

WiMate: Location-independent Material Identification Based on Commercial WiFi Devices

Yu Gu

School of Computer and Information
Hefei University of Technology
Hefei, China
yugu.bruce@ieee.org

Yanan Zhu

School of Computer and Information
Hefei University of Technology
Hefei, China
2019110993@mail.hfut.edu.cn

Jie Li

Dept. of Computer Science and Engineering
Shanghai Jiao Tong University
Shanghai, China
lijie@cs.sjtu.edu.cn

Yusheng Ji

Information Systems Architecture Research Division
National Institute of Informatics
Tokyo, Japan
kei@nii.ac.jp

Abstract—Material identification is playing an increasingly important role in our daily lives such as public security checks. X-ray-based technologies are highly radioactive because they rely on specialized devices to transmit high-frequency signals. Ultrasound-based technologies are cumbersome due to their large size. RF-based approaches necessitate the use of RFID which is usually expensive to be used in home and office environments. To this end, WiFi-based material identification approach has emerged recently as a low-cost yet effective alternative. In this paper, we propose WiMate, a noncontact material identification system leveraging only off-the-shelf WiFi devices. The key enabler of WiMate is a novel theoretical model we build to characterize how the electromagnetic wave decays when penetrating different materials. Our model identifies a unique feature for each material that only depends on the material itself. Consequently, we can leverage this feature coupling with the machine learning techniques for robust and accurate material identification. We prototype WiMate using low-cost commodity WiFi devices and evaluate its performance in real-world. The empirical study shows that WiMate can identify six different materials, i.e., board, paperboard, nickel, wood chip, iron and titanium, with an average accuracy of 96.20%.

Index Terms—Channel State Information (CSI), WiFi, Material Identification

I. INTRODUCTION

Material identification has a wide range of applications in commercial and industrial fields, such as distinguishing water from salt water without tasting [1], analyzing the explosion products after explosion [2], adjusting the grip according to the hardness of the material [3].

Traditional material identification methods include X-ray-based, ultrasonic-based, CT/MRI-based and RF-based. In general, X-ray-based [4], ultrasound-based [5], CT/MRI-based [6] methods are similar in that they both require specialized hardware to transmit signals at extremely high frequencies. Despite their high recognition accuracy, these required hardwares are costly and heavy, making them unsuitable for business scenarios. Furthermore, the radioactivity of X-rays causes damage to the human body. The RF-based methods [1],

[7], [8] use RFID and UWB signals for material identification to achieve high accuracy. While they can overcome the issue of large scale, it necessitates the use of costly RFID reading equipment.

WiFi devices have been used more and more widely because of their low cost and non-destructive advantages. Currently, the primary applications of WiFi-based technologies can be split into two categories: the one for humans and the other one for materials. In the former case, the main principle is that the signal diffracted when it encountered a human body. Gu et al. took advantage of this phenomenon to realize a non-contact behavior analysis system based on signal processing and computational intelligence of WiFi CSI [9], [10]. In the latter case, the key is that when the signal comes into contact with the material, it transmits and fades to varying degrees. Yang et al. [11] presented a non-destructive and economic wheat moisture detection system with commodity WiFi. Although [11] had a high accuracy rate for detecting wheat moisture, it only conducted research on wheat, and the types of materials evaluated were few.

We propose WiMate, a non-destructive passive material recognition system, based on the above observations. It calculates the feature value only related to the material itself, based on the material's effect on the signal amplitude and phase, and uses the obtained feature value to realize the material identification. Our prototype system is made up of two miniPCs, each with an Intel 5300 WiFi NIC. The transmitting has one antenna and the receiving has two antennas. WiMate has a 96.2% recognition accuracy for six materials at different sites according to experimental results and is robust to different experimental environments. This paper's main contributions are summarized as follows:

- We use a low-cost WiFi device to achieve location-independent material identification.
- We quantify a feature value that is only connected to the material itself, relying on the theory that different materials cause different signal attenuation.

- We implement a prototype system for material identification. The experimental results indicate that the detection performance of this method is up to 96.2%.

