2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)

WiMi: Target Material Identification with Commodity Wi-Fi Devices

Chao Feng[†♯], Jie Xiong[‡], Liqiong Chang[†], Ju Wang[§], Xiaojiang Chen[†♯], Dingyi Fang[†], Zhanyong Tang[†]
[†]*Northwest University,*[‡]*University of Massachusetts Amherst,*[§]*University of Waterloo*
[♯]*Shaanxi International Joint Research Centre for the Battery-free Internet of things*

***Abstract*—Target material identification is playing an important role in our everyday life. Traditional camera and video-based methods bring in severe privacy concerns. In the last few years, while RF signals have been exploited for localization, activity tracking and even respiration, very little attention has been paid in material identification. This paper introduces WiMi, a non-contact target material identification system, implemented on ubiquitous and cheap commercial off-the-shelf (COTS) Wi-Fi devices. The intuition is that different materials produce different amounts of phase and amplitude changes when a target appears on the line-of-sight (LoS) of a radio frequency (RF) link. However, due to multipath and hardware imperfection, the measured phase and amplitude of the channel state information (CSI) are very noisy. We thus present CSI pre-processing schemes to address the multipath and hardware noise issues before they can be used for material sensing. We also design a new material feature which is only related to the material type and is independent of the target size. Comprehensive real-life experiments demonstrate that WiMi can achieve fine-grained material identification with cheap commodity Wi-Fi devices. WiMi can identify 10 commonly seen liquids at an overall accuracy higher than 95% with strong multipath indoors. Even for very similar items such as Pepsi and Coke, WiMi can still differentiate them at a high accuracy.**

***Keywords*-Wi-Fi; CSI; Material Identification, contactless;**

I. INTRODUCTION

Target material identification plays an important role in many IoT applications. At the airport, explosive detection is critical at the security checkpoint [1]. With material identification, a robot can smartly reduce its strength when picking up a fragile item such as an egg [2]. Fine-grained material identification can even be applied to differentiate very similar items such as Pepsi and Coke without a taste [3]. With fine-grained material sensing, expired liquid such as milk can be detected without requiring to open the bottle or taste it.

Existing material identification systems such as Radar [4], X-Ray [5], CT/MRI [6] and B-scan ultrasonography [7] utilize dedicated hardware to emit extremely high frequency signals. Although they can provide a high identification accuracy, these systems are usually big in size and expensive to be used in home and office environments. Camera based material identification techniques have been proposed [8, 9]. However, camera based approaches rely on good lighting conditions and bring in severe privacy concerns. Recently, we have witnessed a new trend of employing radio frequency (RF) signals such as Wi-Fi and millimeter wave signals for localization [10–12], gesture recognition [13, 14], and motion tracking [15]. Wi-Fi infrastructure is already ubiquitously deployed around us while millimeter wave may be employed in the next generation 802.11 Wi-Fi standard. However, we notice that very little attention has been paid to material identification with RF signals.

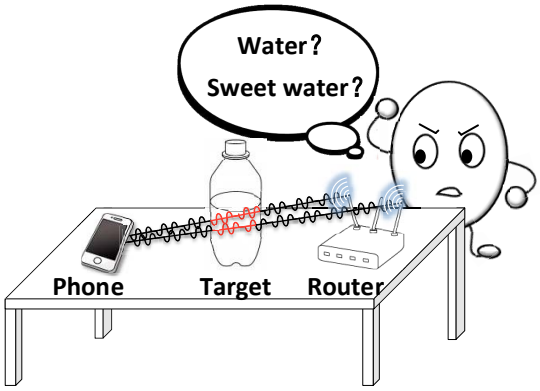


Figure 1: Example application of WiMi.

Recently, two material sensing systems Tagscan and liuqID [3, 16] are proposed to explore the possibility of employing RFID and UWB signals for material identification, achieving high accuracies. However, Tagscan needs an expensive RFID reader (Impinj R420 reader is about 1500 USD) to identify target material, and LiuqID requires UWB kits (the price is about 300 USD [16]) to perform material identification by synchronizing the transmitter and receiver with a wire connection. Compared to RFID and UWB devices, we believe Wi-Fi is a more promising candidate as it is ubiquitous and no any dedicated infrastructure is needed. Therefore, in this work, we ask the question: *can we accurately identify the material type of a target with cheap commodity Wi-Fi devices* ? If this is feasible, as shown in Fig. 1, we may able to employ just a smartphone in the future to sense target material wherever Wi-Fi exists.

In this paper, we introduce WiMi, a commodity Wi-Fi based contactless target material identification system. To identify the target material, the key observation is that different materials cause different amounts of phase and amplitude changes when wireless signal penetrates through [3]. However, different from the previous method which employs

2575-8411/19/$31.00 ©2019 IEEE
DOI 10.1109/ICDCS.2019.00074

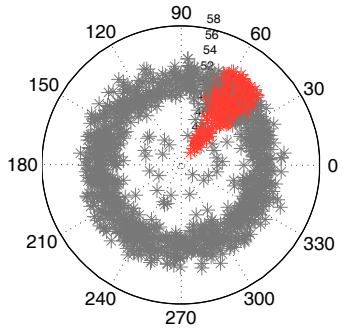


Figure 2: The raw CSI phase values vs. Phase difference values between antennas.

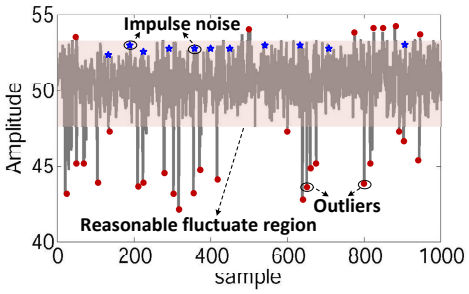


Figure 3: The raw CSI amplitude values.

directional RFID transmissions for material identification, commodity Wi-Fi employs omni-directional antennas and thus multipath effect is more severe. Furthermore, both signal phase and amplitude readings from commodity Wi-Fi devices are very coarse and noisy, which make them not directly suitable for fine-grained material identification. Two main challenges need to be solved before we can perform accurate material sensing with commodity Wi-Fi devices.

The first challenge is that the raw channel state information (CSI) readings retrieved from Wi-Fi hardware are coarse and noisy. We discuss about the CSI phase and amplitude readings separately. Specifically, the raw phase readings are corrupted by asynchronization between the transceivers, hardware noise and multipath [17, 18]. As shown in Fig. 2, the measured CSI phase changes at one subcarrier (marked as grey dots) of a Wi-Fi channel from the Wi-Fi card are very random across different packets. For a same material, if we measure the phase readings multiple times, the variations are large. To deal with this challenge, we observe that the phase difference between the CSI readings obtained from two closely-placed antennas are much more stable and multiple antennas are widely available at commodity Wi-Fi access points. Hence, we propose to utilize a second antenna to extract CSI phase difference for material identification. We further observe that among the 64 subcarriers of a 20MHz Wi-Fi channel, some subcarriers may be greatly affected by multipath while the rest may not due to frequency diversity [12]. Specifically, the phase difference at a 'clean'[1] subcarrier across multiple subsequent packets in the time domain are more stable and have a smaller variance.

