IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 12, 2017, accepted February 15, 2017, date of publication February 24, 2017, date of current version March 28, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2674967

User Capacity of Wireless Physical-Layer Identification

WENHAO WANG[1,2], (Student Member, IEEE), ZHI SUN[2], (Member, IEEE), KUI REN[3], (Fellow, IEEE), AND BOCHENG ZHU[1]

[1]School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China
[2]Department of Electrical Engineering, University at Buffalo, The State University of New York, Buffalo, NY 14260, USA
[3]Department of Computer Science and Engineering, University at Buffalo, the State University of New York, Buffalo, NY 14260, USA

Corresponding author: W. Wang (wenhaowang@pku.edu.cn)

The joint-training Ph.D. study of W. Wang was supported by the China Scholarship Council. The work of Z. Sun was supported by US NSF under Grant CNS-1547908. The work of K. Ren was supported by US NSF under Grant CNS-1318948. The work of B. Zhu was supported by the China National Key Research and Development Project under Grant 2016YFB0502103.

ABSTRACT A wireless physical layer identification (WPLI) system aims at identifying or classifying authorized devices of different users based on the unique radio frequency fingerprints (RFFs) extracted from their radio frequency signals at physical layer. Most existing works mainly focus on demonstrating feasibility of system by presenting the classification performance of a fixed-user-number network. However, the real-world WPLI systems are expected to work in various scenarios with different scales of user numbers, dynamic changes in user numbers, and various choices of user combinations. Hence, an important question needs to be answered: what's the user number that a WPLI system can support, i.e. the user capacity, under the condition that the minimum system performance is guaranteed. In thispaper, we theoretically characterize the user capacity of WPLI. A theoretical approach is proposed based on ensemble mutual information between RFF and user identity. With this approach and one-time RFF training, the user capacity of WPLI can be characterized under various real-world constraints. Extensive experiments are conducted to validate the accuracy and tightness of the theoretically derived WPLI user capacity.

INDEX TERMS Radio frequency fingerprinting, physical-layer security, mutual information, wireless communication.

I. INTRODUCTION

A Wireless Physical Layer Identification (WPLI) system aims at identifying or classifying authorized devices of different users based on the unique Radio Frequency Fingerprints (RFFs) extracted from their radio frequency signals at physical layer [2], [3]. Because the software-level device identities (e.g., IP or MAC address) are vulnerable to be manipulated, while the physical layer feature cannot be modified without significant efforts, WPLI has emerged as a promising wireless security solution. Fig. 1 illustrates the processing procedures of WPLI and typical application scenarios. Each user in a WPLI system is assigned with one device to represent his or her identity, i.e., the identity of each device is the user identity. The signal, transmitted by users within the network, firstly pass through the communication channel. The signal is then obtained by the identification system. Next, a feature extraction module is to obtain selected kinds of fingerprinting feature to form a fingerprint. A fingerprint matcher compares the fingerprints with reference fingerprints stored in database according to authorized identities. (Those reference fingerprints and identities are collected through system training procedure.) The identities are finally classified and assigned to devices using dimensionality reduction classification technique. Here, two application scenarios are involved in WPLI: (i) identification scenario is between all unauthorized imposters and the whole authorized users. (ii) classification scenario is the N-class identification between all authorized users within this network [2], [4]. In these scenario, WPLI can be utilized to successfully defend impersonation attacks within network and fake-identity injection attacks from the outside intruder respectively.

In essence, RFFs are random variables carrying discriminating information of user identity which can be expressed in different feature domain, such as frequency domain, time domain, wavelet transform domain etc. These features have non-Gaussian and non-uniform feature distributions [5], [6]. Due to different channel effects, in-band noises, feature extraction algorithms, and classification techniques,

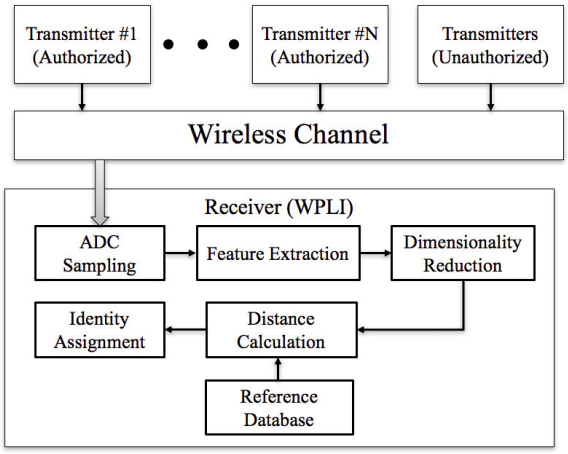


FIGURE 1. Typical physical-layer processing procedure of WPLI.

the extracted RFFs are involved with more noise and interferences, which bring more randomness to the feature distributions [2], [7], [8]. If more users are kept adding into the WPLI system, the feature distribution intervals of different identities are more likely to overlap within the classification dimensionality. Hence, the uncertainty between feature and identity is increased, causing higher classification errors. For example, as reported in [9], 6000 RFID tags of 12 typical models are tested using the phase RFF named TagPrint. From their results, we can observe that there exists an obvious classification performance drop when the user number is increased: the classification success rates drop from 99.58% for 50 users to 55.31% for 6000 users. Moreover, since RFF's distribution is non-uniform, RFFs from different device combinations within the same user number can have significantly different distribution interval overlaps, resulting in different classification results. Currently, the existing works of WPLI mainly focus on the experimental performance of a scenario with a fixed user number and user combination [5], [10]–[12]. However, the real-world WPLI systems are expected to work in various scenarios with different scales of user numbers, dynamic changes in user numbers, and various choices of user combinations. Hence, it is of great importance to have a thorough understanding on what's the user number that a WPLI system can support, i.e., the user capacity, under the condition that the minimum system performance is guaranteed. It should be noted that user capacity is the concept for classification scenario of WPLI. Because the identification capability of outside attackers is based on how close between the features from attackers and genuine users, which we've already modeled in our previous work [7]. Identification procedure can still be working even when user capacity is full, as long as RFFs from attackers are differentiable comparing to all the RFFs from all genuine users. Hence in this work, the user capacity is proposed to describe the N-class identification capability between all the genuine users and defend all the attacks within the network.

It is challenging to characterize the WPLI user capacity based on existing solutions. First, system designers have to run lots of experiments to check the classification error rates of different user numbers, from small values increased to large ones. Second, given a large user number N, the classification results of different combinations of k users can be different. Consequently, to derive the user capacity, the WPLI system designers have to conduct experiments with all C_N^k combinations of genuine users and obtain all the classification error rates. The final user capacity should be determined by the utmost user number, of which the worst classification error rate of its user combinations is within guaranteed error rates. Third, the WPLI systems using different classification algorithms or different classifier subspace dimensionality have different WPLI error rates [5], [11], [13]. The user capacity should be derived under the condition that the best possible classification algorithm is selected, which requires the system designer tests all possible classification algorithms combined with all possible subspace dimensionality in the experiments. Fourth, as the modeling in our previous work [7], many real-world factors, including the receiving equipment and the wireless channel, can dramatically affect the performance of WPLI. Hence, the user capacity of WPLI can be easily affected by those real-world constraints. For example, if high-end measurement equipment is used and the wireless channel is a line-of-sight (LOS) AWGN channel with high SNR [10]–[12], the WPLI user capacity is expected to be high. In contrast, if cheaper off-the-shelf wireless devices are used in the noisy multipath channel, the user capacity can be much lower. The experimental measurement based-on existing solutions cannot have the repeatable results with high accuracy due to different application scenarios, experiment setups. Hence, lots of experiments would need to be conducted for various real-word constraints of application scenarios and experiment setups.

Due to the above challenges, existing solutions are extremely inefficient, if not infeasible, to comprehensively characterize the WPLI user capacity. A theoretical approach rather than the experimental solution is highly desired. In this paper, we address the above challenges, by deducting a theoretical approach to characterize the user capacity of WPLI within targeted error performance, under typical real-world constraints, despite different classification techniques and user combinations. In this approach, mutual information and ensemble mutual information are utilized as a fundamental metric to describe the uncertainty between feature and user identity in WPLI, despite different selections in feature types and dimensionality [10], [14]. With these theoretical models, we only need to utilize the one-time collection of RFF samples from all possible genuine user devices using a good-quality receiver (e.g., we use a USRP N210 software-defined radio) in the near region (i.e., good LOS channel with high SNR). Based on the modeling of the effects of typical real-world constraints (e.g., different quality receiver device or different wireless channel), WPLI user capacity can be accurately captured. Because the classification error rate is theoretically bounded by entropy and mutual information [15], the derived WPLI user capacity indeed

represents the performance upper-bound for different user combinations, despite various classification techniques adopted in WPLI. In practical usage, these raw training data can be collected from database during the roll-in procedure of each WPLI system. Then, the WPLI user capacity can be theoretically characterized for the targeted scenarios under major real-world constraints, which saves great efforts on conducting lots of field experiments. Moreover, this theoretically derived user capacity can set an upper bound on the performance of designed WPLI system, which can help WPLI system designers to select the most significant RFF feature, compare different algorithms, determine equipment settings, and improve the system designing for target application scenario.

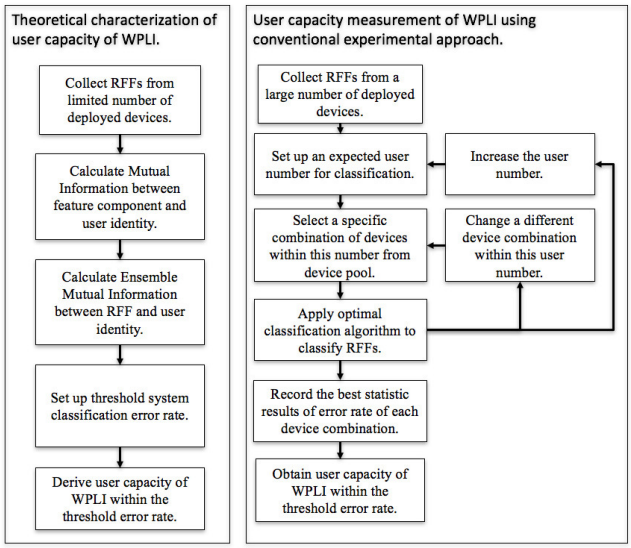


FIGURE 2. Approach comparison between theoretical characterization and conventional measurement of user capacity of WPLI.

The rest of this paper is organized as follows. In section II, related works are summarized. In section III, the user capacity is then characterized using ensemble mutual information between multi-dimensional RFF and identity, which is based on the modeling of typical signal processing procedures in WPLI. After that, we present application cases to illustrate usage of this theoretical characterization as the left side of Fig. 2 shows. In section IV, we conduct experiments on classification error performance of a practical WPLI system to measure the user capacity and validate our user capacity characterization as the right side of Fig. 2 shows. Meanwhile, we conduct the experiments under different application case setting to validate the user capacity characterization under various real-world constraints. The paper is finally concluded in section V.