The remainder of this paper is organized as follows: in the following section, we provide a summary of our preliminary work, and in section III, we present the system design. In section IV, we quantify the experimental effects. Finally, we bring our work to a close in section V.

II. PRELIMINARIES

A. Channel State Information

The WiFi standard 802.11n adopts orthogonal frequency division multiplexing technology to communicate [12]. Unlike received signal strength (RSS), which only captures the superposition of multipath signals, CSI reveals fine-grained information such as different time delays, amplitude attenuation and phase shifts of multipath signals on each subcarrier as the signals are transmitted between each transmitting and receiving antenna pair. The 30 subcarriers contained in the CSI measurements can be exported through CSI Tools. The channel frequency responses can usually be expressed as:

$$H(f, t) = H_l(f, t) + H_m(f, t). \quad (1)$$

Where $H(f, t)$ represents the complex value channel frequency responses in the format of CSI, H_l is the signal passing on the line-of-sight (LOS) path, and H_m represents the total impact of other multipath signals. $H(f, t)$ can also be expressed by the following formula:

$$H(f, t) = h(f, t) \cdot e^{j\theta}, \quad (2)$$

where $h(f, t)$ and θ respectively represent the amplitude and phase of the transmitted signal. So equation (1) can be rewritten as equation (3),

$$h(f, t) \cdot e^{j\theta} = h_l(f, t) \cdot e^{j\theta_l} + h_m(f, t) \cdot e^{j\theta_m}. \quad (3)$$

B. Preliminary Experiments

[Prototype] Fig. 1 shows our prototype, which is made up of two commodity MiniPCs and an Intel Network Interface Controller (NIC) 5300. One is connected to one external antenna as the sender, while the other is connected to two antennas as the receiver. The antennas are mounted on tripods. The sampling frequency is set to 1000 Hz.

[Experimental Materials] The experimental materials included one piece of board, paperboard, nickel, wood chip, iron and titanium of $100 \text{ mm} \times 100 \text{ mm} \times 1 \text{ mm}$ is required as shown in Fig. 2.

[Environment] The preliminary experiment was carried out in an $8.3 \text{ m} \times 10 \text{ m}$ hall as shown in Fig. 1. There are chairs, bookcases, computer desks, sofas and other furniture in the hall. During the preliminary experiment, there are only experimenters in the room.

[Settings] Set up various materials and place them on a non-metal platform between the T_X and the R_{X_1} . The platform is a $20 \text{ cm} \times 20 \text{ cm}$ bracket. The duration of each experiment is 20 seconds.

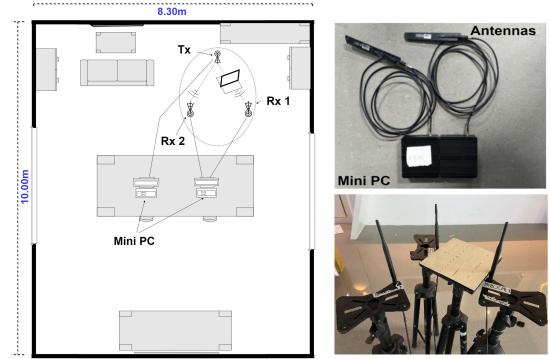


Fig. 1. Prototype of WiMate.

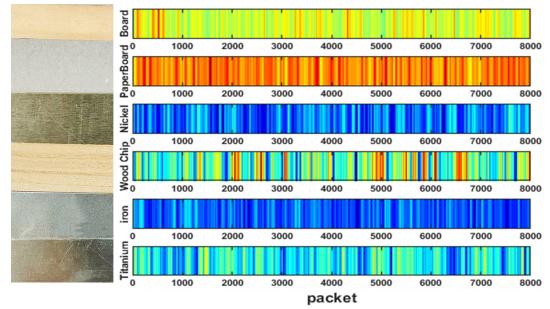


Fig. 2. 16_{th} subcarrier data of different materials (in the middle position).

Fig. 2 shows the amplitude example of the 16_{th} subcarrier of 8000 data packets measured by different materials at the middle position between T_X and R_{X_1} . The Different color shows the material's influence on the signal to varying degrees. It clearly shows that different materials have different effects on CSI.