Besides the phase readings, the amplitude readings at one subcarrier are also very noisy as shown in Fig. 3. To extract the accurate CSI amplitudes, we observe that the raw CSI amplitudes contain substantial outliers and impulse noise. In particular, the impulse noises occur irregularly and instantaneously, and are usually comparative to the useful signals. To deal with this challenge, we first remove the outliers by setting a signal fluctuation threshold. Then based on the observation that the signals are highly-correlated but the noises are weakly-correlated (or uncorrelated) at different frequencies as shown in Sec. III-C, we can remove the impulse noise and reconstruct the useful signals by integrating the correlated signals. Even with the above steps, there are still signal noise left, we thus employ the second antenna again to obtain the amplitude ratio which is more stable than the amplitude readings from each individual antenna.

With the schemes described in the previous session, we obtain clean and stable phase difference and amplitude ratio information. The second challenge is how to utilize these values to design a material feature which is unique for each material and at the same time independent of target size. If we are not able to obtain a unique feature independent of the target size, then the same target with different sizes will be identified as different materials. Inspired by the method proposed in [3], we design a material feature which is only related to the material type but not the target size with just the phase difference and amplitude ratio information. Different from [3], our proposed feature works with multi-antenna system which is common for 802.11n/ac Wi-Fi APs.

We implement our prototype with commodity Wi-Fi devices equipped with multiple antennas, and evaluate the system performance in three different indoor environments: a library, a lab (office) and an empty hall corresponding to high, medium and low multipath environments. Extensive experiments demonstrate that WiMi is able to identify 10 commonly seen liquid materials at higher than 95% average accuracy in all the three environments.

Contributions: The main contributions are as follows:

- To our best knowledge, WiMi is the first Wi-Fi-based material identification system hosted on commodity Wi-Fi devices. WiMi employs the CSI phase difference and amplitude ratio to obtain the unique *material feature* independent of the target size for material identification.
- We propose multiple signal processing schemes to remove the noise to improve the material identification accuracy.

[1]Here 'clean' means the subcarrier is not affected by multipath

- We propose a material feature working with multi-antenna Wi-Fi hardware to achieve fine-grained material identification.
- We design and implement WiMi on commodity Wi-Fi devices. Extensive experiments demonstrate the effectiveness and robustness of WiMi. WiMi is able to differentiate very similar items such as Pepsi and Coke at higher than 90% accuracy.

II. Preliminary

In this section, we first briefly introduce the background information. Then we present the key challenges.

A. Channel State Information

802.11 Wi-Fi physical layer employs OFDM scheme to transmit data across orthogonal subcarriers [19]. For each received packet, we can extract the Channel State Information (CSI) at each subcarrier, which represents the channel characteristic of the communication link. The channel frequency responses can be expressed as:

$$H = (H(f_1), H(f_2) \cdots H(f_K)), \tag{1}$$

where $H(f_k)$ is the channel frequency response for subcarrier k, and K is the number of subcarriers. $H(f_k) = \|H(f_k)\| e^{j \cdot \angle H(f_k)}$ is represented by the amplitude $\|H(f_k)\|$ and the phase $\angle H(f_k)$. The CSI measurements at $K = 30$ subcarriers can be exported from the COTS Intel 5300 Wi-Fi card with a publicly available tool [20].

B. Phase and Amplitude Change When A Wi-Fi Signal Penetrates Through The Target

The *CSI phase change* is the difference between two phase readings, which are measured before and after a target appears at the LoS link. The wavelengths of the Wi-Fi signal are different in different materials. Thus, as Fig.4 shows, with a same transmission distance in the target, the phase changes will be different when the signal penetrates through different target materials of the same size. Let ϕ_{tar} and ϕ_{free} be the phase readings when there is a target and when there is no target, the phase change $\Delta\phi$ can now be calculated as:

$$\begin{aligned}\Delta\phi &= \phi_{tar} - \phi_{free} \\ &= \left[2\pi \cdot \frac{(L-D)}{\lambda_{free}} + 2\pi \cdot \frac{D}{\lambda_{tar}} - 2\pi \cdot \frac{L}{\lambda_{free}}\right] \\ &= D \cdot 2\pi \left(\frac{1}{\lambda_{tar}} - \frac{1}{\lambda_{free}}\right),\end{aligned} \tag{2}$$

where λ_{tar} and λ_{free} are the wavelengths in the target and in the air, respectively. L is the distance between the transmitter and the receiver and D is the path length inside the target. We define $\beta_{tar} = \frac{2\pi}{\lambda_{tar}}$ and $\beta_{free} = \frac{2\pi}{\lambda_{free}}$ as the signal phase constant in the target and in the air, respectively. Then, Equation (2) can be simplified as:

$$\Delta\phi = D(\beta_{tar} - \beta_{free}) \tag{3}$$

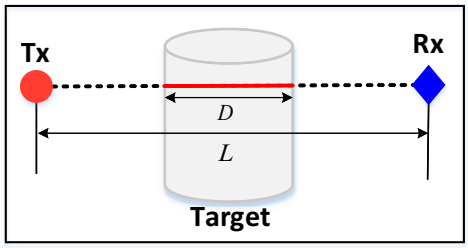


Figure 4: The brief illustration for a signal travels through target.

The *CSI amplitude change* is the difference between the two amplitude readings measured before and after a target appears at the LoS link. For different target materials, the amplitude changes are also different. Generally, the amplitude has an $e^{-\alpha}$ attenuation over a unit prorogation distance inside the target, where α is the attenuation constant related to the target material. We define the measured signal amplitudes before and after the target appears on the RF link as A_{free} and A_{tar}, we can obtain the RSS change (ratio) in dB as:

$$\begin{aligned}\Delta R &= 20 \times \log \frac{A_{tar}}{A_{free}} \\ &= 20 \times \log \frac{A_s e^{-\alpha_{free}(L-D)} e^{-\alpha_{tar} D}}{A_s e^{-\alpha_{free} L}} \\ &= 20 \times \log e^{-D(\alpha_{tar} - \alpha_{free})},\end{aligned} \tag{4}$$

where A_s denotes the amplitude of the original transmitted signal. α_{tar} and α_{free} represent the signal attenuation constants in the target and in the air, respectively.

C. Challenges and Verifications

In this part, we discuss the three challenges we need to tackle before our method can work.

1) Unstable Phase Values: Accurate phase readings are critical to our material identification method. Unlike RFID devices [3] which can provide fine-grained phase readings, the phase information retrieved from the commodity Wi-Fi devices are very unstable because of the non-synchronization between transceivers, hardware noise, and rich multipath in the indoor environment. To explore the distribution of raw phase measurements, we conduct an experiment in a lab office environment with one transmitter and one receiver spaced at a distance of one meter. As shown in Fig. 2, the raw phases are randomly distributed from 0 to 2π. Hence, we need to carefully process the raw phase readings before then can be employed for material sensing.