II. RELATED WORK

In the field of information theory, most existing works focus on the channel capacity of a secure communication system, but not the WPLI user capacity that focuses on the classification performance. In [16]–[19], the maximal achievable data rate of a secure communication system is analyzed and modeled based on Shannon channel capacity theory. Some other works in security field use mutual information concept to improve the performance of WPLI for a fixed user network, which have no study on the user capacity of WPLI to date. In [10], mutual information concept is utilized to select the most significant feature point to improve classification accuracy of WPLI. Similarly, in [6], entropy is utilized to evaluate the randomness and uncertainty of different RFF features as a reference to choose the optimal RFF feature.

The state of art WPLI techniques are comprehensively surveyed in [2] and [3]. Various WPLI solutions have been proposed with different type fingerprinting source selection, which include clock jitter [20], [21], the DAC sampling error [22], the power amplifier non-linearity [22]–[24], the mixer and frequency synthesizer phase error [25], [26], device antenna (including polarization) [4], modulator subcircuit (if the analog modulation is used) [12], multipath wireless channel characteristics [27]–[29], among others. Various physical-layer identification feature extraction algorithms have been proposed, which utilize RFFs from different parts of signal and interpret the RFFs in different ways. In [12] and [30], the RFF is represented by the modulation errors (the distortions of amplitude, phase, and frequency errors in time domain) of a transmitted data packet or the packet preamble. Meanwhile, in [4], [21], and [31]–[33], the signal spectrum of the data/preamble is used as the RFF. Moreover, in [11], [34], [35], [36], and [37], the turn-on/off transient of the transmitted signal is utilized as the RFF. However, the above works focus on demonstrating the feasibility of proposed WPLI system by presenting the experimental classification performance of a fixed-user-number network, not theoretically characterizing their user capacities. The theoretical analyses on WPLI system are relatively less in numbers comparing to the solutions. In [22], [24], [38], and [39], hypothesis testing techniques are utilized to theoretically analyze the identification problem. In [8], the receiver distortions on WPLI performance is investigated and a theoretical bound of identification is discussed. Those works focus on the theoretically analyses on the WPLI performance for a given fix-number of users with selected type of feature, not the WPLI user capacity that we will discuss in this paper.

In our previous work [7], a theoretical framework is established to systematically analyze the complete procedures of WPLI. The classification and identification error rates of given genuine users and features can be theoretically calculated using the proposed framework. The uniqueness of RFF feature is mathematically modeled in details in order to theoretically evaluate the performance of WPLI. Special attention is paid to effects of real-world constraints on the performance of WPLI, such as the influence of multipath channel and low-end receiver devices on the WPLI error rates. It should be noted that the developed theoretical framework in [7] highly depends on the knowledge of the RFF feature model, i.e., we need to derive an accurate model of the

selected RFF feature (e.g., the hardware imperfection of RF circuit), which is very difficult in practice. Moreover, that paper still focuses on the error rate performance of fixed number and combination of user devices, not the WPLI user capacity. In this paper, the ensemble mutual information is utilized as the fundamental metric for all types of features. We do not have to theoretically model RFF feature itself. The user capacity of WPLI using multiple types of features can be characterized based on our proposed theoretical approach, which is easy to use in the real-world applications of WPLI. Meanwhile, the dynamic changes of user capacity of WPLI in different application scenarios are characterized, analyzed and validated to present a complete survey of user capacity of WPLI.

In order to propose the basic idea of this work, a previous conference version was published in [1]. Comparing to the conference version, more comprehensive theoretical modeling is presented in this paper. We model the whole physical-layer processing procedure of WPLI as identity in transmitter, communication channel effects, and receiver effects respectively in Section III. While in the previous conference version, only the sampling procedure in the receiver is modeled. Moreover, we study user capacity characterization under large scale fading effects and multipath channel effects in this paper. In the previous conference version, only AWGN channel is applied in the application cases and only the effects of AWGN noise level on user capacity is analyzed. While in this paper, typical indoor fading and multipath channel parameters are utilized in the channel model to obtain the user capacity of WPLI over Rician and Rayleigh channels. Besides these, in the experimental part of this work, we add a series of experiments to measure the user capacity of a WPLI system without the pre-knowledge of the characterization results, as the comparison and validation for our theoretical results. Also, we add two experiment cases to validate the user capacity characterization under fading channel. Hence, more comprehensive theoretical analyses, more advanced application cases, and more complete experiments are developed in this paper.

III. THEORETICAL CHARACTERIZATION OF USER CAPACITY OF WPLI

In this section, we first discuss briefly the typical WPLI processing procedure [2], [7], [12], which is illustrated in Fig. 1. The influential factors within those procedures are analyzed according to their effects on the WPLI performance. Then we discuss the proposed theoretical approach to characterize the user capacity of WPLI, as the left side of Fig. 2. The mutual information between fingerprinting feature component and user identity is modeled. Then ensemble mutual information between RFF and identity is derived. Based on that the WPLI user capacity is characterized. After that, we use several application cases to illustrate the usage of the developed theoretical approach. The user capacity of WPLI is characterized under practical constraints of different application case settings.

A. PHYSICAL-LAYER PROCESSING PROCEDURE OF WPLI

For WPLI, each user is assigned with one device to represent his user identity. As the discussion in existing works, the identity of each user device is granted at the transmitter end. The hardware imperfections in the RF chain can be modeled as a fingerprint channel according to the modeling in [17]. This channel is not a real physical communication channel, but a general model to represent the RF chain within each transmitter. Various RFF features are added into ideal signal during the signal processing procedures in RF chain. Hence, the transmitted RF signal embedded with identity feature can be generally expressed as

$$F(t) = T_{Fingerprint}(A(t), W_f), \tag{1}$$

where $T_{Fingerprint}(.)$ is the equivalent function of fingerprint channel of RF chain at transmitter, $A(t)$ is the ideal transmitting signal, W_f is the fingerprint channel parameters which are to represent various RFF features discussed in Section II, e.g., RF nonlinearity, timing jitters. $F(t)$ is the final transmitted signal embedded with identity feature. This model is different from our previous work [7] or other existing works [22], [24], [38], in which different types of features are mathematically modeled in details. Here, we use a general formulation to represent the fingerprinting feature, because the proposed theoretical approach in this paper does not depend on the pre-knowledge or model of any particular RFF feature, meanwhile no existing work can perfectly model the combined effects of all types of imperfections within transmitter.

The transmitted signal then goes through the physical communication channel. The RFF features are affected by the channel effects. Hence, more uncertainty and randomness are involved in feature distributions through this procedure. Because the typical WPLI applications are designed for mainly indoor scenarios, here we consider channel model to be a stationary channel model including the classical large scale fading model and multipath impulse model from [40], [41], which can be denoted as,

$$G(t) = C_{Physical}(F(t), H(t, \alpha_l, L_p, \theta_l, \tau_l), \alpha_{PL}, \eta); \tag{2}$$

where $C_{Physical}(.)$ is the equivalent function of physical communication channel, $H(t) = \sum_{l=1}^{L_p} \alpha_{PL}\alpha_l e^{-j\theta_l}\delta(t-\tau_l)$ is the multipath channel impulse response function, which includes the effects from number of paths L_p, small scale amplitude factor α_l, phase shift θ_l and time delay τ_l of each path, large scale fading factor α_{PL} can be calculated with the large scale fading equation of $PL(d) = PL_0 + 10n\log(\frac{d}{d_0})[dB]$, and η is the AWGN noise.

The signal is then captured by the receiver of WPLI and affected by different receiver setting. In our previous work [7], we model, analyze and conclude that the effects of different receiver setting can have different impacts on performance of WPLI, such as receiver carrier demodulation and ADC sampling etc. As the discussion in Section II, this received signal at WPLI receiver can be either baseband

signal or passband signal. Consequently, different types of signal are utilized to extract the fingerprints, which are baseband preamble [21], [31] or passband transient signal respectively [11], [37]. Hence, for a more general application, in this paper, we neglect the possible analog carrier demodulation and only discuss the sampling procedure of ADC. After ADC sampling procedure, the next procedure is to obtain the RFF from the signal, which can be modeled as,

$$\mathbf{X}_{1:M} = R_{Receiver}(G(t), f_s, Q_{ADC}, N_S), \tag{3}$$

where $R_{Receiver}(.)$ is the equivalent function of signal processing procedures at WPLI receiver, $\mathbf{X}$ is the fingerprint set with the feature dimensionality M. Meanwhile $\mathbf{X}$ can be seen as a projection of fingerprint channel model parameters W_f in certain feature extraction algorithm, for example, the FFT or wavelet transformation adopted by WPLI receiver. f_s is the sampling rate of ADC at receiver. N_S is the number of signal samples which often decides the resolutions and amount of information of the RFF, e.g., $N_S = N_{FFT}$ controls the FFT resolution for spectral RFFs [7]. Q_{ADC} is the ADC quantization bits, for Q-bit ADC quantization and input dynamic range U Vp-p, the maximum quantization error is $\delta_{ADC} = 2^{-Q}U$. It should be noted that $\mathbf{X}$ is not constrained as a single value nor a specific type of feature. Recently, more high-dimensional feature and multiple combined feature are widely utilized in the feature selection, e.g., high-dimensional spectral feature [11], channel state information (CSI) or other channel uniqueness feature [29], a combined feature of TIE error and average signal power [21], a combined feature of various modulation errors [12]. Hence, $\mathbf{X}$ here is a general representation form for RFFs.

After that, different dimensionality reduction techniques are applied to reduce the computation burden and find more discriminant subspaces which highlight the relevant features that may be hidden in noise [2]. The identities are finally classified and assigned to devices according to reference RFFs in database. Although various techniques are utilized here, only classification results of a fixed number of users can be obtained for each round, which are far less enough to characterize the real-world performance of WPLI as we discussed in previous section. In order to obtain the user capacity for various feature dimensionality selection and despite various classification approaches of WPLI, the following theoretical approach is developed by us based on mutual information and ensemble mutual information.