Place the material at the center of T_X and R_{X_1} (the distance between T_X and R_{X_1} is 20 cm). Fig. 3 shows an example of the amplitude and phase difference of all 30 subcarriers of R_{X_1} in the time domain (The phase difference example in Fig. 3 is phase adjusted, which will be introduced in Section III). It can be seen from Fig. 3 that there are relatively obvious fluctuations in the amplitude and phase difference of the first 5 seconds. This is caused when the experimenter leaves the experimental environment after the experiment has started. After 5 seconds, the signal gradually stabilized. For each subcarrier, there is still some random noise in the data, and we use time-averaging technology to smooth the noise. To ensure the continued stability of the signal and avoid the influence of human factors, the experiment uses CSI data between 10s and 20s.

In summary, preliminary studies have confirmed that different materials have different effects on channel response data. In the following section, we will go over the specifics of the system design.

C. Support Vector Machine (SVM) Classification

Non-linear SVM is employed to classify the processed CSI data for material identification. For the nonlinear classification

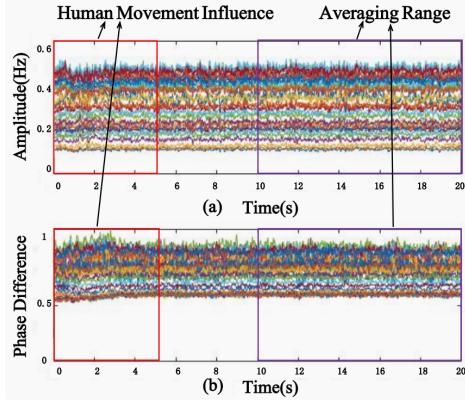


Fig. 3. R_{X_1} CSI data for all 30 subcarriers. (a)Amplitude. (b)Phase Difference

problem in the input space, it can be transformed into a linear classification problem in some dimensional feature space by nonlinear transformation, and the linear SVM can be learned in the high-dimensional feature space. We randomly divide the processed data into two groups to train and test, and then find a hyperplane in the data space. The training procedure is as follows.

Step 1 Input data set:

$$T = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_N, y_N)\}, \quad (4)$$

and randomly divided into training set and test set.

Where x_i represents the feature value of the material, N represents the total number of data set, $y_i = 1, 2, 3, \dots, n$, and n is the kinds of the materials tested.

Step 2 Select the appropriate kernel function $K(a, b)$ and a penalty coefficient C , $C > 0$, construct the constrained optimization problem as:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) - \sum_{i=1}^N \alpha_i \\ \text{s.t.} \quad & \sum_{i=1}^N \alpha_i y_i = 0 \\ & 0 \leq \alpha_i \leq C, i = 1, 2, \dots, N \end{aligned} \quad (5)$$

Get the optimal solution:

$$\boldsymbol{\alpha}^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_N^*)^T \quad (6)$$

Step 3 Choose a component α_j^* of $\boldsymbol{\alpha}^*$ and calculate:

$$b^* = y_j - \sum_{i=1}^N \alpha_i^* y_i K(\mathbf{x}_i, \mathbf{x}_j) \quad (7)$$

Step 4 Classification decision function:

$$f(x) = \text{sign} \left(\sum_{i=1}^N \alpha_i^* y_i K(x, \mathbf{x}_i) + b^* \right) \quad (8)$$

Due to the complexity of the indoor environment, this problem cannot be separated linearly. We can use Gaussian RBF as the kernel function to map the processed CSI data to the high-dimensional eigenvector space.

III. SYSTEM DESIGN

WiMate is a passive material identification system based on common WiFi devices. It just makes use of the transmitter to deliver the signal and the receiver to receive it. The phase and amplitude of the CSI will change when the target appears on the LOS link between the transmitter and the receiver. To identify different materials, we use these changes to create a material feature value that is unique to each material type and regardless of the target location. Fig. 4 shows the system architecture.