2) Noisy Amplitude Values: WiMi leverages the CSI amplitude information to establish the relationship with the propagation path distance, as shown in Equation (4). Thus, an accurate amplitude reading is also important in our proposed scheme. In real deployment setups, there exist substantial outliers and impulse noise in the CSI amplitude

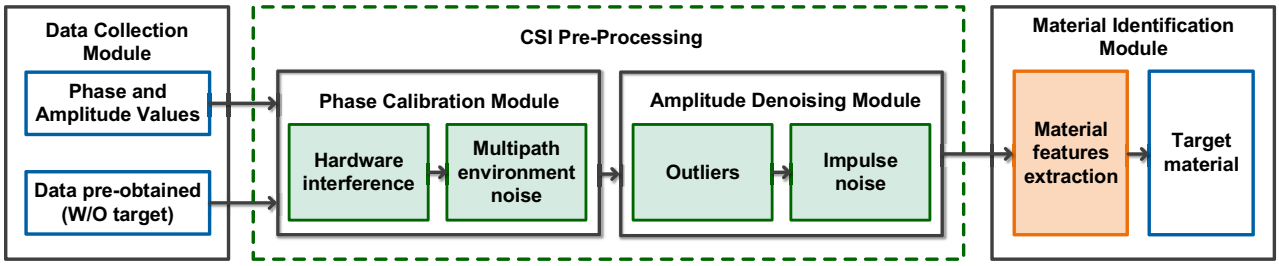


Figure 5: System workflow of WiMi.

measurements. As shown in Fig. 3, the outliers exceed the reasonable fluctuation region. The impulse noises appear at random time points and their amplitudes are comparable to the useful signals. If we do not remove these outliers and impulse noise, the material identification performance will be greatly affected.

3) Target Size Dependency: Different target materials cause different signal phase and amplitude changes. However, different target sizes also cause different amounts of phase and amplitude changes. Thus, we can not directly employ the phase and amplitude readings for material identification but need to find a new parameter which is uniquely related to the target material and at the same time, independent of the target size for material identification.

III. SYSTEM DESIGN

WiMi is a device-free material identification system built on commodity Wi-Fi devices. It only uses a transmitter for sending signals and a receiver for receiving signals. As a target appears at the LoS RF link between the transmitter and receiver, the CSI phase and amplitude will change, and we utilize these changes to design a material feature uniquely related to each material type and independent of target size to identify different materials.

A. System Overview

Now, we present the key components of WiMi followed by a summary of the system workflow as shown in Fig. 5.

- **Data Collection Module:** When there is no target (i.e., W/O target) between the transceiver pair, WiMi extracts a set of phase and amplitude readings from the measured CSI as the baseline data. Then WiMi captures another set of data when the target appears at the LoS link between the transceiver pair.
- **CSI Pre-Processing Module:** To calibrate the unstable CSI phase values, WiMi first uses two receiver antennas to remove the hardware noise and then eliminates the multipath effects by selecting 'good' subcarriers. For the CSI amplitude values, WiMi first rejects the outliers and then addresses the impulse noise. Further, we employ the second antenna again to obtain the stable amplitude ratio. Finally, the processed phase difference and amplitude ratio information are fed into the material identification module.
- **Material Identification Module:** After the above two steps, we use the processed phase difference and amplitude ratio readings to design a unique material feature, which is only determined by the target material and independent of the target size. The target material will then be identified through the classification algorithm employed.

B. CSI Phase Calibration

As discussed in [18, 21], we know that the raw phase readings reported at the Wi-Fi NICs are corrupted by asynchronization between transceivers and hardware noise, such as the Packet Boundary Delay (PBD), Sampling Frequency Offset (SFO) and Carrier Frequency Offset (CFO). To address these phase errors, we calibrate the phase values with a second antenna which is widely available at access points (APs) nowadays. The details are as follows.

The measured phase for the k-th subcarrier on the i-th antenna $\tilde{\phi}_{k,i}$ can be expressed by:

$$\tilde{\phi}_{k,i} = \phi_{k,i} + k\left(\lambda_b + \lambda_s\right) + \beta + Z \tag{5}$$

where λ_b is the PBD, λ_s is the SFO, β is the CFO, and Z is measurement noise, respectively. To deal with this problem, the key observation is that the commodity Wi-Fi devices are equipped with more than one antenna, and these antennas on the same board share the same sampling and oscillator clock, which means that they suffers from the same sampling frequency offset, carrier frequency offset and packet boundary delay [17, 22]. Thus, we can obtain the phase difference $\Delta\tilde{\phi}_k$ between two receiving antennas as follows:

$$\Delta\tilde{\phi}_k = \Delta\phi_k + \Delta Z \tag{6}$$

where $\Delta\phi_k$ is the theoretical phase difference, and ΔZ is the noise difference between two antennas, which follows a Gaussian distribution, which can be removed by averaging it over a time window. The results are shown in Fig. 2 and we can clearly see that the phase difference values of the one subcarrier (marked as red dots) are now aggregated in a relatively smaller area, ranging around 18 degrees. However, it is still has a relatively large variance of the phase difference

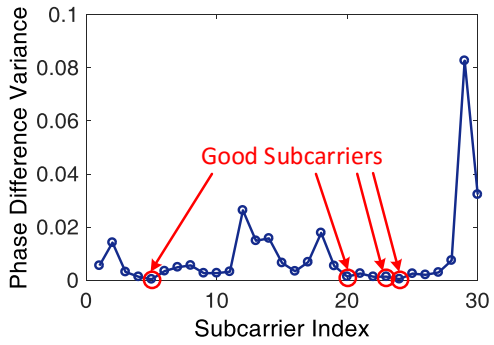


Figure 6: Phase difference variance for each subcarrier.

due to the multipath in indoor environments, which cannot be directly employed for target material identification.

Hence, to address this problem, our insight is that different subcarriers are affected differently by multipath. Specifically, we find that the phase differences at those subcarriers affected less by multipath are more stable and have smaller variances. Therefore, our goal is to select the subcarrier(s) with smaller variance across packets. In particular, the variance for the k-th subcarrier across M continuous CSI packets is calculated by the following equation:

$$\sigma_k^2 = \frac{1}{M}\sum_{m=1}^{M}\left(\Delta\tilde{\phi}_k(m) - \frac{1}{M}\cdot\sum_{m=1}^{M}\Delta\tilde{\phi}_k(m)\right)^2. \quad (7)$$

The variances at different subcarriers are illustrated in Figure 6. We can find P 'good' subcarriers which have the smallest phase difference variance. If we set $P = 4$, subcarrier 5, 20, 23, 24 are selected for material identification.