B. MUTUAL INFORMATION BETWEEN FEATURE AND IDENTITY

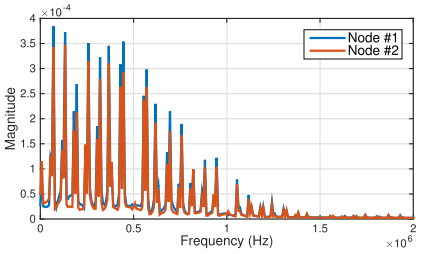
To characterize the user capacity N_C, we firstly need to calculate the mutual information between the extracted RFF feature $\mathbf{X}$ and its user identity. In information theory, entropy is defined as a measure of the uncertainty on the values taken by a feature component. Meanwhile, as the discussion in [10] and [14], mutual information can be seen as the reduction in the uncertainty of one feature component due to the knowledge of user identity. Hence, we utilize mutual information between RFF feature and user identity as a metric to evaluate the uniqueness and randomness of RFFs' feature distribution. It should be noted that some related works utilize other metrics to evaluate RFF feature distribution. In [42], Kolmogorov-Smirnov test and F test statistic values of features are calculated to evaluate their uniqueness. In [43], a table of confusion matrix of 12 testing devices are shown to illustrate the uniqueness and similarity between their RFFs. However, these metrics are either difficult to calculate or inconvenient to understand for readers comparing to utilizing mutual information as the metric. Mutual information can be a general metric widely applied to evaluate the relevance of the feature to its identity.

Specifically, variable X is the one single dimensional component of the RFF feature, i.e., $X \in \mathbf{X}$. Y denotes the device identity of this RFF feature, which is also representing the unique user identity assigned with this device. The value of X and Y varies for each testing RFF sample received by WPLI. The number of X values is N_X. The entropy of feature values can be calculated as, $H(X) = -\sum_{i=1}^{N_X} p(x_i)\log(p(x_i))$. N_Y is the number of Y, which is also the quantity of users connected to WPLI. The classification procedure of WPLI is to utilize lots of samples of X with different identities to decide the user identity Y. Hence, the conditional entropy, which describe the uncertainty remaining in X after obtained the outcome of Y, can be calculated as $H(X|Y) = -\sum_{j=1}^{N_Y} p(y_j)\sum_{i=1}^{N_X} p(x_i|y_j)\log(p(x_i)|y_j)$. The mutual information between X and Y can be finally derived as,

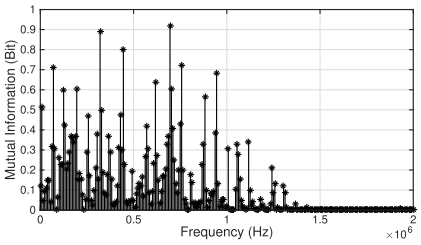
$$\begin{aligned} I(X;Y) &= I(Y;X) = H(X) - H(X|Y) \\ &= \sum_{j=1}^{N_Y}\sum_{i=1}^{N_X} p(x_i y_i)\log\left(\frac{p(x_i y_i)}{p(x_i)p(y_j)}\right). \end{aligned} \tag{4}$$

For instance, if the whole signal spectrum is used as feature $\mathbf{X}$ for RFF, each frequency point is one feature component of $\mathbf{X}$, and each value of frequency point can be seen as the one variable X for the spectral feature. x_i is the specific magnitude value of each frequency point of the ith RFF sample. Hence, mutual information between each frequency point and identity can be measured and calculated using large number of tests. To have a clear understanding of mutual information between feature component and identity, in Fig. 3 we present two PSDs of signal preambles of two Micaz sensor nodes and corresponding mutual information between each frequency point and signal identity calculated with equation (4). The difference of spectrums is utilized to classify the identities. From the figures, we can see that the frequency points which have larger differences in spectrum also show larger mutual information value with their identity. Just as our previous discussion, mutual information is a significant metric to characterize the relevance of the feature to its identity.

If only single dimensional feature is utilized to form a RFF, the mutual information between RFF and identity can already be calculated using equation (4). However, as our previous discussion, even single dimensional features are combined to



(a)



(b)

FIGURE 3. (a) Preamble PSDs of Micaz sensor nodes. (b) Mutual information between spectral feature component and identity.

form a multi-dimensional feature $\mathbf{X}$ to form a RFF. Hence, the ensemble mutual information between ensemble feature and user identity, $I(\mathbf{X}; Y)$, needs to be calculated to characterize the relation between the ensemble RFF and the identity. In [44] and [45], one definition of ensemble mutual information between high-dimensional feature and user identities is given as,

$$\begin{aligned} I(\mathbf{X}; Y) &= \sum_{N_Y} \int p(\mathbf{x}, y) \log \frac{p(\mathbf{x}, y)}{p(\mathbf{x})p(y)} d\mathbf{x} \\ &= \sum_{N_Y} p(y) \mathbb{E}_{\mathbf{x}|y} \left[\log \frac{p(\mathbf{x}|y)}{p(\mathbf{x})} \right], \end{aligned} \tag{5}$$

However, in practice, the increase of dimensionality of ensemble feature would cause very large numbers of possible variable values, complicated dependency of each feature component, and huge computation burden. All these factors would make the calculation from natural definition infeasible. Hence approximate calculation methods are needed [46]. The pdfs $p(\mathbf{x}|y)$, $p(\mathbf{x})$ and the conditional expectation $\mathbb{E}_{\mathbf{x}|y}$ can be approximately calculated using non-parametrical Kernel Density Estimator (KDE) with $K(.)$ as the kernel [47],

$$\begin{aligned} I(\mathbf{X}; Y) &\approx \sum_{N_Y} \frac{p(y)}{N_{\mathbf{x}}^y} \sum_{j=1}^{N_{\mathbf{x}}^y} \log \frac{(1/N_{\mathbf{x}}) \sum_{i=1}^{N_Y} K(\mathbf{x}_j^y - \mathbf{x}_i^y)}{(1/N_{\mathbf{x}}^y) \sum_{i=1}^{N_Y} K(\mathbf{x}_j^y - \mathbf{x}_i)} \\ &\approx \sum_{N_Y} \frac{p(y)}{N_{\mathbf{x}}^y} \sum_{j=1}^{N_{\mathbf{x}}^y} \log \left[\frac{\bar{\varphi}^T(\mathbf{x}_j) \bar{\Lambda} \bar{\boldsymbol{\mu}}_y}{\bar{\varphi}^T(\mathbf{x}_j) \bar{\Lambda} \bar{\boldsymbol{\mu}}} \right] \end{aligned} \tag{6}$$

where the kernel $K(.)$ can be calculated with the eigenvectors $\boldsymbol{\Phi}(\mathbf{x})$ and the eigenmatrix $\bar{\boldsymbol{\Phi}}_{\mathbf{x}} = [\boldsymbol{\Phi}(\mathbf{x})_1, \ldots, \boldsymbol{\Phi}(\mathbf{x})_N]$, $\bar{\boldsymbol{\mu}}_y = (1/N_{\mathbf{x}}^y) \bar{\boldsymbol{\Phi}}_{\mathbf{x}} \mathbf{m}_y$ is the average eigenvector for user y, $\bar{\boldsymbol{\mu}} = (1/N_{\mathbf{x}}) \bar{\boldsymbol{\Phi}}_{\mathbf{x}} \mathbf{1}$ is the average eigenvector for all the training samples, $N_{\mathbf{x}}^y$ is the number of RFF feature samples of user y, N_Y is number of user identities, and $N_{\mathbf{x}}$ is the number of all RFF feature samples.

C. USER CAPACITY OF WPLI

With the ensemble mutual information $I(\mathbf{X}; Y)$ obtained, the important property of mutual information related to classification error rate P_e can be utilized to derive the user capacity of WPLI. In practice, the WPLI system finally assigns the user identity of testing RFF according to minimal feature distance scores between reference RFFs. After a large number of tests, the WPLI classification performance can be evaluated using average classification error rate as the metric [4], [21]. However, in [14], [15], and [48], the information-theoretic bounds for classification error rate are given in details using Fano's Inequality (note that the inequality is valid for three or more classes scenario). Hence, the classification error rate of WPLI can be bounded as,

$$P_e \geq \frac{H(Y) - I(\mathbf{X}; Y) - H(P_e)}{\log(N_Y - 1)}. \tag{7}$$

The bound determines that no classifier can possibly achieve better than this lower error bound. This means despite the various classification algorithm adopted by WPLI, the classification error rate is information-theoretically restricted by two terms. One is the ensemble mutual information between feature and identity $I(\mathbf{X}; Y)$, the other is user identity entropy $H(Y)$.

Based on these, we originally propose the definition of user capacity of WPLI. Considering a specific scenario of WPLI, the stable mutual information between feature component and identity, can be measured with a large number of test samples [10] and the ensemble mutual information can be calculated using equation (5). Hence, the classification error rate can be bounded by the user identity entropy which is directly related to the user number N_Y of WPLI. Considering an equal-identity-probability WPLI system, i.e., $H(Y) = \log(N_Y)$, the user capacity can then be derived as,

$$\begin{aligned} N_C = \max\ &(\mathbf{N}_Y), \\ s.t.\ &(\frac{\log(N_Y) - I(\mathbf{X}; Y) - H(\lambda)}{\log(N_Y - 1)} \leq \lambda). \end{aligned} \tag{8}$$

where N_C is the user capacity, which is the maximum number of $\mathbf{N}_Y$, the set of all possible user numbers that satisfy the Fano's Inequality. Y is the user identity, and λ is the performance threshold for classification error rate P_e. In theory, the performance threshold λ can be any desired value in the value range of probability and the corresponding user capacity can be derived. However, in practice λ is often determined by desired system performance, which is alway a very low error probability for a typical security system. In [9] and [42], the classification success/error rates of typical WPLI methods are tested and summarized. The best performance can be less $P_e = 0.42\%$, while the average performance varies in $P_e = 4\% \sim 20\%$. Although it varies a lot in different

scenarios for different system, to have a far discussion in this work, we choose both $\lambda = 1\%$ and $\lambda = 10\%$ as two threshold λ values to derive the user capacity as instances. By far, the theoretical approach to characterize the user capacity of WPLI is given.

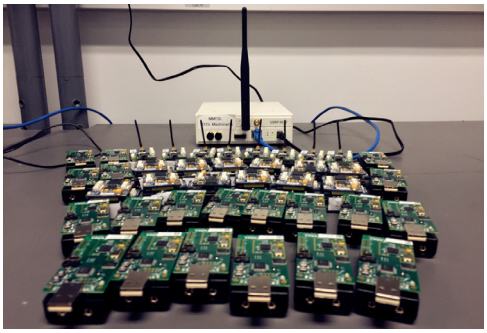


FIGURE 4. The experimental equipment.

TABLE 1. Case overview.