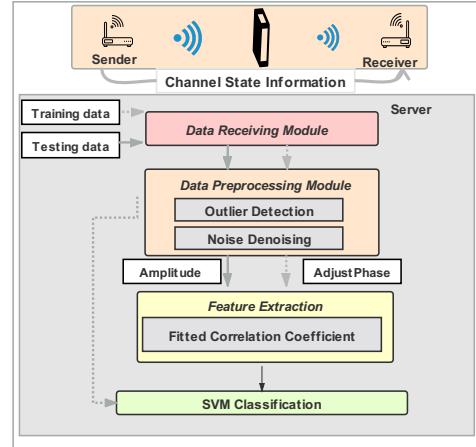


Fig. 4. System architecture of WiMate.

A. Theoretical Model

When the microwave propagates through a dissipative medium, its amplitude is attenuated by $e^{-K_I \cdot D}$. Meanwhile, the medium's phase transition is measured as $K_R \cdot D$ [13], where D is the propagation distance, which is equal to the material thickness. K_I , K_R can be calculated as follows:

$$K_I = \omega \sqrt{\mu_0 \varepsilon_0 \varepsilon_r} \left[\frac{1}{2} \left(\sqrt{1 + \frac{\sigma^2}{(\varepsilon_0 \varepsilon_r \omega)^2}} - 1 \right) \right]^{1/2}, \quad (9)$$

$$K_R = \omega \sqrt{\mu_0 \varepsilon_0 \varepsilon_r} \left[\frac{1}{2} \left(\sqrt{1 + \frac{\sigma^2}{(\varepsilon_0 \varepsilon_r \omega)^2}} + 1 \right) \right]^{1/2}. \quad (10)$$

Among them, ω represents the angular frequency of the wave, μ_0 and ε_0 are the magnetic permeability and the dielectric constant under vacuum respectively. ε_r and σ are the relative dielectric constant of the medium and the electrical conductivity of the medium respectively. K_I and K_R are only related to the material.

When placing the material to be identified on the LOS path between T_X and R_{X_1} , (3) can be rewritten as follow:

$$\begin{aligned} h_{R_1}(f, t) \cdot e^{j\theta_{R_1}} = & h'_l(f, t) \cdot e^{-K_I \cdot D} \cdot e^{j(\theta'_l - K_R \cdot D)} \\ & + h_m(f, t) \cdot e^{j\theta_m}, \end{aligned} \quad (11)$$

where $h'_l(f, t)$ and θ'_l represent signals that are not affected by the material.

In WiMate, the clock of transmitter and receiver is out of sync. However, the sampling clock of different antennas at the receiving end is the same. To solve the problem of the transmitter and receiver clock being out of sync, all channel response data is multiplied by a factor $e^{-j\theta_{R_2}}$, which is equivalent to phasing the data so that the phase of R_2 is zero. After phasing, all phases are synchronized to R_1 . The phase difference $\Delta\theta_{R_1} = \theta_{R_1} - \theta_{R_2}$ represents the phase change affected by the material. So equation (11) can be rewritten as follow:

$$h(f, t) \cdot e^{j\Delta\theta_{R_1}} = h'_l(f, t) \cdot e^{-K_l \cdot D} \cdot e^{j(\Delta\theta_l - K_l \cdot D)} + h_m(f, t) \cdot e^{j\Delta\theta_m}, \quad (12)$$

where $\Delta\theta_l = \theta_l - \theta_{R_2}$, $\Delta\theta_m = \theta_m - \theta_{R_2}$, $\Delta\theta_{R_1} = \theta_{R_1} - \theta_{R_2}$.

The multiplication of transmitted signal using the channel frequency responses in the frequency domain can also express the received signals:

$$R(f) = S(f) \cdot H(f). \quad (13)$$

Therefore, it can be seen that:

$$h(f, t) \cdot e^{j\Delta\theta_{R_1}} = S(f) \cdot \|H(f)\| \cdot e^{j\Delta\theta_{R_1}}, \quad (14)$$

$$h'_l(f, t) \cdot e^{j\Delta\theta_l} = S(f) \cdot \|H_l(f)\| \cdot e^{j\Delta\theta_l}, \quad (15)$$