C. CSI Amplitude Denoising

As shown in Sec. II-C, the impulse noise occur irregularly and instantaneously, and the amplitudes are usually comparative to the useful signals. To eliminate the impact of the noise, we propose a denoising method, which contains two steps. The first step is the outlier removal. We first calculate the mean value μ_A and standard deviation σ_A of the measured CSI amplitudes. Then we preserve the data samples whose amplitude values are within the region $[\mu_A - 3\sigma_A, \mu_A + 3\sigma_A]$, and filter out the outliers which are outside of this region.

After this step, however, there still remain a lot of impulse noise. To remove them, our key observation is that the wavelet coefficients of the useful signals are often strongly correlated in different scales (frequencies), while the wavelet coefficients of the noises are weakly correlated or uncorrelated. Then by multiplying the wavelet coefficients of adjacent scales, the original useful signals can be reconstructed since they will be larger than the noise. We illustrate our observation through the following proof. Assume that the noise $n(x)$ obeys Gaussian distribution [21] with a mean value of 0 and a variance of σ^2, we can get the correlation of the noise:

$$R_n(u,v) = E[(n(u)\,n(v))] = \sigma^2, \quad (8)$$

where E denotes the expectation, u and v are different time points. Based on the discrete wavelet transform (DWT) [23], the power of wavelet transform noise can be given as:

$$\|W_l n(x)\|_2^2 = \sum_T\sum_T n(u)\,n(v)\,\psi_l(x-u)\,\psi_l(x-v), \quad (9)$$

where l denotes the wavelet scale and T is the discrete time number. Then we can obtain the expectation of $\|W_l n(x)\|_2^2$:

$$E(\|W_l n(x)\|_2^2) = \sum_T\sum_T \sigma^2\psi_l(x-u)\,\psi_l(x-v) = \frac{\|\psi\|_2^2\,\sigma^2}{l}, \quad (10)$$

where ψ is the wavelet function. From Equation(9) and (10), we observe that $\|W_l n(x)\|_2^2$ decreases when the frequency (scale l) increases, which demonstrates that the noises are weakly-correlated at different frequencies, and it is the opposite of the general characteristic of useful signals. That is to say, the useful signals are highly correlated at different frequencies.

Thereafter, we reconstruct the useful signal by integrating data from multiple frequencies. Consider a group of amplitude values $A_k = [a_k(1), \cdots, a_k(M)]$ calculated from the CSI measurements at the selected k-th subcarrier, the corresponding wavelet coefficients $W_k^l = [w_k^l(1), \cdots, w_k^l(M)]$ can be extracted by the DWT algorithm. Then the correlation coefficients at scale l can be formulated as:

$$Corr_k^l = W_k^l W_k^{l+1}. \quad (11)$$

To normalize the equation, we first calculate the power of the l^{th} wavelet coefficients $PW_k^l = \|W_k^l\|_2^2$ and the power of the l^{th} correlation coefficients $PCorr_k^l = \|Corr_k^l\|_2^2$. Hence, the normalized correlation coefficients can be given as:

$$NCorr_k^l = Corr_k^l\sqrt{PW_k^l / PCorr_k^l} \quad (12)$$

Then we compare the absolute difference values between W_k^l and $NCorr_k^l$, and the new wavelet coefficient of the m-th sample is:

$$Nw_k^l(m) = \begin{cases} w_k^l(m), & |w_k^l(m)| \geq |NCorr_k^l(m)| \\ 0, & |w_k^l(m)| < |NCorr_k^l(m)| \end{cases} \quad (13)$$

Note that, if $|w_k^l(m)| \geq |NCorr_k^l(m)|$, the coefficient $w_k^l(m)$ will be substituted by 0, otherwise, $w_k^l(m)$ remains unchanged. We repeat the aforementioned process until PW_k^l is less than or equals to the noise threshold at scale l, the choice of the threshold is based on robust median estimation [24].

After extracting the new wavelet coefficients at each scale, we can use them to reconstruct the original signal amplitude $\overline{A}_k$ by the inverse wavelet transform, which effectively

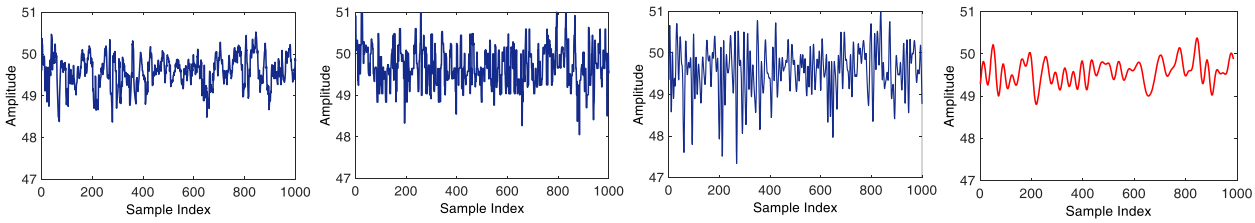


Figure 7: Performance of CSI amplitude denoising. (a)-(c) show three general filter methods, (d) shows our proposed method.

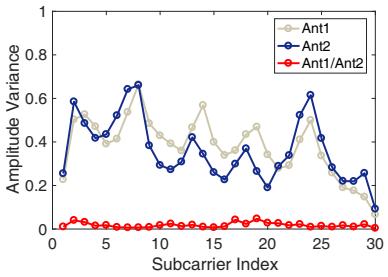


Figure 8: Amplitude variance of each subcarrier for each antenna.

removes the troublesome noises. As Fig. 7 shows, we apply the median filter, slide filter, butterworth filter, and our proposed method to eliminate the noise. It is clear that our method has the best noise removal performance than the other three traditional filters.

Furthermore, we find an interesting observation that the ratio of amplitude between two antennas is more stable than the amplitude of each antenna. As illustrated in Fig. 8, one can see the ratio of amplitude has much smaller variance than that of each individual antenna. The is because close-by antennas experience similar multipath, so the division operation not only removes the hardware noise but also partially removes the variations caused by the environmental multipath. Thus, we use the more stable amplitude ratio information to identify target material.

D. Target Material Identification

After the pre-processing for the CSI phase and amplitude, we obtain clean and stable phase difference and amplitude ratio for material identification. However, there exists another challenge. The material identification feature introduced in [3] does not work with commodity Wi-Fi devices equipped with multiple antennas. This is because the accurate absolute phase readings and amplitude readings can be obtained from commodity RFID devices but not from commodity Wi-Fi devices. For commodity Wi-Fi devices, as discussed in the previous two subsections, we can only obtain the stable and accurate phase difference and amplitude ratio but not the stable absolute phase and amplitude readings. Thus, we need to design a new material identification feature for the commodity Wi-Fi device.