Item	Instruction
Feature selection	FFT spectrum of baseband preamble, [4], [10], [21], [31].
Transmitter selection	Micaz, Imote2 and TelosB sensor nodes (with same ZigBee Protocol radio chip). [49], [50], [51].
Battery selection	All the transmitters are equipped with new AA Batteries of the same brand.
Receiver selection	USRP N210 with SBX daughter board (14-bit ADC) [52].
Communication channel	indoor AWGN channel (SNR= 20~30 dB).
Sampling rates	2 M~10 MS/s.
Number of FFT points	64~2048 p.
Number of transmitters	40 in all.
Number of signal samples	2000 samples per user.

D. USER CAPACITY CHARACTERIZATION

To apply this theoretical approach, we use several application cases to illustrate its usage for WPLI. We use raw training RFF samples collected from real devices to calculate ensemble mutual information between RFF feature and device identity. Then the user capacity is characterized under practical constraints of different application case settings. The stability of this characterization is analyzed at first. Then, the effects of key system parameters on user capacity are evaluated and analyzed. It should be noted that most existing works try to present the best performance with the high quality receiving equipment. If high-end measurement equipment is used, the WPLI user capacity is expected to be high. In contrast, if cheaper off-the-shelf wireless devices are used, the user capacity can be much lower. In order to obtain more achievable results for real-world applications, we try to derive the user capacity and evaluate the system feasibility using the existing approach under practical constraints of off-the-shelf devices. The equipment used for this application case study are presented in Fig. 4, the details of which are given in Table 1. After one-time collection of the raw signals from sensor nodes, training samples of RFF can be obtained. The samples can be reprocessed in Matlab to regenerate corresponding RFFs due to desired typical system setting and scenario setting using the channel models and receiver models provided in our previous work [7]. In each target scenario, we calculate ensemble mutual information for each RFF and its user identify using the non-parametrical Kernel Density Estimation approach in equation (6). With the ensemble mutual information obtained, the next step is to characterize the user capacity for this scenario using equation (8), as the left side of Fig. 2 shows.

Since we obtain the raw RFF samples from limited number of transmitters and limited number of RFF samples to characterize the user capacity, firstly, we characterize user capacity results using RFF samples (randomly picked 2000 samples for each device) from 3 to 40 training transmitters to find out what's the least number of user identities we need to characterize a stable user capacity for this system in one application scenario, i.e. to study the stability of this characterization method. (It should be noted that this subsection is the validation of theoretical user capacity not the classification results.) Subsequently, we use all the samples from 40 nodes to characterize user capacity under different constraints of key parameters in section I, including in-band AWGN noise level, ADC quantization bits Q, number of FFT points N_{FFT} and sampling rate of receiver f_s. We set the parameters due to different typical application scenarios of WPLI and derive the user capacity with targeted performance as the following figures. In each case, we present the ensemble mutual information (EMI) (blue curve) together with user capacity under 1% and 10% classification error rate, i.e., $N_C|(P_e \leq 1\%)$ (black curve) and $N_C|(P_e \leq 10\%)$ (red curve).

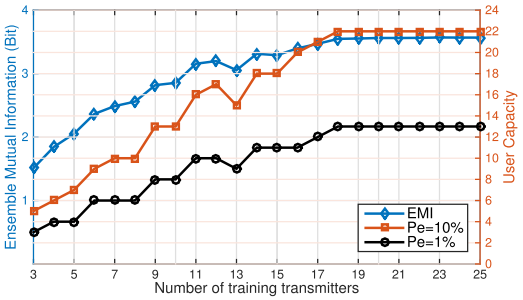


FIGURE 5. User capacity under different numbers of training transmitters.

1) STABILITY OF USER CAPACITY CHARACTERIZATION

After we collect the raw RFF samples and choose one typical system setting for ideal close distance indoor scenario, as $f_s = 4$MS/s, $N_{FFT} = 512$p, and SNR= 24dB. As we use ZigBee protocol, of which signal bandwidth is 2MHz, here we set baseband sampling rate as 4 MS/s, which is typical value to cover the main spectrum of signal. The more

information on the relation between spectral RFF and sampling rate can be found in our previous work [7]. In Fig. 5, the user capacity is characterized using samples from 3 to 25 training transmitters (the results from 25 to 40 stay the same). The ensemble mutual information directly related to classification error rate and user capacity, the relation of which can be easily observed in the figures. In the beginning, when the number of training transmitters is too small, the obtained ensemble mutual information is also small which results in the user capacity is near to N_Y. As the increase of N_Y, the results are increased unstably. After the number of training transmitters is larger than 18, the characterization result becomes a stable and reliable result which is $(N_C = 13)|(P_e \leq 1\%)$ and $(N_C = 22)|(P_e \leq 10\%)$. Hence, in order to characterize the user capacity we at least should use 18 nodes to collect the raw training RFF samples.

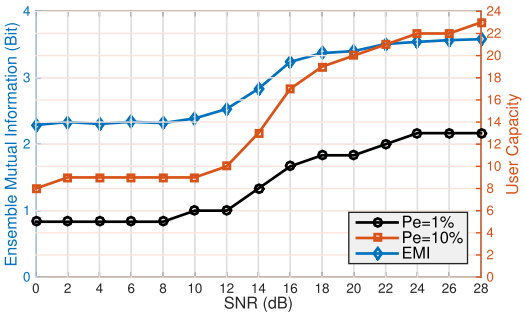


FIGURE 6. User capacity under different AWGN level.

2) EFFECTS OF AWGN NOISE LEVEL

We firstly present user capacity due to effects of AWGN noise level. Here, we set the other parameters as $f_s = 4MS/s$, $N_{FFT} = 512p$. As the discussion in equation (2), the AWGN noise level significantly affects the classification performance of WPLI resulting in the decrease of user capacity we finally obtained. We simulate the noise feature value within SNR= $0 \sim 28$dB, which cover most common SNR levels for the wireless sensor networks we set up. The user capacity results are presented in Fig. 6. The ensemble mutual information and user capacity are decreased synchronously as the AWGN SNR level decreases. It should be noted that, according to equation (8), the user capacity we obtained is the upper bound for all classification methods and classifiers using all these RFF samples. Hence, in high SNR situations, the typical classification procedure of WPLI can easily achieve the user capacity quite accurately. However, in the extremely low SNR scenarios, it is hard to use a single method or feature to achieve the upper bound of user capacity. Hence, more combined features extracted from RFF for multiple classifiers should be utilized to achieve the upper bound of user capacity, as the work in [12] and [21]. Moreover, in the low SNR scenarios, the error in signal acquisition procedure of WPLI can also contribute to worsen the classification performance [31], which is out of the scope of this paper and can be discussed in future works.

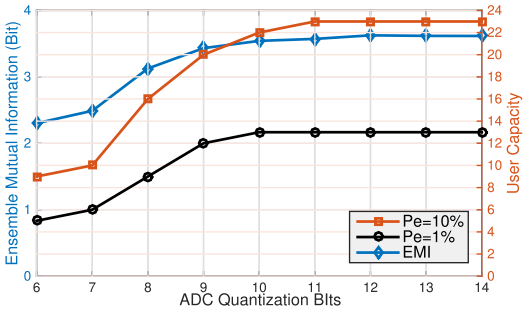


FIGURE 7. User capacity under different ADC quantization bits.

3) EFFECTS OF ADC QUANTIZATION BITS

In order to study the effects of ADC quatization bits, we need to choose a scenario with rarely high SNR and no channel effects. We choose SNR= 29dB and simulate the feature value of ADC quantization error within $Q = 6 \sim 14$bits. The corresponding user capacity results are presented in Fig. 7. The user capacity becomes stable when the ADC quantization bits are increased to 10 bits. In practical applications, the effects of channel noise level usually more significant than the effects of ADC quantization error, while the effects of ADC quantization error can be significant when channel SNR level is very high. Here we only can simulate the feature value for 14 bits quantization due to the constraints of USRP daughter board, while in practical, 16 or more quantization bits can also be found in higher standard equipment. With the development of device resolutions, ADC quantization noise is irrelevant and can be neglected.

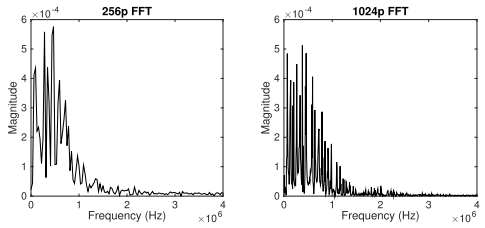


FIGURE 8. Preamble spectrum of signal under different number of FFT points.

4) EFFECTS OF NUMBER OF FFT POINTS

Here we present user capacity due to various number of FFT points within the same sampling rate setting. The number of FFT points N_{FFT} is the key parameter to form the spectral RFF which decides the resolution of the spectral feature $\mathbf{X}$ in equation (3) and consequently reflects the distribution of RFF feature. The specific modeling about the number of FFT points of spectral feature can be found in [11]. In Fig. 8, we present the preamble spectrum obtained with two different number of FFT points to present difference of resolution. In order to eliminate the effects of channel noise, we choose a scenario with extremely high SNR. Here, we set the other parameters as $f_s = 8$MS/s, SNR= 29dB, $Q = 14$bits. We simulate the result within number of FFT points

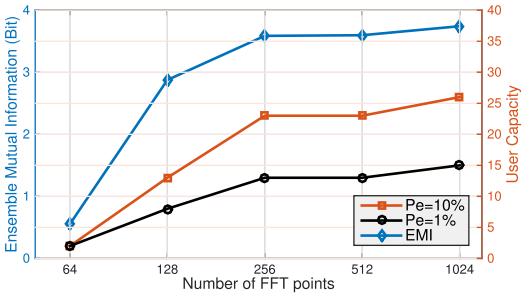


FIGURE 9. User capacity under different numbers of FFT points.

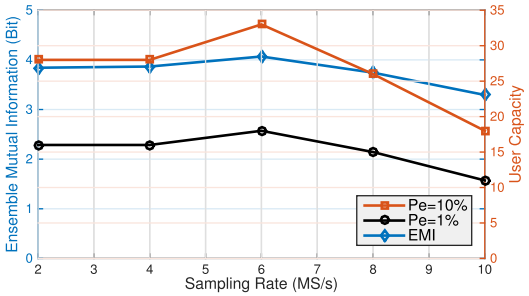


FIGURE 11. User capacity under different sampling rates.

$N_{FFT} = 64 \sim 1024$p. The user capacity results are presented in Fig. 9. From the results, we can see the larger number of FFT points can increase the ensemble mutual information of feature and improve the performance and user capacity. However, when the resolution is accurate to some extend, the improvement of performance is not so significant. Since the increase of FFT points can cause greater computation burden for WPLI, here involves a trade off for system designer.

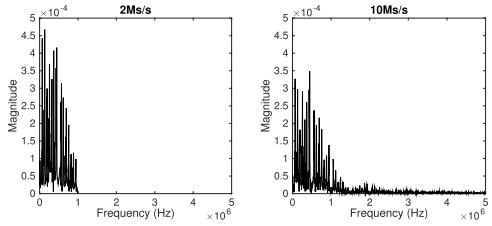


FIGURE 10. Preamble spectrum of signal under different sampling rates.