$$h'_m(f, t) \cdot e^{j\Delta\theta_m} = S(f) \cdot \|H_m(f)\| \cdot e^{j\Delta\theta_m}, \quad (16)$$

where $S(f)$ and $R(f)$ respectively represent the spectrum of the transmitted and received signal, $H(f)$ is the channel response. combining equation (14) ~ (16), we can rewrite equation (12) as:

$$\|H(f)\| \cdot e^{j\Delta\theta_{R_1}} = \|H_l(f)\| \cdot e^{j\Delta\theta_l} \cdot e^{-K_l \cdot D - j \cdot K_l \cdot D} + \|H_m(f)\| \cdot e^{j\Delta\theta_m}, \quad (17)$$

$$e^{-K_l \cdot D - j \cdot K_l \cdot D} = \frac{\|H(f)\| \cdot e^{j\Delta\theta_{R_1}} - \|H_m(f)\| \cdot e^{j\Delta\theta_m}}{\|H_l(f)\| \cdot e^{j\Delta\theta_l}}, \quad (18)$$

let the left side of the equation be ζ ,

$$\zeta = e^{-K_l \cdot D - j \cdot K_l \cdot D}. \quad (19)$$

Obviously, ζ is a variable that is only correlated to the material itself, different material has different ζ .

B. Fitted Correlation Coefficient

$\|H_m(f)\| \cdot e^{j\Delta\theta_m}$ and $\|H_l(f)\| \cdot e^{j\Delta\theta_l}$ in (18) are fixed values under the premise that the environment remains unchanged. $\|H(f)\| \cdot e^{j\Delta\theta_{R_1}}$ can be extracted from the measured CSI data. We can't get $\|H_m(f)\| \cdot e^{j\Delta\theta_m}$ and $\|H_l(f)\| \cdot e^{j\Delta\theta_l}$ directly, so we have to substitute the materials of known K_l and K_m into (18) to fit them.

It can be seen that relative dielectric constant and conductivity of different concentrations of alcohol aqueous solutions from [14] are shown in Table 1. These data are measured by professional hardware at 5.32 GHz frequency.

The experimental materials included a 210 mm × 148 mm × 2 mm storage box and an alcohol aqueous solution with a concentration of 0% - 90% and an interval of 10%. Place a

TABLE I
MEASUREMENT RESULTS FOR DIFFERENT ABVs OF ETHANOL/WATER MIXTURES

ABV	ϵ_r	σ
0%	73.38	6.41
10%	57.12	8.33
20%	50.89	8.64
30%	40.64	8.57
40%	30.66	7.71
50%	24.74	6.82
60%	18.48	5.54
70%	13.72	4.32
80%	9.93	3.15
90%	6.85	2.02

storage box filled with various concentrations of alcohol/water solutions in the center of T_X/R_{X_1} . We only use the average amplitude and average phase of the 16th subcarrier difference since the CSI data of the 16th subcarrier is closer to the center frequency of 5.32 GHz. Then we do least-squares fitting to these ten sets of data to calculate the correlation coefficient.

C. Data Preprocessing Module

The data preprocessing module includes outlier detection and noise cancellation to calibrate the captured CSI data.

- Outlier Detection: There are usually some outliers in the captured CSI amplitude and phase difference tracking. Anomaly detection is performed to detect bad data values which should be replaced from the original CSI data. WiMate adopts the Pauta criterion method to detect and remove outliers.
- Noise Denoising: The raw channel data may contain abnormal samples caused by background noise or hardware glitches. Therefore, before applying feature extraction techniques, we chose Butterworth Filters as in [15] to further eliminate ambient noise. The sampling rate of CSI is $F_s = 1000$ samples/s and the cut-off frequency is $\omega_c = \frac{2\pi \cdot f}{F_s} = \frac{2\pi \cdot 15}{1000} = 9.42$ rad/s.

D. Feature Extraction

In this module, the amplitude and phase difference data of CSI data that have been preprocessed are divided into a group of 1000 data packets, and the average value within the group is calculated as the amplitude and phase difference value of the group. Then the average value within the group is substituted into (19) to calculate the feature value within the group.