Now we consider a two-antenna AP communicates with a single antenna Wi-Fi device. D_1 and D_2 respectively represent the propagation distances inside the target for antenna1 and antenna2. Based on Equation (2) and (4), we can easily get the following equations for each individual antenna:

$$\Delta\tilde{\phi}_1 = \tilde{\phi}_{tar1} - \tilde{\phi}_{free1} = D_1\left(\beta_{tar} - \beta_{free}\right) \tag{14}$$

$$\Delta A_1 = \frac{A_{tar1}}{A_{free1}} = e^{-D_1(\alpha_{tar}-\alpha_{free})} \tag{15}$$

and

$$\Delta\tilde{\phi}_2 = \tilde{\phi}_{tar2} - \tilde{\phi}_{free2} = D_2\left(\beta_{tar} - \beta_{free}\right) \tag{16}$$

$$\Delta A_2 = \frac{A_{tar2}}{A_{free2}} = e^{-D_2(\alpha_{tar}-\alpha_{free})} \tag{17}$$

Then, we calculate the phase difference and amplitude ratio between the two antennas to obtain the following equations:

$$\begin{aligned}\Delta\Theta = \Delta\widetilde{\phi}_1 - \Delta\widetilde{\phi}_2 &= \left(\widetilde{\phi}_{tar1} - \widetilde{\phi}_{tar2}\right) - \left(\widetilde{\phi}_{free1} - \widetilde{\phi}_{free2}\right) \\ &= \left(D_1 - D_2\right)\left(\beta_{tar} - \beta_{free}\right)\end{aligned} \tag{18}$$

and

$$\begin{aligned}\Delta\Psi = \frac{\Delta A_1}{\Delta A_2} &= \frac{A_{tar1}}{A_{tar2}} \cdot \frac{A_{free2}}{A_{free1}} \\ &= e^{-(D_1-D_2)(\alpha_{tar}-\alpha_{free})}\end{aligned} \tag{19}$$

where we successfully convert unstable $\tilde{\phi}_{tar1}-\tilde{\phi}_{\text{free1}}$ and $\tilde{\phi}_{tar2}-\tilde{\phi}_{\text{free2}}$ into stable $\tilde{\phi}_{tar1}-\tilde{\phi}_{tar2}$ and $\tilde{\phi}_{free1}-\tilde{\phi}_{free2}$. The same effect also applies to amplitude processing. We can now employ the processed stable phase difference and amplitude ratio to build the material feature and the details are introduced in the next section.

E. Addressing Target Size Issue

Besides the phase difference caused by target material, different target sizes also bring in phase differences. Without removing the target-size related phase difference, we are not able to employ the phase change for material identification.

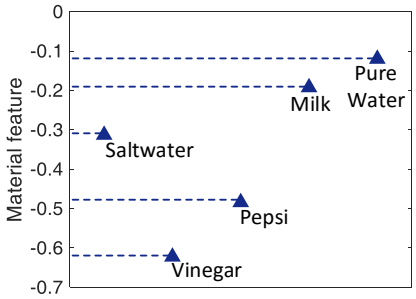
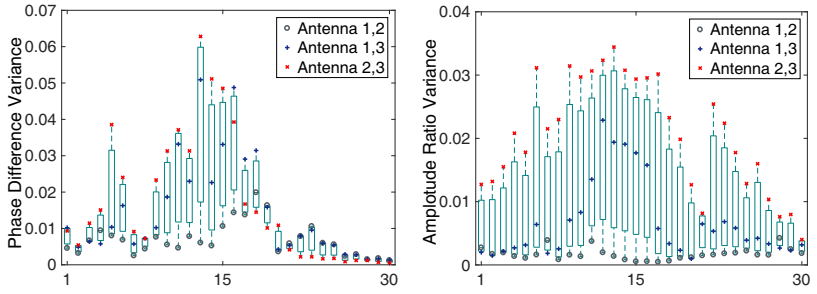


Figure 9: The material features for five different materials.



(a) Phase difference variance of different antenna combinations.

(b) Amplitude ratio variance of different antenna combinations.

Figure 10: Variances are different for each antenna combination.

Inspired by [3], we design a parameter $\bar{\Omega}$ which can be calculated with just the amplitude ratio and phase changes. This parameter $\bar{\Omega}$ is uniquely related to the target material and is independent of the target size. Based on Equation (18) and (19), we get the following equation:

$$D_1 - D_2 = \frac{\Delta\Theta + 2\gamma\pi}{\beta_{tar} - \beta_{free}} = \frac{-\ln \Delta\Psi}{\alpha_{free} - \alpha_{tar}}, \tag{20}$$

where γ is an integer. Then, we can obtain the following equation with parameter $\bar{\Omega}$ calculated as:

$$\bar{\Omega} = \frac{-\ln \Delta\Psi}{\Delta\Theta + 2\gamma\pi} = \frac{\alpha_{free} - \alpha_{tar}}{\beta_{tar} - \beta_{free}}, \tag{21}$$

Note that $\Delta\Theta$ is the phase difference and $\Delta\Psi$ is amplitude ratio we can directly calculate with measurements from the commodity Wi-Fi device. β_{free} and α_{free} are constants. For a given material, the values of β_{tar} and α_{tar} are also constants. As for γ, it is an integer value related to the signal propagation distance inside the target. If the container's size changes in a small range, γ does not change. As an integer value, γ can be accurately estimated with the coarse CSI amplitude readings.

In Equation (21), We can see that, for the obtained parameter $\bar{\Omega}$, D_1 and D_2 disappear which means the parameter is independent of the target size. The signal phase constant β_{tar} and attenuation constant α_{tar} are decided by the target material. Thus, $\bar{\Omega}$ is also uniquely related to the target material type. What is more important, $\bar{\Omega}$ can be obtained as shown in Equation (21) with just phase difference ($\Delta\Theta$) and amplitude ratio ($\Delta\Psi$) which can be measured and directly at commodity Wi-Fi devices. We name $\bar{\Omega}$ *material feature* and employ this unique feature to identify a target's material.

Further, we conduct a group of benchmark experiments to demonstrate the availability of the material feature $\bar{\Omega}$ in the office environment. The test liquids are 'Saltwater', 'Vinegar', 'Pepsi','Milk' and 'Pure water'. The results in Fig.9 illustrates that the material features $\bar{\Omega}$ can be taken as an effective reference to identify the target material. Finally, we put the extracted feature values into the material database. Then, when identifying a test material, WiMi collects the phase and amplitude change measurements, and incorporates the material database and the SVM classifier to identify the target material.

F. Selecting Antenna Pairs

Although WiMi can use two receiver antennas to perform material identification, current commercial Wi-Fi devices are equipped with more than two antennas. Given the receiver has p antennas, we can extract $\frac{p(p-1)}{2}$ phase differences and amplitude ratio values. As Fig. 10 presents, one can clearly see the variances of phase difference values and amplitude ratio values are different for different combinations of antennas. Hence, we can improve the stability of material identification by selecting the appropriate antenna combination.

IV. IMPLEMENTATION

Hardware setup: We build a prototype of WiMi using a commodity wireless router as transmitter and a laptop with an Intel 5300 Wi-Fi NIC with three antennas as receiver. The router works in 802.11n AP mode at 5 GHz frequency band. The laptop has a 3.6 GHz CPU (Intel i7-4790) and 8 GB memory. The laptop is employed to receive CSI measurements every 10 ms and run our material identification algorithm.