5) EFFECTS OF SAMPLING RATE

Here we present user capacity due to various sampling rates of receiver with the same frequency resolution, which are the key parameters to determine the bandwidth of the spectral feature $\mathbf{X}$ in equation (3) and consequently reflects its distribution. In Fig. 10, we present the preamble spectrum obtained with two different sampling rates to present difference of spectrum bandwidth. In our previous work [7], we discussed the bandwidth of spectral RFF is $BW = f_s/2$ according to Nyquist sampling rate. As we previous discussion, we use ZigBee protocol, of which baseband symbol rate is 1MS/s. In our case, when sampling rate is $f_s = 2$MS/s, the RFF bandwidth can cover the main lobe of signal PSD. While higher sampling rates can cover more side lobe information of signal PSD, which are more beneficial for WPLI performance. However, the choice of sampling rates also involves a trade off that with the increase of bandwidth, the bandwidth of noise is also increased which can result in the decreasing of signal SNR which worsen the WPLI performance [7]. This phenomenon can be observed in the user capacity characterization and also experimental validations. Here, we still choose a ideal scenario with extremely high SNR to avoid the effects of channel noise. We set the other parameters as, SNR= 29dB for $f_s = 8$Ms/s, $Q = 14$bits. We simulate the result within different sampling rates, $f_s = 2 \sim 10$MS/s, with the same spectrum resolution ($N_{FFT} = 512$p for $f_s = 4$Ms/s), the user capacity results of which are presented in Fig. 11. It can be inferred that when low-noise devices are applied in high SNR scenarios, a higher sampling rate (over-sampling) should be applied for WPLI system to increase user capacity. While, for the low-SNR scenarios, a sampling rate which tightly covers the main lobe of spectrum (Nyquist sampling) should be chosen, in case that more noise rather than signal information are involved to decrease user capacity.

6) EFFECTS OF FADING CHANNELS

Here we discuss the user capacity due to large scale fading path loss fading effects and small scale multipath fading effects including Rayleigh model and Rician model. The Rayleigh model is frequently used to model multipath fading with no direct line-of-sight (LOS) path. The Rician model is often used when there is one strong direct line-of-sight component and many random weaker components. According to our experimental scenario (2.4GHz, indoor), we utilize the typical channel parameters from [53] to characterize user capacity under the fading channel effects.

We firstly choose the typical LOS ($PL_0 = 30$dB, $n = 1.21$, $\epsilon_{rms} = 0.93$, $K = -4.38$dB, $\mu_\tau = 3.624$ns) and NLOS ($n = 2.11$, $\epsilon_{rms} = 1.61$, $\mu_\tau = 26.06$ns) parameters for stimulated channel models. Then we apply the fading channel effects to the collected raw AWGN RFF samples to generate the RFF samples under various fading channel effects. In Fig. 12, the mutual information between each spectral feature component and identity, which are calculated using equation (4), are shown under the fading channels effects at 6 meters away. Comparing with Fig. 3, it is obvious that the fading channel effects significantly decrease the mutual information between spectral components and user identities which would definitely worsen the classification performance of WPLI.

The user capacities over typical fading channels with communication distance ranging from 1 to 10 meters are consequently characterized in Fig. 13. Comparing with the user capacity in Fig. 6, it is clear to see that the fading channel effects can decrease the user capacity more rapidly especially the NLOS Rayleigh channel effects. However, there still exists possibilities to classify a few users under the fading channel effects. in Fig. 13, we can see there still exist a user capacity of 5. That's because among these devices, there

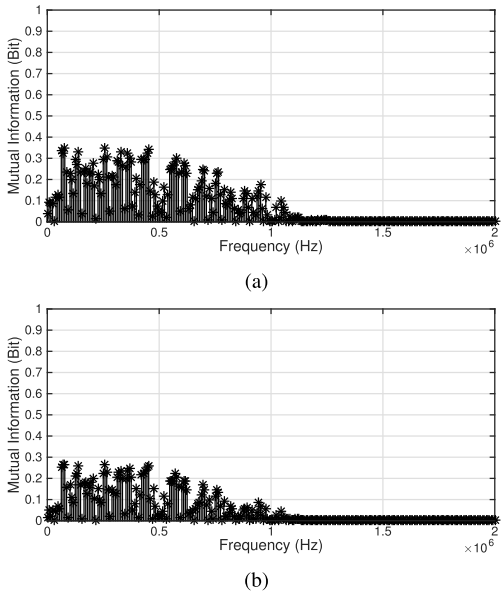


FIGURE 12. Mutual information between spectral feature component and identity, (a) at 6m over Rician channel. (b) at 6m over Rayleigh channel.

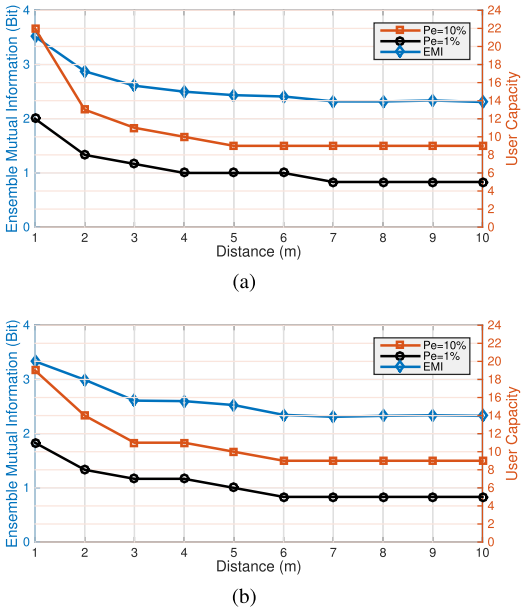


FIGURE 13. User capacity over typical fading channel. (a) over Rician channel. (b) over Rayleigh channel.

are 5 devices have much different genuine features, e.g., transmitting power, battery level, antenna direction, which still remain differently under the same channel model, which makes it possible to distinguish those devices. As the conclusion in previous discussion, with more combined feature selection, extraction and classification techniques, WPLI system can still have chances to work against the fading channel effects. Currently, the existing WPLI systems are vulnerable to fading channel effects. Hence, this direction still needs to be explored for related researchers. Because our theoretical method can find the upper bound on the performance of the designed WPLI system, it can definitely help researchers or system designers to select the most significant RFF feature, compare different algorithms, determine equipment settings, and improve the system designing for target application scenario.

IV. EXPERIMENTAL VALIDATIONS FOR USER CAPACITY

In this section, we utilize the equipment in section III to conduct the conventional in-field experiments of WPLI as the right side of Fig. 2 shows. We present the classification error rates of different user numbers to measure the user capacity using existing experimental approach as the comparison and validation for our theoretical approach. Later we conduct field experiments on classification error performance according to the different case setting in section III to validate the various user capacity characterizations under various real-world constraints.

For classification procedure of WPLI, we select the Fisher LDA [11] as the feature dimensionality reduction technique and the Mahalanobis distance as the distance metric [4], [11], [21]. According to [11], the LDA subspace dimensionality can significantly affect the classification performance. The increase of LDA subspace dimensionality can improve the classification performance to some extends. However, the computation burdens are also significantly raised, which results in a trade off in practical applications. In [13], the typical classification algorithms are compared due to different dimensionality, the results of which show all PCA, LDA and MMI classification converge to the same accuracy if the subspace dimensionality is large enough. Hence, to pursue the optimal classification performance and achieve the upper bound user capacity, we set a large LDA subspace dimensionality $\kappa = 150$ despite the computation time. This LDA subspace dimensionality is quite enough to classify all the features, the larger dimensionality won't improve the classification error more than 0.5%. In each individual time of experiment, each user is assigned with one transmitter to represent his identity. We use newly collected 1000 samples per user to train the LDA training matrix and another 1000 samples per user to be tested for classification respectively, which are extremly large enough to obtain stable results. All these samples are randomly picked from our sample pool collected within a very short period (less than 30s), which the channel can be considered as being stationary.

We present the classification error performance of WPLI with selected user number near the user capacity we obtained. Here we use P_e to denote the classification error rate, λ for the targeted threshold for P_e, N_Y for number of users (one transmitter per user), N_C for the user capacity characterization results of our theoretical approach, and $|$ for the condition symbol. Hence, if the classification error rate is larger than the threshold error rate, i.e., $(P_e|N_Y) > \lambda$, when the user number is larger than the user capacity, i.e., $N_Y > (N_C|\lambda)$, the user capacity is validated. Meanwhile, we also present the classification error performance when user number is near to the user capacity bound, to show the tightness of this bound.

A. USER CAPACITY MEASUREMENT USING EXPERIMENTAL APPROACH

At the beginning of our experiments, we present the results of user capacity measurement without the knowledge of our user capacity characterization methods, but totally relying on an experimental approach, which is similar to existing works. With these results, readers can have a clear idea on the difference, efficiency, and accuracy between our theoretical approach and existing experimental approaches. As the discussion in Section I, the classification performance will get worse with the increasing number of users. Even with the same number of users, the different device combinations can result in different classification error rates. Hence, we design our experimental approach according to [9].

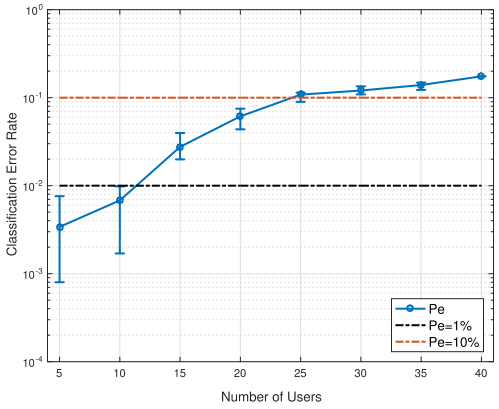


FIGURE 14. User capacity measurement using experimental approach.

