IEEE Sensors Letters VOL. 7, NO. 5, MAY 2023

IEEE Sensors Council

Sensor signal processing

Efficient Carrier Frequency Offset Estimation in Wireless Sensor Networks

Siqi Liu and Shaowei Wang*

School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China
**Senior Member, IEEE*

Manuscript received 23 November 2022; revised 6 February 2023; accepted 7 April 2023. Date of current version 24 April 2023.

Abstract—The carrier frequency offset (CFO) estimation is a crucial problem in orthogonal frequency division multiplexing (OFDM) systems, especially for the enabling technologies in the Internet of Things, which demand stringent limitations of low cost, low power consumption, and wide ranges, like wireless sensor networks. In this letter, we propose an efficient CFO estimation method for OFDM systems with in- and quadrature-phase (IQ) imbalance, which combines the emerging multitask learning (MTL) with the channel residual energy (CRE) to reduce the average inferring time greatly. The MTL estimator can extract the correlated features between the CFO and the IQ imbalance so as to mitigate the intractable coupling effect. On the other hand, the CRE technique can guide the neural network and loss function design to improve the estimation accuracy and the training efficiency. Numerical results demonstrate that our proposal achieves high estimation accuracy while the inferring time is significantly reduced compared with other methods.

Index Terms—Sensor signals processing, carrier frequency offset (CFO), in- and quadrature-phase (IQ) imbalance, multitask learning, orthogonal frequency division multiplexing (OFDM), wireless sensor networks.

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is the most representative multicarrier modulation scheme, and thus, is widely adopted in the enabling technologies of the Internet of Things, including Wi-Fi, wireless sensor networks, and the narrowband Internet of Things [1]. However, these systems are highly sensitive to hardware impairments resulting from the imperfection of radio-frequency components, especially for the carrier frequency offset (CFO) caused by the frequency deviation of oscillators. The frequency offset jeopardizes the orthogonality among subcarriers, introduces the intercarrier interference, and greatly degrades the demodulation performance, which needs to be estimated fast and accurately. Nevertheless, the CFO is highly coupled with the in- and quadrature-phase (IQ) imbalance that is caused by imperfect mixers, making the CFO estimation a challenging task in practice.

Moreover, the hardware impairments are aggravated in wireless sensor networks, where a wide range of cheap devices equipped with low-cost radio-frequency components and chips are deployed [2]. The low-cost components mean poor manufacturing process, leading to severe hardware impairment issues, and thus, the coupling effect is further amplified [3]. Beyond that, the limitations of low cost and low power consumption drive the estimation problem more challenging. The low-cost chips utilize low-capacity on-chip flashes equipped with limited memory and computing resources, which make most existing estimation methods intractable due to their heavy computation loads. The low power consumption is achieved by the transmit power control and sleep scheduling [4]. The power control makes most receivers deployed in the area work in low signal-to-noise ratio (SNR) regions, which raises a large estimation bias in existing estimators [5]. Hence, fast estimation methods using limited computing resources are urgently needed.

Both blind and data-aided estimation methods have been investigated in the literature [6], [7], [8], [9]. Since data-aided methods generally yield high estimation accuracy and low computational complexity compared to blind ones, we focus on the data-aided schemes in this work. In [7], a low-complexity serial method is studied, which can estimate different parameters in a predefined order. However, it has high estimation bias due to the noise impact and the coupling effect. Iterative process methods are proposed to address the coupling effect issue, which can obtain high estimation accuracy; however, they are usually time-consuming due to the large number of iterations [8], [9]. In brief, these methods are difficult to achieve fast and accurate estimation with limited bandwidth resources as the coupling effect and noise cannot be described accurately in an analytical form.

Recently, deep learning (DL) has drawn much attention for its outstanding feature extraction ability, which is exceptionally suited to solving the problems without analytical models [10]. The DL methods demonstrate impressive performance improvement as compared with conventional signal processing methods in dealing with hardware impairments, including the CFO, the power amplifier nonlinearity, and the IQ imbalance [11]. However, most existing DL methods cope with single impairment and cannot address the coupling effect issue [12]. Besides, they are purely data-driven, whose computation loads are heavy since a lot of network parameters need to be trained with a large volume of samples [13].