Experimental environments: To evaluate the material identification performance of WiMi, we conduct experiments in three typical indoor environments: an empty hall, a laboratory, and a library. Note that the three environments represent low, medium, and high multipath environments, respectively. In each environment, the laptop is placed 2 m from the router, and the tested target is placed at the LoS RF link. Note that we first extract a set of phase and amplitude values as the baseline data when the empty plastic beaker is placed at the LoS RF link, and then we pour the same amount of each tested liquid into the empty plastic beaker to measure another set of data. We wait a few seconds to let tested liquid become stable before each test. The

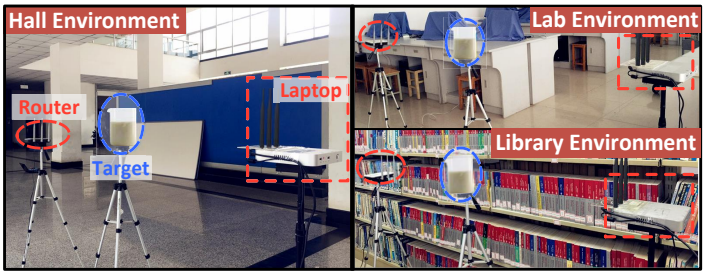


Figure 11: Deployment in three different environments.

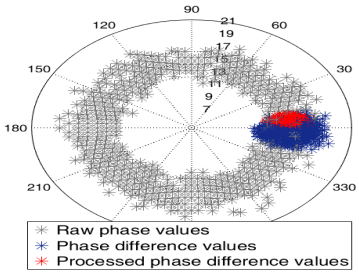


Figure 12: Performance of phase calibration scheme.

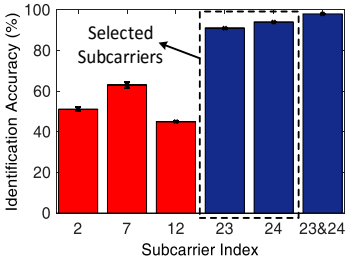


Figure 13: Identification performance with different subcarriers.

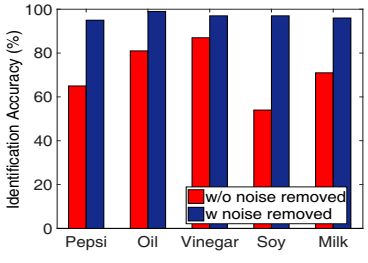


Figure 14: Identification performance of the amplitude denoising method.

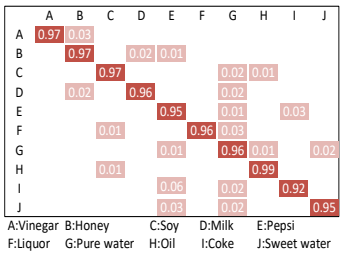


Figure 15: Identification performance for 10 liquids.

plastic beaker has a diameter of 14.3 cm and a height of 23 cm. Fig. 11 shows the deployment setup in three different environments.

Tested targets: To evaluate the material identification performance of WiMi, we test 10 different liquids including 'Vinegar', 'Honey', 'Soy', 'Milk', 'Pepsi', 'Liquor', 'Pure water', 'Oil', 'Coke', and 'Sweet water'.

V. PERFORMANCE EVALUATION

A. Microbenchmark

We conduct two benchmark experiments to validate the practicability and effectiveness of the pre-processing methods for the CSI phase and amplitude.

Verification of phase calibration: We conduct experiments in the library environment to test the phase calibration scheme. For each tested material, we collect CSI readings for 10 seconds, and repeat this process 20 times.

Fig. 12 illustrates the phase distributions by using the proposed calibration scheme. We can clearly see that the raw phase values are randomly distributed from 0 to 2π. Then after we remove the hardware interference with the phase difference, the obtained new phase difference values assemble to a small region (angular fluctuation is around 18 degrees). At last, we select the 'good' subcarriers whose phase difference values are more stable and have a more concentrated distribution. The angular fluctuation is further decreased to 5 degrees.

Further, to verify the efficacy of the subcarrier selection scheme, we randomly select 3 subcarriers (2, 7, 12) and compare the material identification performance with the two 'good' subcarriers (23, 24)[2]. The performance comparison is illustrated in Fig. 13. Since the phase values at the randomly selected subcarriers have large variances, the two 'good' subcarriers achieve a much higher identification accuracy. Furthermore, when we combine subcarrier 23 with 24, the identification accuracy is further increased, higher than applying only one of them. These results demonstrate the effectiveness of our proposed subcarrier selection method.

Verification of amplitude denoising: To further verify the impact of our two-antenna denoising method on target material identification, we respectively use the raw amplitude ratio values and the amplitude ratio values with both the outliers and impulse noise removed to identify the target material. The results are presented in Fig. 14 and it is clear that the material identification performance with our two-antenna denoising method is consistently better. These results demonstrate the effectiveness of proposed denoising method.

B. Overall Performance

We evaluate the material identification performance in the lab environment. For each tested target, we repeat collecting the measurements 20 times. Fig. 15 presents the identification results for ten different liquids. WiMi achieves an average accuracy of 96%. To evaluate WiMi's identification performance for the same type of liquid but with different concentrations, we pour three different concentrations of

[2]In our test experiments, the default target material is milk.

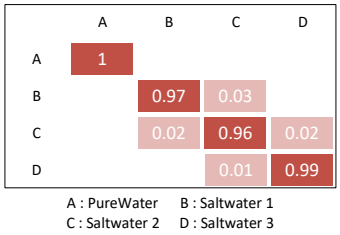


Figure 16: Identification performance with different concentrations of saltwater.

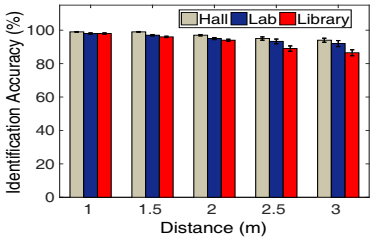


Figure 17: Identification performance with varying distance between transmitter and receiver.

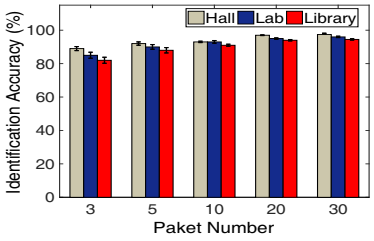


Figure 18: Identification performance with varying number of packets.

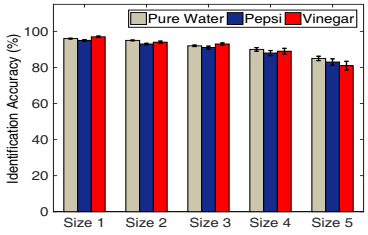


Figure 19: Identification accuracy with varying Container size.

