection of radio transmitter turn-on transients," *Electronics Letters*, vol. 35, no. 23, pp. 1996–1997, Nov 1999
- [10] L. Peng, A. Hu, Y. Jiang, Y. Yan, and C. Zhu, "A differential constellation trace figure based device identification method for zigbee nodes," in 2016 8th International Conference on Wireless Communications Signal Processing (WCSP), pp. 1–6, Oct 2016
- [11] S. S. Hanna and D. Cabric, "Deep learning based transmitter identification using power amplifier nonlinearity," in 2019 International Conference on Computing, Networking and Communications (ICNC), pp. 674–680, Feb 2019
- [12] Q. Wu, C. Feres, D. Kuzmenko, D. Zhi, Z. Yu, X. Liu, and X. 'Leo' Liu, "Deep learning based rf fingerprinting for device identification and wireless security," *Electronics Letters*, vol. 54, no. 24, pp. 1405–1407
- [13] J. Yu, A. Hu, F. Zhou, Y. Xing, Y. Yu, G. Li, and L. Peng, "Radio frequency fingerprint identification based on denoising autoencoders," in 2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), pp. 1–6, Oct 2019
- [14] J. Yu, A. Hu, G. Li, and L. Peng, "A robust rf fingerprinting approach using multisampling convolutional neural network," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6786–6799, Aug 2019
- [15] Y. Yuan and L. Peng, "Wireless device identification based on improved convolutional neural network model," in 2018 IEEE 18th International Conference on Communication Technology (ICCT), pp. 683–687, Oct 2018
- [16] L. Peng, J. Zhang, M. Liu, and A. Hu, "Deep learning based rf fingerprint identification using differential constellation trace figure," *IEEE Transactions on Vehicular Technology*, pp. 1–1, 2019
- [17] "IEEE standard for local and metropolitan area networks—part 15.4: Low-rate wireless personal area networks (lr-wpans)," *IEEE Std 802.15.4-2011 (Revision of IEEE Std 802.15.4-2006)*, pp. 1–314, Sep. 2011
- [18] F. Gomez-Donoso, A. Garcia-Garcia, J. Garcia-Rodriguez, S. Orts-Escolano, and M. Cazorla, "Lonchanet: A sliced-based cnn architecture for real-time 3d object recognition," in 2017 International Joint Conference on Neural Networks (IJCNN), pp. 412–418, May 2017
- [19] Z. Jiang, P. L. Chazot, M. E. Celebi, D. Crookes, and R. Jiang, "Social behavioral phenotyping of drosophila with a 2d–3d hybrid cnn framework," *IEEE Access*, vol. 7, pp. 67 972–67 982, 2019
- [20] S. Miao, Z. J. Wang, and R. Liao, "A cnn regression approach for real-time 2d/3d registration," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1352–1363, May 2016
- [21] L. Wang, W. Li, W. Li, and L. Van Gool, "Appearance-and-relation networks for video classification," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018