In this letter, we incorporate the multitask learning (MTL) [14] with the channel residual energy (CRE) and develop an efficient two-step CFO estimation procedure for the OFDM systems with IQ imbalance. The first step is for coarse estimation. We train an MTL estimator based on a shared convolutional neural network (CNN) to address the CFO and the potential IQ imbalance issue in one shot, which significantly decreases the inferring time. The shared network is exploited to extract correlated features between the two impairments, which effectively alleviate the coupling effect and noise impact. The second step is for fine estimation, where we expand the CRE into a fixed three-layer neural network. The expanded form is also applied to derive an additional regularization item, which enhances the feature extraction ability and promotes the training efficiency. Numerical

Corresponding author: Shaowei Wang (e-mail: wangsw@nju.edu.cn).
Associate Editor: F. Falcone.
Digital Object Identifier 10.1109/LSENS.2023.3266430

results show that our method achieves higher estimation accuracy than other data-aided methods, while the inferring time is decreased by an order of magnitude.

II. SYSTEM MODEL

Consider an OFDM system with K subcarriers and G lengths of cyclic prefix (CP). Assume that the channel order $L < G$ and $\mathbf{H}$ is the circulant channel matrix, where the first column is $[h(0) \cdots h(L)\, 0 \cdots 0]^T$. Denote $\mathbf{s} \in \mathbb{C}^K$ as the pilot sequence in time domain. If the normalized CFO denoted by θ causes the phase rotation $e_k = e^{j2\pi\theta(k+G-1)/K}$, the received sequence $\mathbf{r}$ is given by

$$\mathbf{r} = \mathbf{E}_\theta \mathbf{H}\mathbf{s} + \mathbf{n} \tag{1}$$

where $\mathbf{n} \in \mathbb{C}^K$ is the additive white Gaussian noise, and $\mathbf{E}_\theta = \mathrm{diag}(e_1, e_2, \ldots, e_K)$ is the phase shift matrix.

Suppose that there exists IQ imbalance with the corresponding imbalance parameters as follows:

$$\mu_i = \frac{1 + \epsilon_i e^{-j\phi_i}}{2}, \qquad \nu_i = \frac{1 - \epsilon_i e^{-j\phi_i}}{2}, \ \text{for } i \in \{t, r\} \tag{2}$$

where t and r stand for the transmitter and the receiver, respectively. ϵ_i and ϕ_i represent the amplitude and phase mismatches, respectively [7]. With the transmitter IQ imbalance, the pilot sequence is

$$\mathbf{s}' = \mu_t \mathbf{s} + \nu_t \mathbf{s}^* \tag{3}$$

and the corresponding received sequence is

$$\mathbf{r}' = \mathbf{E}_\theta \mathbf{H}\mathbf{s}' + \mathbf{n} = \mathbf{E}_\theta \mathbf{H}(\mu_t \mathbf{s} + \nu_t \mathbf{s}^*) + \mathbf{n}. \tag{4}$$

Due to the receiver IQ imbalance, the received sequence $\tilde{\mathbf{r}}$ is

$$\tilde{\mathbf{r}} = \mu_r \mathbf{r}' + \nu_r \mathbf{r}'^*. \tag{5}$$

By substituting (4) into (5), $\tilde{\mathbf{r}}$ is rewritten as

$$\tilde{\mathbf{r}} = \mu_r[\mathbf{E}_\theta \mathbf{H}(\mu_t \mathbf{s} + \nu_t \mathbf{s}^*) + \mathbf{n}] + \nu_r[\mathbf{E}_\theta \mathbf{H}(\mu_t \mathbf{s} + \nu_t \mathbf{s}^*) + \mathbf{n}]^*. \tag{6}$$