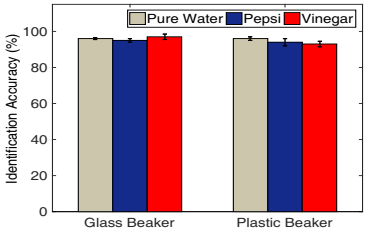


Figure 20: Identification accuracy with varying Container material.

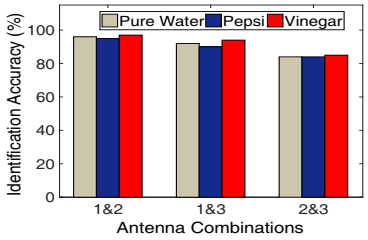


Figure 21: Performances for different antenna combinations.

saline water (1.2g/100ml, 2.7g/100ml and 5.9g/100ml) into the same plastic container. The identification results are shown in Fig. 16 and we can see that WiMi can still achieve higher than 95% accuracy.

The impact of distance between transmitter and receiver: Different distances between the transmitter and receiver would result in different material identification accuracy, since the amount of multipath and diffraction increase as the distance increases. As Fig. 17 shows, when the distance increases from 1 m to 3 m at a step size of 0.5 m, the identification accuracy decreases from 98% to 87.3%. We can thus conclude that the larger distance degrades the system performance but WiMi can still achieve around 90% accuracy when the transceiver pair is separated by 3 m.

The impact of packet number: We now evaluate WiMi's performance with different numbers of packets. As Fig. 18 illustrates, WiMi achieves an increasing material identification accuracy when the packet number grows from 3 to 30. The reason is that the phase and amplitude errors are reduced more by the two pre-processing schemes with a larger number of packets. However, it does not mean the performance will keep improving with more packets. We can see that WiMi achieves a similar accuracy when the number of packets is increased from 20 to 30. Note that it takes more time to receive more number of packets. Thus, in our implementation, WiMi employs 20 packets for target material sensing.

The impact of container's size: To verify the impact of different container sizes, we run a set of experiments by pouring test liquids into five glass beakers with different sizes. We pour the liquid until the same height of 23 cm. The diameters of the five beakers are 14.3 cm, 11 cm, 8.9 cm, 6.1 cm, and 3.2 cm, referred as Size 1, Size 2, Size 3, Size 4, and Size 5, respectively. As Fig. 19 shows, the identification accuracy fluctuates in the range of 95% to 91% when the diameter is decreased from 14.3 cm to 8.9 cm. When the container's diameter is further decreased to 3.2 cm, the identification accuracy is decreased obviously. We believe this is because when the diameter is smaller than the wavelength (6cm) of the signal, diffraction degrades the identification accuracy.

The impact of different material containers: We evaluate the impact of container material on WiMi's performance. We pour the test liquids into a plastic beaker and a glass beaker respectively, both of which have a same diameter of 14.3 cm and a height of 23 cm. The results are shown in Fig. 20. We can clearly see that the identification accuracies are similar for the same liquid with different containers. This is because we collect a set of base data before the liquid is poured into the container and then collect another set of data when the target liquid is poured in. By taking the difference, we actually remove the effect of container so the container material has little effect on the identification performance as long as the container is not too small. Note that, when the container is metallic or covered with a foil paper, the RF signal will be essentially reflected back rather than penetrating through the target, causing our system to stop working.

The impact of Antenna combinations: As shown in Fig. 21, one can see that different antenna combinations

have slightly different material identification accuracies. The combination of antenna 1 and antenna 2 achieves the best performance than the other two pairs. We believe this is because different antenna pairs experience different multipath and environmental noise, which result in different degree of stability of phase difference and amplitude ratio. Thus, we could select appropriate antenna pair to achieve better material identification performance when there are more than two antennas.

VI. DISCUSSION

The current version of WiMi has several limitations. First, we cannot identify the target's material if it is comprised of two or more materials. In this work, we mainly focus on single-material target. Second, WiMi requires the target to be placed between the transmitter and receiver with LoS signal penetrating through the target. Thus, WiMi can only identify one target at a time with one WiFi transmitter-receiver pair. However, we believe there will be more and more devices equipped with a Wi-Fi module in our surrounding environment in the future, all of which will be connected to the Wi-Fi AP. Hence, more Wi-Fi links can be available to be employed for material sensing. Third, our current system can only identify the material type of a static liquid. When the target is moving or the liquid is flowing, it will cause extra changes of signal phase and amplitude and it is then challenging to perform material identification. In the future, we would like to dig deep in this interesting direction and employ deep learning techniques to address the fluid dynamics for sensing.

VII. RELATED WORK

Material Identification: Existing literatures on material identification are mainly divided into two groups. The first group is the RF-based methods, such as the RSA system [25], which uses millimeter wave signal to identify a target's surface material. RadarCat [26] touches target and leverages 60 GHz signal to identify the material. Although they have a high material identification accuracy, both of them require to use high frequency millimeter wave signal and the sensing range is very small. Also, recently, TagScan system [3] utilizes the RFID signal for accurate target material identification owing to the very stable absolute phase and RSS readings from RFID device, which is not available on commodity Wi-Fi devices. Another latest work LiquID [16] uses UWB device to perform material identification achieving a high accuracy by connecting a wire between the transmitter and receiver. Compared to UWB and RFID devices, Wi-Fi is more ubiquitous and already widely deployed in our surrounding environment.

The second group is based on dedicated devices, such as Radar [4], X-Ray [5], CT/MRI [6], These systems require dedicated hardware to emit high frequency and high bandwidth signals. These equipments are usually not only huge in size but is also expensive despite these systems can achieve a high material identification accuracy. Systems further employ visible light camera [9, 27] to identify target material, and they only work with good-lighting conditions and can not identify targets inside a non-transparent container. On the other hand, WiMi is built on commodity Wi-Fi device which can identify the the material type of targets inside a non-transparent container and has no requirement on the lighting conditions.

Phase Calibration and Amplitude Denoising: The issues of unstable phase and noisy amplitude are unavoidable with commodity Wi-Fi hardware. For phase calibration, PinLoc [21] and PhaseU [18] use a linear transform to remove the phase offsets caused by hardware. Unlike these approaches, WiMi addresses the random phase problem based on the observation that the random phase jumps are the same on closely-spaced antennas on a board. Furthermore, we identify 'good' subcarriers which are less contaminated by the multipath in indoor environment for material identification. For amplitude calibration, current works such as CARM and Wikey [13, 18] use PCA technology to remove the environmental noise effectively for single antenna, which is still not stable enough for our system. This is because we need more stable amplitude information to perform fine-grained material sensing. WiMi first uses a threshold to filter out the outliers and utilizes the fact that noises are uncorrelated in different frequencies to remove the impulse noise effectively. Then WiMi applies more stable amplitude ratio between closely-spaced antennas to perform material identification.