Six kinds of materials in the preliminary experiment were placed at the distance of T_X 4cm, 8cm, 10cm, 12cm and 16cm respectively. Explore whether the characteristic values of the same material are consistent in different locations. The result is shown in Fig. 5. It can be seen from the figure that the eigenvalues of the same material have very little difference at different positions. In Fig. 5, solid and dashed lines represent nonmetallic and metallic materials, respectively. The difference between the two types is extremely obvious. It can be seen that the characteristic values proposed in this paper can distinguish different types of materials better.

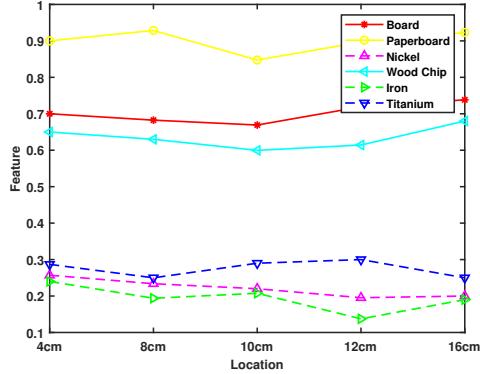


Fig. 5. Feature value of six materials at different locations.

IV. PERFORMANCE EVALUATION

This section systematically evaluates the performance of WiMate via real-world experiments.

A. Hardware Setup

The system is realized by using off-the-shelf hardware equipment. Two miniPCs are used as sending and receiving equipment, with the transmitter having one antenna and the receiver having two antennas.

Our antenna setup is shown in Fig. 6. The distance between the transmitting antenna and the receiving antenna is 20 cm and they are all placed horizontally above the floor. The material to be tested is placed on the LOS path of T_X and R_{X_1} . Our transmission equipment has a transmission speed of 1000 packets per second. When we collected the data, there were no other irrelevant people present.

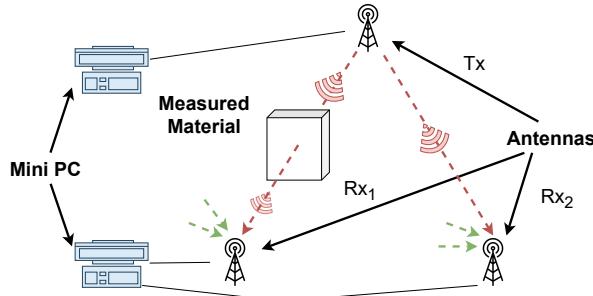


Fig. 6. Antenna setup.

B. Experimental Environment and Data Collection

In our experiment, the test and training data sets were collected from the environment shown in Fig. 1. Since the design goal of our system is to identify different materials, we conducted experiments of 30 seconds for six different materials (the same as the previous experiment) at 5 different positions, and the first 5 seconds of this 30 s will not be used on experimental data.

Since the experimental material is static, we take every 1000 data packets as a group to calculate the feature value of the material. Finally, we collected 6 (number of materials) \times 5 (number of locations) \times 25 (experimental group) data files. We use SVM to classify the processed CSI data. The data were randomly divided into two groups for training and testing to observe the classification effect under different positions.

C. Real-world Evaluation

In this section, we use the collected 750 sets of experimental data, and randomly select 80% as the training set and 20% as the test set. The final results are shown in Fig. 7. We used SVM to achieve an average resolution of 96.2% for six materials at five locations.

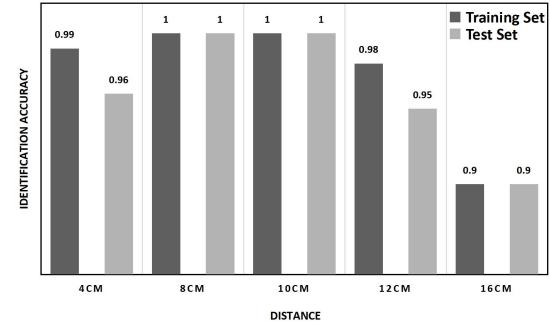


Fig. 7. Classification Accuracy.