To simplify the notations, we define two complex parameters for IQ imbalance and an equivalent channel matrix as follows:

$$\alpha \triangleq \frac{\nu_t}{\mu_t}, \quad \beta \triangleq \frac{\nu_r}{\mu_r^*}, \quad \widetilde{\mathbf{H}} \triangleq \mu_t \mu_r \mathbf{H}. \tag{7}$$

Then, the received sequence is rewritten as

$$\tilde{\mathbf{r}} = (\mathbf{E}_\theta \widetilde{\mathbf{H}}(\mathbf{s} + \alpha \mathbf{s}^*) + \mathbf{n}) + \beta(\mathbf{E}_\theta \widetilde{\mathbf{H}}(\mathbf{s} + \alpha \mathbf{s}^*) + \mathbf{n})^* \tag{8}$$

and the estimation problem is formulated as

$$\max_{\theta, \beta, \alpha} \Pr(\tilde{\mathbf{r}} | \mathbf{s}). \tag{9}$$

III. EFFICIENT TWO-STEP CFO ESTIMATION

We develop a trainable CNN-based MTL estimator and a fixed three-layer neural network for coarse and fine estimation, respectively. The former infers all parameters in one shot, which can address the coupling effect with negligible inferring time. The latter embeds the CRE method to improve the estimation accuracy, which is also utilized to derive a regularization item to promote the training efficiency.

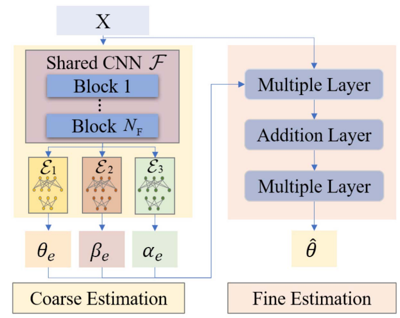

Fig. 1. Proposed CFO estimation procedure.

A. Coarse Estimation

The coarse estimation is composed of a shared CNN $\mathcal{F}$ and three fully connected subnetworks $\{\mathcal{E}_i\}_{i=1,2,3}$, as illustrated in Fig. 1. The network $\mathcal{F}$ is composed of N_{F} feature extraction blocks, which learn to extract the shared features preferred by all subtasks from the input $\mathbf{X}$. The N_{F} blocks share the same structure, which consists of two convolutional layers, one batch normalization layer, one activation layer with the rectified linear unit activation function, one addition layer, and one pooling layer. The input $\mathbf{X}$ is a $2K \times 2$ matrix consisting of the real and imaginary parts of the transmitted pilot sequence $\mathbf{s}$ and the corresponding received sequence $\tilde{\mathbf{r}}$, that is

$$\mathbf{X} = \begin{bmatrix} Re(s_1) & Re(\tilde{r}_1) & \cdots & Re(s_K) & Re(\tilde{r}_K) \\ Im(s_1) & Im(\tilde{r}_1) & \cdots & Im(s_K) & Im(\tilde{r}_K) \end{bmatrix}^T \tag{10}$$

with which network $\mathcal{F}$ outputs the shared feature $\mathcal{F}(\mathbf{X})$.

Subnetworks $\mathcal{E}_1$, $\mathcal{E}_2$, and $\mathcal{E}_3$ correspond to the estimation tasks of the CFO, the receiver and the transmitter IQ imbalance, respectively. They utilize the feature $\mathcal{F}(\mathbf{X})$ as the input, and output the corresponding estimated parameters θ_e, β_e, and α_e. Note that β_e and α_e are complex numbers, thereby we assign two neurons to the last layer of $\mathcal{E}_2$ and $\mathcal{E}_3$ to match their real and imaginary parts independently. By concatenating all outputs, we obtain the general output

$$\mathbf{y}_p = (\theta_e,\ Re(\beta_e),\ Im(\beta_e),\ Re(\alpha_e),\ Im(\alpha_e)). \tag{11}$$