We firstly fix our the experimental scenario and system setting measured as, $SNR = 24$dB, $f_s = 4$MS/s, $N_{FFT} = 512$p. Then we increase the number of users N_Y from 5 to 40 with an equal small interval of 5 to obtain the classification error rate for different user numbers. As a result of this approach, the theoretical combinations of users is $C_{40}^{5} + C_{40}^{10} + \ldots + C_{40}^{40} \approx 2.2 \times 10^{11}$, which is a huge number that impossible for implementation in practical. Hence, we calculate 100 random combinations for each user number (except that there's only one combination for 40 users), which is a number large enough to present the characteristics. We show the maximal, mean, and minimal values of the classification error rates, among which, the maximal classification error rates should be used to validate the user capacity. These classification error rates are presented in Fig. 14 with a log scale and in Table 2 with exact values. In Fig. 6, the user capacity for 1% and 10% error rates are characterized as 12 and 22, i.e., $(N_C = 12)|(P_e \leq 1\%)$ and $(N_C = 22)|(P_e \leq 10\%)$. From the results in Fig. 14 and Table 2, the max P_e for 10 classes is 0.98%, i.e., $(P_e = 0.98\%)|(N_Y = 10)$ and $(P_e = 3.97\%)|(N_Y = 15)$. Hence, the user capacity for $P_e \leq 1\%$ should be within 10 to 15 users. Similarly, the user capacity for $P_e \leq 10\%$ should be within 20 to 25 users. Hence, the user capacity results of our theoretical approach are validated also through the measurement results using experimental results.

TABLE 2. Classification error rates of different user numbers.

Experimental setting			
$SNR = 24dB,\ f_s = 4MS/s, N_{FFT} = 512p$			
Practical combinations		Theoretical combinations	
$7 \times 100 + 1$		2.2×10^{11}	
Number of users, N_Y	Mean P_e	Max P_e	Min P_e
5	0.34%	0.76%	0.08%
10	0.68%	0.98%	0.08%
15	2.76%	3.97%	1.99%
20	6.14%	7.50%	4.37%
25	10.83%	11.41%	8.97%
30	12.06%	13.46%	10.89%
35	13.92%	14.85%	12.26%
40	17.48%	17.48%	17.48%

It should be noted that, Fig. 14 is displayed in a log scale, although these starting values are small, however it seems to have larger change range. This phenomena is because when smaller number of transmitters are choosed for classification, more combinations of different devices are validated. Some devices are more similar to each other than others, which make the results change a lot for each time classification. While for the larger number of transmitters, although more classification errors appear because of user capacity, less devices combinations exists, leading to less changes in classification error rate.

The whole process of this experimental approach is quite complicated in operation, time and computation consuming in calculation, and vulnerable to mistakes in data collections. The whole experiments almost took us a whole week to calculate the results by utilizing a high profile workstation with a large memory, while these results are still approximation results with a large reduction of user numbers interval and user combinations. As a comparison, by utilizing our theoretical approach, the user capacity can be accurately characterized after one-time data collection. The user capacity for all users can be calculated in minutes for a common desktop. Hence, our theoretical approach can be a much more efficiency, stable, accurate, and low-cost solution than the user capacity measurement using experimental approach.

B. VALIDATION FOR USER CAPACITY CHARACTERIZATION UNDER REAL-WORLD CONSTRAINTS

After the comparison our theoretical approach with conventional experimental approach, we conduct field experiments on classification error performance according to the different case setting in section III. These experimental results can validate the user capacity characterization under various real-world constraints in section III. Meanwhile the influence of different practical constraints can be analyzed through these experiment results. The only exception here, is the case of ADC quantization bits, because the ADC quantization bits is impossible to change for given hardware setting, we can only fix the ADC quantization bits to Q=14bit according

to USRP daughter board setting. The classification results are shown in the following figures where the x-axis is the number of test samples, the y-axis is the minimal distance score between test sample and its reference, and the z-axis is the user/class identity number assigned to the test samples. Besides the color of each sample is to present its true identity which can help the reader to compare classified identities of test samples.

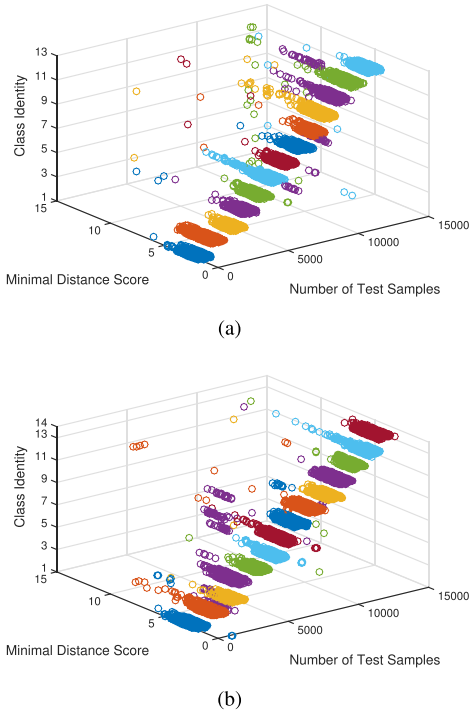


FIGURE 15. SNR=26dB, $f_s = 4MS/s$, $N_{FFT} = 512p$. (a) Classification results of 13 users ($P_e = 0.91\%$). (b) Classification results of 14 users ($P_e = 1.09\%$).

1) EFFECTS OF AWGN NOISE LEVEL

Here, we use the experiment results to validate the user capacity characterization for AWGN noise level case. As the case setting in Fig. 6, we conduct the experiments at SNR=26dB, where the 1% error rate user capacity is 13, i.e., $(N_C = 13)|(P_e \leq 1\%)$. In Fig. 15(a), classification results of 13 users, are shown of which the classification error rate is $(P_e = 0.91\%)|(N_Y = 13)$. While in Fig. 15(b), the classification results of 14 users are shown of which the classification error rate is $(P_e = 1.09\%)|(N_Y = 14)$. Hence, we can see the user capacity characterization at this point is validated accurately. Similarly, we change SNR situation and threshold error rate to see validate another point of our user capacity curves. In Fig. 6, when SNR=22dB, the user capacity is 12 with 1% error rate, i.e., $(N_C = 12)|(P_e \leq 1\%)$. In Fig. 16(a), classification results of 11 users, are shown of which $(P_e = 0.75\%)|(N_Y = 11)$. While in Fig. 16(b), the classification results of 12 users are shown of which $(P_e = 1.79\%)|(N_Y = 12)$. The user capacity characterization

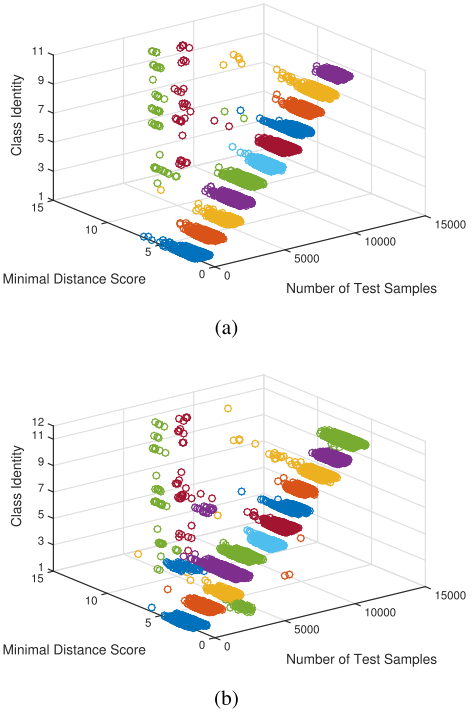


FIGURE 16. SNR=22dB, $f_s = 4MS/s$, $N_{FFT} = 512p$. (a) Classification results of 11 users ($P_e = 0.75\%$). (b) Classification results of 12 users ($P_e = 1.79\%$).

at this case is slightly larger than the experimental result. As we discuss in the previous section, with the decrease of SNR level, single classifier and single feature selection are not enough achieve the upper bound user capacity. Hence, more advanced or combined feature selection and classification techniques should be utilized in WPLI.

2) EFFECTS OF NUMBER OF FFT POINTS

We use the experiment results to validate the user capacity characterization for number of FFT points case. We conduct the experiments as the case setting for Fig. 9. As the user capacity characterization under $N_{FFT} = 64p$ is, $(N_C = 2)|(P_e \leq 10\%)$. In Fig. 17(a), the classification results of 3 users are shown with $(P_e = 15.67\%)|(N_Y = 3)$ which is out of user capacity. For $N_{FFT} = 256$p, $(N_C = 13)|(P_e \leq 1\%)$. In Fig. 17(b), the classification results of 14 users are shown with $(P_e = 1.25\%)|(N_Y = 14)$ which is still out of user capacity. For $N_{FFT} = 1024$p, $(N_C = 15)|(P_e \leq 1\%)$. In Fig. 17(c), the classification results of 15 users are shown with $(P_e = 0.47\%)|(N_Y = 15)$ which is within user capacity. Hence, the experimental results match the discussion in previous section very well.

3) EFFECTS OF SAMPLING RATE

We use the experiment results to validate the user capacity characterization for sampling rate case. We conduct the experiments as the case setting for Fig. 11. As the user capacity characterization under $f_s = 2$MS/s is, $(N_C = 16)|(P_e \leq 1\%)$. In Fig. 18(a), the classification results

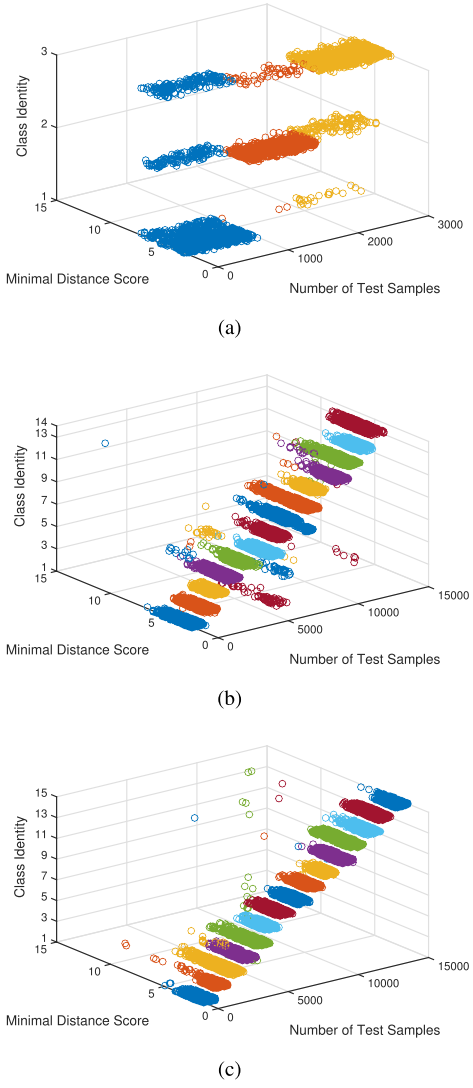


FIGURE 17. SNR=29dB , $f_s = 8MS/s$. (a) Classification results of 3 users, $N_{FFT} = 64$ ($P_e = 15.67\%$). (b) Classification results of 14 users, $N_{FFT} = 128$ ($P_e = 1.25\%$). (c) Classification results for 15 users, $N_{FFT} = 1024$ ($P_e = 0.47\%$).