VIII. CONCLUSION

We present WiMi, the first COTS Wi-Fi based material identification system, which works with the noisy CSI readings. WiMi proposes novel CSI phase calibration and CSI amplitude denoising methods to obtain clean and stable phase difference and amplitude ratio information. WiMi leverages the phase difference and amplitude ratio to design a new material feature, which is only related to the target material and independent of target size. Extensive experimental results demonstrate the effectiveness and high accuracy of the proposed system. We believe the proposed CSI denoising methods can be applied to benefit other sensing applications.

ACKNOWLEDGMENT

This work is partially supported by the National Natural Science Foundation of China under Grant Nos. 61802307, 61572402 and 61602382, the ShaanXi Science and Technology Innovation Team Support Project under grant agreement 2018TD-O26, and the Postgraduate independent innovation project of NWU under Grant Nos. YZZ17184. Xiaojiang Chen is the corresponding author.

REFERENCES

[1] C. A. Kendziora, R. A. Mcgill, R. Furstenberg, M. Papantonakis, V. Nguyen, and J. M. Byers, "Detecting traces of explosives," *Spie Newsroom*, 2015.

[2] S. Cussatblanc and J. Pollack, "Cracking the egg: Virtual embryogenesis of real robots," *Artificial Life*, vol. 20, no. 3, pp. 361–383, 2014.

[3] J. Wang, J. Xiong, X. Chen, H. Jiang, R. K. Balan, and D. Fang, "Tagscan: Simultaneous target imaging and material identification with commodity rfid devices," in *Proc. ACM International Conference on Mobile Computing and Networking. (MobiCom)*, 2017.

[4] S. M. J. Mortazavi, T. Shahram, and N. Dehghan, "Alterations of visual reaction time and short term memory in military radar personnel," *Iranian journal of public health*, vol. 42, no. 4, p. 428, 2013.

[5] K. Ogawa, T. Hirokawa, and S. Nakamura, "Identification of a material with a photon counting x-ray ct system," in *Proc. IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, 2010, pp. 2582–2586.

[6] H. Zhou, Y.-X. J. Wang, H.-y. Lou, X.-j. Xu, and M.-m. Zhang, "Hepatic sinusoidal obstruction syndrome caused by herbal medicine: Ct and mri features," *Korean journal of radiology*, vol. 15, no. 2, pp. 218–225, 2014.

[7] F. De Bats, B. Wolff, M. Mauget-Faÿsse, C. Scemama, and L. Kodjikian, "B-scan and en-face spectral-domain optical coherence tomography imaging for the diagnosis and followup of acute retinal pigment epitheliitis," *Case reports in medicine*, vol. 2013, 2013.

[8] S. Nanda, S. Manna, A. K. Sadhu, A. Konar, and D. Bhattacharya, "Real-time surface material identification using infrared sensor to control speed of an arduino based car like mobile robot," in *Third International Conference on Computer, Communication, Control and Information Technology*, 2015, pp. 1–6.

[9] J. Ševek, S. Katuščák, L. Dubinyová, and M. Fikar, "An automatic identification of wood materials from color images," in *Cybernetics & Informatics (K&I)*, 2016, pp. 111–116.

[10] F. Adib and D. Katabi, "See through walls with wifi!" in *Acm Sigcomm Conference on Sigcomm*, 2013.

[11] J. Xiong and K. Jamieson, "Arraytrack: a fine-grained indoor location system," in *Usenix Conference on Networked Systems Design and Implementation*, 2013, pp. 71–84.

[12] J. Wang, H. Jiang, J. Xiong, K. Jamieson, X. Chen, D. Fang, and B. Xie, "Little human-effort for device-free localization with fine-grained subcarrier information," in *Proc. ACM International Conference on Mobile Computing and Networking. (MobiCom)*, 2016, pp. 243–256.

[13] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of wifi signal based human activity recognition," in *Proc. ACM International Conference on Mobile Computing and Networking. (MobiCom)*, 2015, pp. 65–76.

[14] J. Zhang, Z. Tang, M. Li, D. Fang, P. Nurmi, and Z. Wang, "Crosssense: Towards cross-site and large-scale wifi sensing," in *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*. ACM, 2018, pp. 305–320.

[15] T. Wei and X. Zhang, "mtrack:high-precision passive tracking using millimeter wave radios," in *Proc. ACM International Conference on Mobile Computing and Networking. (MobiCom)*, 2015, pp. 117–129.

[16] Y. Z. H. H. R. R. C. Ashutosh Dhekne, Mahanth Gowda, "Liquid: A wireless liquid identifier," in *Pro. ACM international conference on Mobile systems, applications, and services. MobiSys*, 2018.

[17] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity wifi," in *Proc. ACM International Conference on Mobile Computing and Networking. (MobiCom)*, 2015, pp. 53–64.

[18] C. Wu, Z. Yang, Z. Zhou, K. Qian, Y. Liu, and M. Liu, "PhaseU: Real-time los identification with wifi," in *Pro. IEEE conference on Computer Communications (INFOCOM)*, 2015, pp. 2038–2046.

[19] R. v. Nee and R. Prasad, *OFDM for wireless multimedia communications*. Artech House, Inc., 2000.

[20] A. S. D. W. Daniel Halperin, Wenjun Hu, "Linux 802.11n csi tool," http://dhalperi.github.io/linux-_80211n-_csitool/faq.html.

[21] S. Sen and Radunovic, "You are facing the mona lisa: spot localization using phy layer information," in *Pro. ACM international conference on Mobile systems, applications, and services. MobiSys*, 2012, pp. 183–196.

[22] J. Xiong, K. Sundaresan, and K. Jamieson, "Tonetrack: Leveraging frequency-agile radios for time-based indoor wireless localization," in *Proc. ACM International Conference on Mobile Computing and Networking. (MobiCom)*, 2015, pp. 537–549.

[23] C. Torrence and G. P. Compo, "A practical guide to wavelet analysis," *Bulletin of the American Meteorological Society*, vol. 79, no. 1, 1998.

[24] Y. Xu, J. B. Weaver, D. M. Healy, and J. Lu, "Wavelet transform domain filters: a spatially selective noise filtration technique," *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 3, no. 6, pp. 747–758, 1994.

[25] Y. Zhu, Y. Zhu, B. Y. Zhao, and H. Zheng, "Reusing 60ghz radios for mobile radar imaging," in *Proc. ACM International Conference on Mobile Computing and Networking. (MobiCom)*, 2015, pp. 103–116.

[26] H. S. Yeo, G. Flamich, P. Schrempf, D. Harris-Birtill, and A. Quigley, "Radarcat: Radar categorization for input & interaction," in *Symposium on User Interface Software and Technology*, 2016, pp. 833–841.

[27] C. Balas, G. Epitropou, A. Tsapras, and N. Hadjinicolaou, "A novel hyperspectral camera and analysis platform for the non-destructive material identification and mapping: An application in paintings by el greco," in *Proc. IEEE International Conference on Imaging Systems and Techniques*, 2016, pp. 211–215.