B. Fine Estimation

The fixed neural network for fine estimation is developed based on the CRE method. Denote $\mathbf{s}_a$ and $\mathbf{s}_b$ as two repeated sequences with the length of $K/2$, which comprise the transmitted sequence $\mathbf{s} = [\mathbf{s_a}\, \mathbf{s_b}]^T$, and the received sequence $\tilde{\mathbf{r}} = [\tilde{\mathbf{r}}_\mathbf{a}\, \tilde{\mathbf{r}}_\mathbf{b}]^T$ is

$$\tilde{\mathbf{r}}_\mathbf{a} \triangleq e^{-j\pi\theta} \tilde{\mathbf{r}}_\mathbf{b}. \tag{12}$$

By substituting (12) into (8), the estimated CFO is

$$\hat{\theta} = \frac{1}{\pi} \angle \{(\tilde{\mathbf{r}}_\mathbf{a} - \beta_e \tilde{\mathbf{r}}_\mathbf{a}^*)^\dagger (\tilde{\mathbf{r}}_\mathbf{b} - \beta_e \tilde{\mathbf{r}}_\mathbf{b}^*)\}. \tag{13}$$

We expand (13) into a neural network to accelerate the processing by taking full advantage of the promising parallel operations, which is rewritten as

$$\begin{aligned} \hat{\theta} &= \frac{1}{\pi} \angle \{(\tilde{\mathbf{r}}_\mathbf{a}^* - \beta_e^* \tilde{\mathbf{r}}_\mathbf{a})^T (\tilde{\mathbf{r}}_\mathbf{b} - \beta_e \tilde{\mathbf{r}}_\mathbf{b}^*)\} \\ &= \frac{1}{\pi} \angle \underbrace{[\tilde{\mathbf{r}}_\mathbf{a}^* \cdot \tilde{\mathbf{r}}_\mathbf{b} + |\beta_e|^2 \tilde{\mathbf{r}}_\mathbf{a} \cdot \tilde{\mathbf{r}}_\mathbf{b}^* - \beta_e (\tilde{\mathbf{r}}_\mathbf{a}^* \cdot \tilde{\mathbf{r}}_\mathbf{b}^* + \tilde{\mathbf{r}}_\mathbf{a} \cdot \tilde{\mathbf{r}}_\mathbf{b})]}_{\triangleq \Gamma}. \end{aligned} \tag{14}$$

Since neural networks prefer real numbers, we further expand complex numbers in (14) into their corresponding real and imaginary parts. To simplify the notations, we define A, B, C, and D as follows:

$$A \triangleq Re(\tilde{\mathbf{r}}_{\mathbf{a}})^T Re(\tilde{\mathbf{r}}_{\mathbf{b}}), \quad B \triangleq Re(\tilde{\mathbf{r}}_{\mathbf{a}})^T Im(\tilde{\mathbf{r}}_{\mathbf{b}})$$
$$C \triangleq Im(\tilde{\mathbf{r}}_{\mathbf{a}})^T Re(\tilde{\mathbf{r}}_{\mathbf{b}}), \quad D \triangleq Im(\tilde{\mathbf{r}}_{\mathbf{a}})^T Im(\tilde{\mathbf{r}}_{\mathbf{b}}). \tag{15}$$

Then, we have

$$\begin{aligned}\hat{\theta} &= \frac{1}{\pi} \arctan \frac{Im(\Gamma)}{Re(\Gamma)} \\ &= \frac{1}{\pi} \arctan \frac{(1 - y_{p,2}^2 - y_{p,3}^2)(B-C)}{(1+y_{p,2}^2+y_{p,3}^2)(A+D)-2[y_{p,2}(A-D)+y_{p,3}(B+C)]}.\end{aligned} \tag{16}$$

According to (16), we develop the fine estimation neural network with three layers. The first layer calculates each item of the numerator and denominator, the second layer adds these items, and the last layer computes the fraction. Finally, $\hat{\theta}$ is given through a neuron with an arctangent activation function.