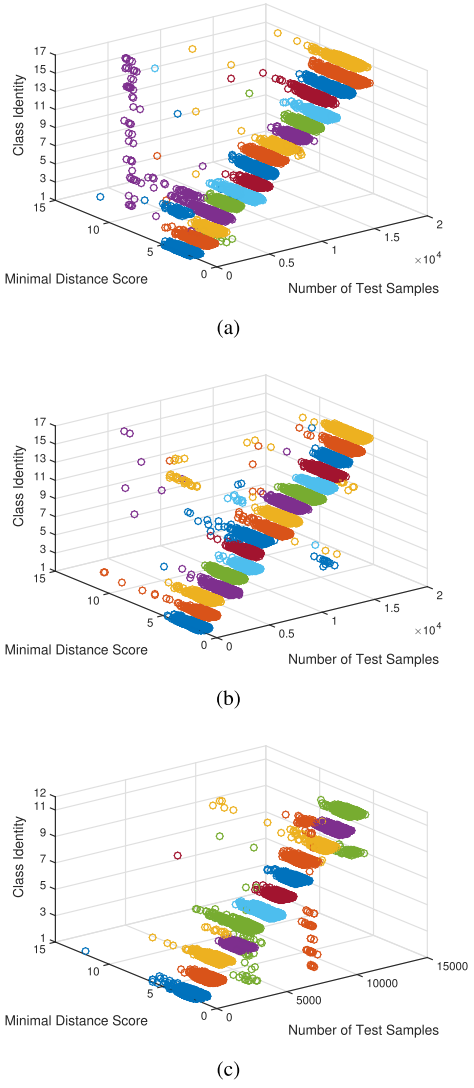


FIGURE 18. SNR=29dB (8MS/s). (a) Classification results of 17 users, $f_s = 2$MS/s ($P_e = 1.17\%$). (b) Classification results of 17 users, $f_s = 6$MS/s ($P_e = 0.58\%$). (c) Classification results for 12 users, $f_s = 10$MS/s ($P_e = 1.51\%$).

of 17 users are shown with $(P_e = 1.17\%)|(N_Y = 17)$ which is out of user capacity. For $f_s = 6$MS/s, $(N_C = 17)|(P_e \leq 1\%)$. In Fig. 18(b), the classification results of 17 users are shown with $(P_e = 0.58\%)|(N_Y = 17)$ which is within the capacity. For $f_s = 10$MS/s, $(N_C = 11)|(P_e \leq 1\%)$. In Fig. 18(c), the classification results of 12 users under $f_s = 10$MS/s are shown with $(P_e = 1.51\%)|(N_Y = 12)$, which is out of the capacity. The analyses in our previous discussion is also validated.

4) EFFECTS OF FADING CHANNELS

We use the experiment results to validate the user capacity characterization for fading channel effects. Due to the uniqueness of channel characteristics, the communication channel cannot be precisely rebuild as the parameters in section III.

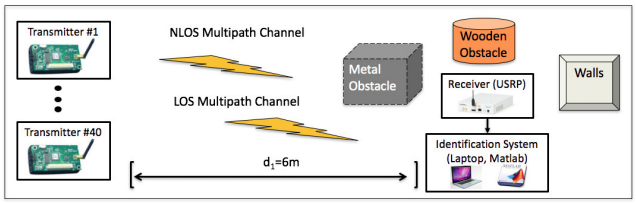
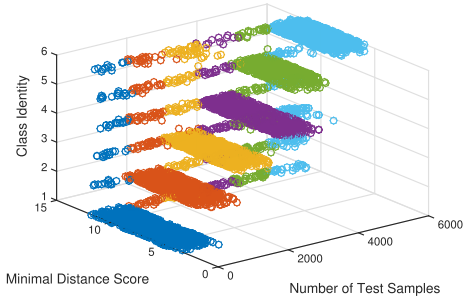
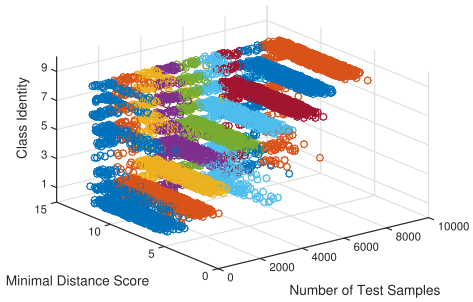


FIGURE 19. Experimental scenarios for fading channel effects.

We try to find the same indoor corridor scenario in our labs and to use the metal obstacles to create multipath channel with LOS or NLOS signal as Fig. 19. We newly collect the RFF samples from different nodes in these scenarios to check the classification performance to validate the user capacity characterized with typical fading channel parameters

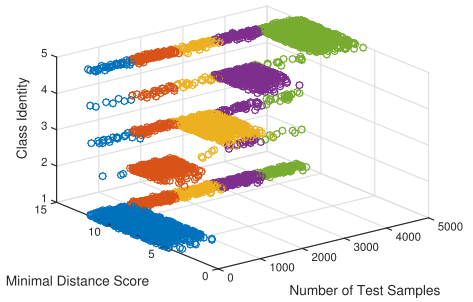


(a)

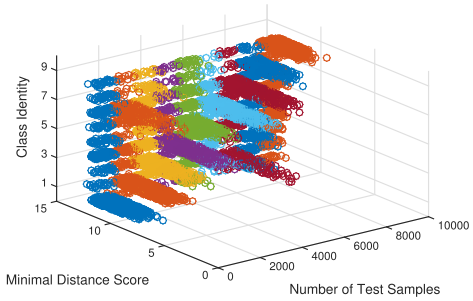


(b)

FIGURE 20. LOS channel, 6 meters away, $f_s = 4$MS/s. (a) Classification results of 6 users ($P_e = 13.73\%$). (b) Classification results of 9 users ($P_e = 23.79\%$).



(a)



(b)

FIGURE 21. NLOS channel, 6 meters away, $f_s = 4$MS/s. (a) Classification results of 5 users ($P_e = 28.54\%$). (b) Classification results of 9 users ($P_e = 35.09\%$).

in section III. In Fig. 13(a), the user capacity is $(N_C = 6)|(P_e \leq 1\%)$ and $(N_C = 9)|(P_e \leq 10\%)$. In Fig. 20, the classification results of 6 users and 9 users at 6m LOS channel scenario are shown, of which the error rates are $(P_e = 13.73\%)|(N_Y = 6)$ and $(P_e = 23.79\%)|(N_Y = 9)$ respectively. In Fig. 13(b), the user capacity is $(N_C = 5)|(P_e \leq 1\%)$ and $(N_C = 9)|(P_e \leq 10\%)$. In Fig. 21, the classification results of 5 users and 9 users at 6m NLOS channel scenario are shown, of which the error rates are $(P_e = 28.54\%)|(N_Y = 5)$ and $(P_e = 35.09\%)|(N_Y = 9)$ respectively. From these results, it is clear to see that the fading channel effects can significantly worsen the classification performance of WPLI. Moreover, using single feature selection and one classification technique only is hard to achieve the user capacity under the fading channel effects. Hence, if some novel approaches or algorithms can be proposed for WPLI to work against the fading channel effects, that would be a great break through in this research direction.

V. CONCLUSION

In this work, we establish a theoretical understanding on user capacity of Wireless Physical-layer Identification (WPLI). We propose a theoretical approach to characterize user capacity based on mutual information between RFF feature and user identity. The user capacity of WPLI are consequently characterized within targeted error performance, under various real-world constraints, despite different classification techniques and user combinations. Various experiments of a practical WPLI system are conducted to measure and validate the user capacity characterization. The classification results of experiments show great accuracy and tightness of user capacity characterization. By applying this approach with one-time collected training RFF samples, lots of inefficient experiments to measure user capacity can be avoided, which saves great efforts for system-designers. With the understanding of user capacity, significant RFF features can be selected to achieve the largest user capacity for a WPLI system. Different classification algorithms can be evaluated by comparing with the guaranteed accuracy of user capacity. The effects of various real-world constraints on the performance of WPLI can be analyzed by checking the changes in user capacity. Hence, this user capacity can be widely utilized as an important metric to evaluate the performance of real-world WPLI applications.

REFERENCES

[1] W. Wang, Z. Sun, K. Ren, and B. Zhu, "User capacity of wireless physical-layer identification: An information-theoretic perspective," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.
[2] B. Danev, D. Zanetti, and S. Čapkun, "On physical-layer identification of wireless devices," *ACM Comput. Surv.*, vol. 45, no. 1, p. 6, 2012.
[3] Q. Xu, R. Zheng, W. Saad, and Z. Han, "Device fingerprinting in wireless networks: Challenges and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 94–104, 1st Quart., 2016.
[4] B. Danev, T. S. Heydt-Benjamin, and S. Čapkun, "Physical-layer identification of RFID devices," in *Proc. USENIX Secur. Symp.*, 2009, pp. 199–214.
[5] H. J. Patel, M. A. Temple, and R. O. Baldwin, "Improving ZigBee device network authentication using ensemble decision tree classifiers with radio frequency distinct native attribute fingerprinting," *IEEE Trans. Rel.*, vol. 64, no. 1, pp. 221–233, Mar. 2015.
[6] H. J. Patel and B. W. Ramsey, "Comparison of parametric and non-parametric statistical features for Z-wave fingerprinting," in *Proc. IEEE Military Commun. Conf. (MILCOM)*, Oct. 2015, pp. 378–382.