Fig. 7 shows the identification accuracy of six materials in 5 positions. The X-axis represents the distance from the material to T_X and the Y-axis represents the accuracy. We can find that the identification accuracy of the material at 16cm is the lowest, only 90%, but the identification accuracy at other positions is as high as 100%. This is probably because the material is too far away from the transmitting antenna. In general, our system can accurately identify materials at any location of LOS. When measuring a new material in a new environment, we only need to repeat the experiment in section III to solve the influence of environmental changes, and then perform the same experiment on the new material.

V. CONCLUSION

In this paper, we demonstrate that we can use WiFi signals to identify materials, using only a pair of WiFi devices. In particular, our system uses CSI extracted from the WiFi physical layer to detect signal attenuation caused by materials and calculates an eigenvalue that is only related to the material itself but independent of the position. We have implemented a prototype system with off-the-shelf equipment, and the experimental results show that it has accurate material identification performance.

ACKNOWLEDGMENT

This work is sponsored by the National Key Research and Development Program Cyberspace Security Special Project “Research on Key Technologies for the Internet of

Things and Smart City Security Assurance” under Grant No. 2018YFB0803403.

REFERENCES

- [1] 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 *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*, 2017, pp. 288–300.
- [2] C. A. Kendziora, R. A. McGill, R. Furstenberg, M. Papanikos, V. Nguyen, and J. M. Byers, “Detecting traces of explosives,” *Spie Newsroom*, 2015.
- [3] S. Cussat-Blanc and J. Pollack, “Cracking the egg: Virtual embryogenesis of real robots,” *Artificial life*, vol. 20, no. 3, pp. 361–383, 2014.
- [4] K. Yin, J. Chen, F. Wang, X. Sun, W. Ren, and Y. Wang, “Study on identification of winding material in distribution transformer based on computed tomography,” in *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*. IEEE, 2016, pp. 1034–1037.
- [5] K. Ohtani and M. Baba, “A simple identification method for object shapes and materials using an ultrasonic sensor array,” in *2006 IEEE Instrumentation and Measurement Technology Conference Proceedings*. IEEE, 2006, pp. 2138–2143.
- [6] H. Zhou, Y. X. 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: official journal of the Korean Radiological Society*, vol. 15, no. 2, pp. 218–225, 1900.
- [7] U. Ha, Y. Ma, Z. Zhong, T.-M. Hsu, and F. Adib, “Learning food quality and safety from wireless stickers,” in *Proceedings of the 17th ACM Workshop on Hot Topics in Networks*, 2018, pp. 106–112.
- [8] R. Suwalak, C. Phongcharoenpanich, D. Torrungrueng, and P. Akkaraekthalin, “Dielectric material determination using the radar equation in rfid sensor applications,” in *2015 IEEE Conference on Antenna Measurements Applications (CAMA)*, 2015, pp. 1–4.
- [9] Y. Gu, X. Zhang, Z. Liu, and F. Ren, “Bessense: leveraging wifi channel data and computational intelligence for behavior analysis,” *IEEE Computational Intelligence Magazine*, vol. 14, no. 4, pp. 31–41, 2019.
- [10] Y. Gu, X. Zhang, C. Li, F. Ren, J. Li, and Z. Liu, “Your wifi knows how you behave: Leveraging wifi channel data for behavior analysis,” in *2018 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2018, pp. 1–6.
- [11] W. Yang, X. Wang, A. Song, and S. Mao, “Wi-wheat: Contact-free wheat moisture detection with commodity wifi,” in *2018 IEEE International Conference on Communications (ICC)*. IEEE, 2018, pp. 1–6.
- [12] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.
- [13] L. Tsang, J. A. Kong, and R. T. Shin, “Theory of microwave remote sensing,” 1985.
- [14] H. Song, B. Wei, Q. Yu, X. Xiao, and T. Kikkawa, “Wieps: Measurement of dielectric property with commodity wifi device—an application to ethanol/water mixture,” *IEEE Internet of Things Journal*, vol. 7, no. 12, pp. 11 667–11 677, 2020.
- [15] K. Ali, A. X. Liu, W. Wei, and M. Shahzad, “Keystroke recognition using wifi signals,” in *ACM MobiCom*, 2015.