C. Loss Function

The loss function consists of main losses of the optimization objective and regularization items, and we adopt the minimum square error (MSE) criterion for the minimization. The main losses are

$$\mathcal{L}_\theta = \frac{1}{N_b} \sum_{n=1}^{N_b} (y_{p,1}^n - y_{l,1}^n)^2 \tag{17}$$

$$\mathcal{L}_\beta = \frac{1}{2N_b} \sum_{n=1}^{N_b} \sum_{i=2}^{3} (y_{p,i}^n - y_{l,i}^n)^2 \tag{18}$$

where N_b and $(\cdot)^n$ stand for the mini-batch size and the nth sample.

The auxiliary task of α estimation comprises one part of the regularization items, which introduces different noise patterns and enhances the data diversity. Thus, the noise resistance and generalization are enhanced. The item is defined as

$$\mathcal{L}_\alpha = \frac{1}{2N_b} \sum_{n=1}^{N_b} \sum_{i=4}^{5} (y_{p,i}^n - y_{l,i}^n)^2. \tag{19}$$

Another regularization item is related to $\hat{\theta}$, which forces the shared network to focus on the correlated features between the impairments by minimizing the difference between θ and $\hat{\theta}$. Hence, the network feature extraction ability is improved. The item is

$$\mathcal{L}_{cr} = \frac{1}{N_b} \sum_{n=1}^{N_b} (y_{p,1}^n - \hat{\theta}^n)^2. \tag{20}$$

Then the general loss function $\mathcal{L}$ is written as

$$\mathcal{L} = \mathcal{L}_\theta + \mathcal{L}_\beta + \gamma(\mathcal{L}_\alpha + \mathcal{L}_{cr}) \tag{21}$$

where γ is regularization coefficient.

IV. NUMERICAL RESULTS

Consider an OFDM system with $K = 64$ subcarriers, $G = 16$ lengths of CP, and the quadrature phase-shift keying modulation. The channel model follows the suggestion of COST 207, where the propagation delays are $[0, 0.1, 0.2, 0.3, 0.4, 0.5]$ μs, the power attenuations are $[0, -4, -8, -12, -16, -20]$ dB, and the user's speed is 1.5 m/s. The number of shared blocks and dense layers are $N_F = 4$ and $N_E = 4$, respectively. The batch size, the learning rate, and the regularization coefficient are given by $N_b = 100$, $\eta = 8 \times 10^{-5}$, and $\gamma = 0.1$, respectively. The well-trained estimator is then applied for online estimation, while the pure DL, the serial estimator [7], the CRE [8], and the isolated transmitter (IsoTx) [9] are also used for performance comparison. The pure DL has a similar structure to the coarse estimator but without embedding the CRE method and the regularization item.

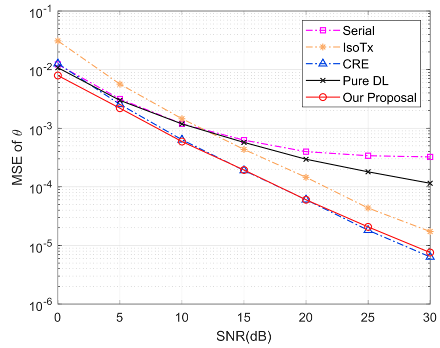

Fig. 2. Estimation accuracy of θ against different SNRs.

In Fig. 2, we depict the estimation accuracy of θ against different SNRs, where the hardware impairment parameters are $\theta = 0.2$, $\epsilon_t = \epsilon_r = \epsilon = 1.1$, and $\phi_t = \phi_r = \phi = 0.18$. As can be seen from Fig. 2, our proposal achieves the lowest MSE compared to other methods in the low SNR region [0, 15] dB, because it can average the noise patterns of different tasks using the joint loss function. In the high SNR region [20, 30] dB, the MSE of the serial method flattens because of the coupling effect, while the MSEs of other methods continue to decrease since they can address the issue. Our proposal and the pure DL apply the shared feature extraction network to extract the correlated features to fit and alleviate the coupling effect, while the CRE and the TxIso iteratively address the issue. Besides, our method outperforms the pure DL since the additional regularization item forces it to focus on the correlated relationship between two impairments, which enhances the feature extraction ability.