[7] W. Wang, Z. Sun, S. Piao, B. Zhu, and K. Ren, "Wireless physical-layer identification: Modeling and validation," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 9, pp. 2091–2106, Sep. 2016.
[8] Y. Huang and H. Zheng, "Theoretical performance analysis of radio frequency fingerprinting under receiver distortions," *Wireless Commun. Mobile Comput.*, vol. 15, no. 5, pp. 823–833, 2015.
[9] L. Yang, P. Peng, F. Dang, C. Wang, X.-Y. Li, and Y. Liu, "Anti-counterfeiting via federated RFID tags' fingerprints and geometric relationships," in *Proc. INFOCOM*, 2015, pp. 1966–1974.
[10] P. Scanlon, I. O. Kennedy, and Y. Liu, "Feature extraction approaches to RF fingerprinting for device identification in femtocells," *Bell Labs Tech. J.*, vol. 15, no. 3, pp. 141–151, Dec. 2010.
[11] B. Danev and S. Čapkun, "Transient-based identification of wireless sensor nodes," in *Proc. Int. Conf. Inf. Process. Sensor Netw.*, 2009, pp. 25–36.
[12] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, "Wireless device identification with radiometric signatures," in *Proc. 14th ACM Int. Conf. Mobile Comput. Netw.*, 2008, pp. 116–127.
[13] K. Torkkola and W. M. Campbell, "Mutual information in learning feature transformations," in *Proc. ICML*, 2000, pp. 1015–1022.
[14] G. Brown, "An information theoretic perspective on multiple classifier systems," in *Multiple Classifier Systems*. Berlin, Germany: Springer, 2009, pp. 344–353.
[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2012.
[16] O. Gungor, C. E. Koksal, and H. El Gamal, "An information theoretic approach to RF fingerprinting," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, 2013, pp. 61–65.
[17] O. Gungor and C. E. Koksal, "On the basic limits of RF-fingerprint-based authentication," *IEEE Trans. Inf. Theory*, vol. 62, no. 8, pp. 4523–4543, Aug. 2016.
[18] O. Gungor and C. E. Koksal, "RF-fingerprint based authentication: Exponents and achievable rates," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Oct. 2014, pp. 97–102.
[19] E. Jorswieck, S. Tomasin, and A. Sezgin, "Broadcasting into the uncertainty: Authentication and confidentiality by physical-layer processing," *Proc. IEEE*, vol. 103, no. 10, pp. 1702–1724, Oct. 2015.
[20] 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, Mar. 2010.
[21] D. Zanetti, B. Danev, and S. Čapkun, "Physical-layer identification of UHF RFID tags," in *Proc. 16th Annu. Int. Conf. Mobile Comput. Netw.*, 2010, pp. 353–364.
[22] A. C. Polak, C. Dolatshahi, and D. L. Goeckel, "Identifying wireless users via transmitter imperfections," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 7, pp. 1469–1479, Aug. 2011.
[23] M.-W. Liu and J. F. Doherty, "Specific emitter identification using nonlinear device estimation," in *Proc. IEEE Sarnoff Symp.*, Apr. 2008, pp. 1–5.
[24] A. C. Polak and D. L. Goeckel, "RF fingerprinting of users who actively mask their identities with artificial distortion," in *Proc. Conf. Rec. 45th Asilomar Conf. Signals, Syst. Comput. (ASILOMAR)*, 2011, pp. 270–274.
[25] J. Toonstra and W. Kinsner, "Transient analysis and genetic algorithms for classification," in *Proc. IEEE WESCANEX, Commun., Power, Comput. Conf.*, vol. 2. May 1995, pp. 432–437.
[26] J. Toonstra and W. Kinsner, "A radio transmitter fingerprinting system ODO-1," in *Proc. IEEE Can. Conf. Elect. Comput. Eng.*, vol. 1. May 1996, pp. 60–63.
[27] Z. Li, W. Xu, R. Miller, and W. Trappe, "Securing wireless systems via lower layer enforcements," in *Proc. 5th ACM Workshop Wireless Secur.*, 2006, pp. 33–42.
[28] L. Xiao, L. Greenstein, N. Mandayam, and W. Trappe, "Fingerprints in the ether: Using the physical layer for wireless authentication," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2007, pp. 4646–4651.
[29] H. Liu, Y. Wang, J. Liu, J. Yang, and Y. Chen, "Practical user authentication leveraging channel state information (CSI)," in *Proc. 9th ACM Symp. Inf., Comput. Commun. Secur.*, 2014, pp. 389–400.
[30] M. Edman and B. Yener, "Active attacks against modulation-based radiometric identification," Dept. Comput. Sci., RPI, Troy, NY, USA, Tech. Rep. 02-09, 2009.
[31] W. C. Suski, II, M. A. Temple, M. J. Mendenhall, and R. F. Mills, "Using spectral fingerprints to improve wireless network security," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov./Dec. 2008, pp. 1–5.
[32] R. Klein, M. A. Temple, M. J. Mendenhall, and D. R. Reising, "Sensitivity analysis of burst detection and RF fingerprinting classification performance," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2009, pp. 1–5.
[33] S. Ur Rehman, K. Sowerby, C. Coghill, and W. Holmes, "The analysis of RF fingerprinting for low-end wireless receivers with application to IEEE 802.11a," in *Proc. Int. Conf. Sel. Topics Mobile Wireless Netw. (iCOST)*, 2012, pp. 24–29.
[34] S. Ur Rehman, K. Sowerby, and C. Coghill, "RF fingerprint extraction from the energy envelope of an instantaneous transient signal," in *Proc. Austral. Commun. Theory Workshop (AusCTW)*, 2012, pp. 90–95.
[35] N. Hu and Y.-D. Yao, "Identification of legacy radios in a cognitive radio network using a radio frequency fingerprinting based method," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 1597–1602.
[36] J. Hall, M. Barbeau, and E. Kranakis, "Enhancing intrusion detection in wireless networks using radio frequency fingerprinting," in *Proc. 3rd IASTED Int. Conf. Commun. Internet Inf. Technol. (CIIT)*, 2004, pp. 201–206.
[37] M. Barbeau, J. Hall, and E. Kranakis, "Detecting rogue devices in Bluetooth networks using radio frequency fingerprinting," in *Proc. 3rd IASTED Int. Conf. Commun. Comput. Netw. (CCN)*, 2006, pp. 4–6.
[38] A. C. Polak and D. L. Goeckel, "Wireless device identification based on RF oscillator imperfections," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 12, pp. 2492–2501, Dec. 2015.
[39] G. Verma, P. Yu, and B. M. Sadler, "Physical layer authentication via fingerprint embedding using software-defined radios," *IEEE Access*, vol. 3, pp. 81–88, 2015.
[40] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels*, vol. 95. Hoboken, NJ, USA: Wiley, 2005.
[41] T. S. Rappaport *et al.*, *Wireless Communications: Principles and Practice*, vol. 2. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996.
[42] T. J. Bihl, K. W. Bauer, and M. A. Temple, "Feature selection for RF fingerprinting with multiple discriminant analysis and using ZigBee device emissions," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 8, pp. 1862–1874, Aug. 2016.
[43] J. Hasse, T. Gloe, and M. Beck, "Forensic identification of GSM mobile phones," in *Proc. 1st ACM Workshop Inf. Hiding Multimedia Secur.*, 2013, pp. 131–140.
[44] U. Ozertem, D. Erdogmus, and I. Santamaria, "Detection of nonlinearly distorted signals using mutual information," in *Proc. 13th Eur. Signal Process. Conf. (EUSIPCO)*, 2005, pp. 1–4.
[45] U. Ozertem, D. Erdogmus, and R. Jenssen, "Spectral feature projections that maximize Shannon mutual information with class labels," *Pattern Recognit.*, vol. 39, no. 7, pp. 1241–1252, 2006.
[46] Z.-H. Zhou, *Ensemble Methods: Foundations and Algorithms*. Boca Raton, FL, USA: CRC Press, 2012.
[47] G. Fasshauer and M. McCourt, *Kernel-Based Approximation Methods Using MATLAB*. New, NY, USA: USA: World Scientific Publishing, 2015.
[48] Z.-H. Zhou and N. Li, "Multi-information ensemble diversity," in *Multiple Classifier Systems*. Springer, 2010, pp. 134–144.
[49] *M. Datasheet*, Crossbow Technol. Inc., San Jose, CA, USA, 2006.
[50] *Imote2: High-Performance Wireless Sensor Network Node*, accessed on Mar. 08, 2017. [Online]. Available: http://wsn.cse.wustl.edu/images/e/e3/Imote2_Datasheet.pdf
[51] *T. Datasheet*, Crossbow Technol. Inc., Milpitas, CA, USA, 2013.
[52] M. Ettus. (2008). Universal software radio peripheral (USRP). Ettus Research LLC. [Online]. Available: http://www.ettus.com
[53] R. de Francisco, "Indoor channel measurements and models at 2.4 GHz in a hospital," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2010, pp. 1–6.

WENHAO WANG (S'16) received the B.S. degree in communication engineering from Southwest Jiaotong University, Chengdu, China, in 2011. He is currently pursuing the Ph.D. degree, under the supervision of Prof. B. Zhu, with the School of Electronics Engineering and Computer Science, Peking University, Beijing, China. He is also a joint-training Ph.D. Student, with Prof. Z. Sun, at the Department of Electrical Engineering, The State University of New York, Buffalo, NY, USA. He was a recipient of the Scholarship from China Scholarship Council for his study in USA in 2014. His research interests include wireless communication, physical-layer security, information-theoretic security, and radio frequency fingerprints.

ZHI SUN (M'11) received the B.S. degree in telecommunication engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2004, and the M.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 2007, and the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2011. He was a Post-Doctoral Fellow with the Georgia Institute of Technology. He is currently an Assistant Professor with the Department of Electrical Engineering, The State University of New York, Buffalo, NY, USA. His research interests include wireless communication and networking in extreme environments, metamaterial enhanced communication and security, physical-layer security, wireless intra-body networks, wireless underground networks, wireless underwater networks, and cyber physical systems. He was a recipient of the NSF CAREER Award in 2017, the Best Paper Award in the IEEE Global Communications Conference in 2010, the BWN Researcher of the Year Award at the Georgia Institute of Technology in 2009, and the Outstanding Graduate Award at Tsinghua University in 2007. He currently serves as an Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

KUI REN (M'07–SM'11–F'16) received the Ph.D. degree from Worcester Polytechnic Institute. He is currently a Professor of Computer Science and Engineering and the Director of the UbiSeC Laboratory, The State University of New York. His current research interest spans cloud and outsourcing security, wireless and wearable systems security, and mobile sensing and crowdsourcing. He has authored 150 peer-reviewed journal and conference papers. His research has been supported by NSF, DoE, AFRL, MSR, and Amazon. He was a recipient of the UB Exceptional Scholar Award for Sustained Achievement in 2016, the SEAS Senior Researcher of the Year in 2015, the Sigma Xi/IIT Research Excellence Award in 2012, and the NSF CAREER Award in 2011. He received several Best Paper Awards, including the IEEE ICNP 2011. He currently serves as an Associate Editor of the IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, the IEEE TRANSACTIONS ON MOBILE COMPUTING, the IEEE WIRELESS COMMUNICATIONS, and the IEEE INTERNET OF THINGS JOURNAL. He is a Distinguished Lecturer of the IEEE, a member of the ACM, and a Past Board Member of the Internet Privacy Task Force, State of Illinois.

BOCHENG ZHU received the Ph.D. degree in electromagnetic field and microwave technology from the Beijing Institute of Technology, Beijing, China, in 1996. He is currently a Professor with the School of Electronics Engineering and Computer Science, Peking University, Beijing, China. He is also a Senior Member with the Chinese Institute of Electronics. He is also a Senior Reviewer of the National High Technology Research and Development Program (863 program) and the National Science and Technology Major Project of China. His research interests include wireless communication, satellite navigation, and microwave technology.