Fig. 3 shows the impact of the coupling effect by presenting the MSE of θ against different CFO and IQ imbalance degrees. For the former, we set $\theta \in [-0.5, 0.5]$, $\epsilon = 1.1$, and $\phi = 0.18$. For the latter, we set $\theta = 0.2$, $\epsilon = 1 + 0.2\delta$, and $\phi = 0.35\delta$, where $\delta \in [0, 1]$ denotes the IQ imbalance degree. In both cases, the SNR is 15 dB. As can be seen from Fig. 3, the MSE of the serial method varies with θ and δ, as it cannot effectively address the coupling effect issue. In contrast, the MSEs of other methods almost remain unchanged as they can address the coupling effect issue due to the same reason in Fig. 2. Besides, our proposed and the IsoTx methods outperform the CRE and the pure DL methods. Our method designs the additional regularization item to enhance the feature extraction ability, while the IsoTx performs more iterations to alleviate the coupling effect.

In Fig. 4, we display the bit-error-rate (BER) performance against different SNRs. The parameter setting is the same as that of Fig. 2, and the least-square and the zero-forcing are employed for channel estimation and signal detection, respectively. The ideal curve with perfect estimation for impairment parameters is also included. As can be seen from Fig. 4, our proposal and the CRE achieve a similar BER for a given SNR, which is significantly lower than others.

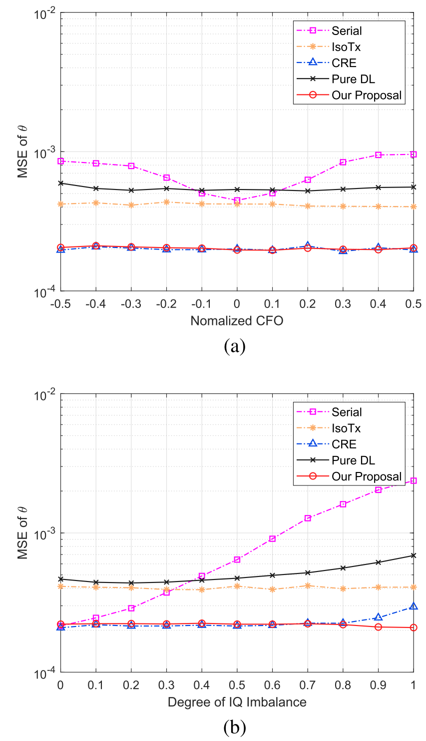

Fig. 3. Estimation accuracy of θ against different impairments. (a) Different CFOs. (b) Different IQ imbalance degrees.

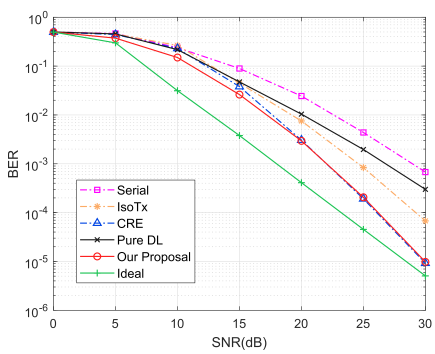

Fig. 4. BER performance against different SNRs.

TABLE 1. Average Inferring Time of Different Methods

CRE	IsoTx	Serial	Pure DL	Our proposal
480 μs	240 μs	30 μs	20 μs	**9 μs**

Table 1 shows the average inferring time, which is 480, 240, 30, 20, and 9 μs for the CRE, the IsoTx, the serial, the pure DL, and our proposed methods, respectively. All the results are averaged by 1000 Monte Carlo simulations. The former two iterative methods consume much more time for convergence than the latter three one-shot methods. Besides, the inferring time of our proposal is the shortest because it embeds the conventional method to guide the network design and training so as to reduce the computation load.

V. CONCLUSION

In this letter, we proposed an efficient CFO estimation method for OFDM systems with IQ imbalance, which is composed of an MTL coarse estimator and a fixed three-layer fine estimator. The former utilizes a shared CNN to extract correlated features between the CFO and the IQ imbalance and infers all parameters in one shot, which effectively addresses the coupling effect issue and reduces the inferring time. The latter increases the estimation accuracy by an order of magnitude using the CRE-based three-layer neural network, which is also applied to derive an additional regularization item to promote the feature extraction ability and the training efficiency. Thus, our proposal realizes faster convergency with fewer network parameters than the pure DL and achieves higher estimation accuracy than other data-aided methods with the shortest inferring time. Future works include the joint estimation of multiple hardware impairments, such as CFO, IQ imbalance, and phase noise.

ACKNOWLEDGMENT

The authors would like to thank the editors and the anonymous reviewers, whose invaluable comments helped improve the presentation of this letter substantially. This work was supported by the National Natural Science Foundation of China under Grant 61931023 and Grant U1936202.

REFERENCES

[1] K. Ashhar, M. O. Khyam, M. Noor-A-Rahim, A. Jayasuriya, and S. Cheong Boon, "Doppler correction in moving narrowband ultrasonic ranging sensors for small-scale motion tracking," *IEEE Sensors Lett.*, vol. 4, no. 9, Sep. 2020, Art. no. 7501304.

[2] A. Sarthi, D. S. Gurjar, C. Sai, P. Pattanayak, and A. Bhardwaj, "Performance impact of hardware impairments on wireless powered cognitive radio sensor networks," *IEEE Sensors Lett.*, vol. 4, no. 6, Jun. 2020 Art. no. 7500704.

[3] T. C. W. Schenk and E. R. Fledderus, "RF impairments in high-rate wireless systems - understanding the impact of TX/RX-Asymmetry," in *Proc. IEEE 3rd Int. Symp. Commun., Control Signal Process.*, 2008, pp. 117–122.

[4] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Comput. Netw.*, vol. 52, no. 12, pp. 2292–2330, Aug. 2008.

[5] S. Liu, T. Wang, and S. Wang, "Joint compensation of CFO and IQ imbalance in OFDM receiver: A deep learning based approach," in *Proc. IEEE/CIC Int. Conf. Commun.*, Xiamen, China, 2021, pp. 793–798.

[6] Y. Meng, W. Zhang, W. Wang, and H. Lin, "Joint CFO and I/Q imbalance estimation for OFDM systems exploiting constant modulus subcarriers," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 10076–10080, Oct. 2018.

[7] J. Tubbax et al., "Joint compensation of IQ imbalance and frequency offset in OFDM systems," in *Proc. IEEE Glob. Telecommun. Conf.*, San Francisco, CA, USA, 2003, pp. 2365–2369.

[8] Y.-H. Chung and S.-M. Phoong, "Joint estimation of I/Q imbalance, CFO and channel response for MIMO OFDM systems," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1485–1492, May 2010.

[9] Y. Zhuang and Y. Wan, "Joint estimation of carrier frequency offset and in-phase/quadrature-phase imbalances for orthogonal frequency division multiplexing systems," *Comput. Elect. Eng.*, vol. 48, no. 5, pp. 1–11, Nov. 2015.

[10] Z. Chen, M. I. AlHajri, M. Wu, N. T. Ali, and R. M. Shubair, "A novel real-time deep learning approach for indoor localization based on RF environment identification," *IEEE Sensors Lett.*, vol. 4, no. 6, Jun. 2020, Art. no. 7002504.

[11] S. Liu, T. Wang, and S. Wang, "Toward intelligent wireless communications: Deep learning based physical layer technologies," *Digit. Commun. Netw.*, vol. 7, no. 4, pp. 589–597, Dec. 2021.

[12] W. Yu, T. Wang, and S. Wang, "Multi-label learning based antenna selection in massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 7, pp. 7255–7260, Jul. 2021.

[13] S. Liu, T. Wang, and S. Wang, "Hardware impairment estimation in NB-IoT: A parallel multi-task learning method," *IEEE Internet Things J.*, vol. 10, no. 8, pp. 6859–6869, Apr. 2023.

[14] Y. Zhang and Q. Yang, "A survey on multi-task learning," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 12, pp. 5586–5609, Dec. 2022.