 doi:10.20944/preprints202103.0434.v1

figuresection

 doi:10.20944/preprints202103.0434.v1

Article

Device-Free Indoor Location Estimation System Using Commodity Wireless LANs

Yuan Zhou, Minseok Kim, Hideaki Momose, Satoru Yasukawa

Graduate School of Science and Technology, Niigata University, Niigata 950-3198, Japan
* Correspondence: mskim@eng.niigata-u.ac.jp

Abstract: In recent years, since the propagation channel characteristics have been effectively used for applications such as motion sensing, position detection, etc. A great deal of attention is attracted to channel sounding methods easy to utilize using low-cost devices. This paper presents a device-free indoor location estimation method using spatio-temporal features of radio propagation channels using the 2.4-GHz band 3-by-3 MIMO channel sounder developed using commodity wireless LANs. The measurement results demonstrated a reasonable performance of the proposed method with small number of antennas.

Keywords: WiFi sounder; CSI; MIMO; indoor location estimation; array signal processing; machine learning; SVM

1. Introduction

The technical development of Internet of Things (IoT) has grown people's demand for location services. In the IoT era, position information plays an important role. Among them, indoor positioning as a service closely related to people's lives has attracted a great deal of attention. Indoor positioning system (IPS) can be widely used such as home surveillance system, nursing care for the elderly peoples and patient monitoring in hospitals, etc [1]. Most of existing surveillance systems use cameras to achieve the purpose of monitoring [2]. However, it has its own problems of insufficient coverage and blind area. In order to improve the performance, it needs to increase the number of cameras, but it also brings a high cost and complex configuration. In addition, privacy is still a problem that cannot be ignored.

To cope with these problems, various methods utilizing radio waves have been taken into consideration. By capturing the fluctuation of wireless communication channel characteristics due to a person's movement and applying machine learning, the estimation of a person's position and behavior pattern can be achieved. However, the positioning in an indoor environment still has many technical challenges owing to different propagation condition from outdoor applications [3]. For example, due to signal fading and multi-path effects caused by building structures, commonly used global positioning system (GPS) do not work effectively (low accuracy, bad signal, etc) in an indoor environment. Therefore, we put our attention to those wireless devices that has already been widely deployed, such as WiFi (WLAN using IEEE802.11 standards). WiFi has been widely used in homes, hotels, cafes, airports, shopping malls, and other various types of large or small buildings, which makes WiFi become very attractive in indoor positioning applications. It is expected that the universality and affordability of WiFi devices enable the popularization of indoor localization systems.

Recently, there is various research on indoor localization system using WiFi device have been reported. For instance, IPS using fingerprint method of received signal strength indication (RSSI) have been discussed [5][6]. However, due to the temporal and spatial variation of the indoor environment [7], several access points (AP) is required to improve the accuracy, and even more, the configuration should become complex [8]. Channel State Information (CSI) based IPS have been developed [9]-[12]. However, due to the WiFi device's limited bandwidth, it requires several APs to maintain the performance of the system [13]. Moreover, a microwave system using an antenna array also had already been developed

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1

[14] where some predetermined events are identified monitoring the fluctuation of the signal subspace spanned by the eigenvectors. Because the eigenvector represents the spatial structure of the multi-path propagation, spatial filtering by first eigenvector can reduce the effects of noise, but a large number of antennas are needed to resolve the multi-path components more precisely. In device-free localization (DFL) techniques the performance is significantly influenced by the strong line-of-sight path, and is apt to biased depending on the person's location [15].

To cope with the mentioned problems, this paper developed an indoor DFL system based on commodity WiFi devices which achieves a reasonable performance by using machine learning of spatio-temporal features of wireless channels. The technical contributions of this paper is as following. First, a 2.4-GHz band 3×3 multiple-input-multiple-output (MIMO) channel sounder was developed where the channel bandwidth was extended to approximately 68 MHz by concatenating CSIs taken at six consecutive WiFi channels to achieve high resolution in delay time domain. Second, we developed an indoor location estimation method using spatio-temporal features of the multi-path propagation characteristics which are obtained by separating the multi-path components into three angle-taps and two delay-tap. That can solve the problem that the performance is largely influenced by the change in the line-of-sight (LoS) path. Third, the support vector machine (SVM) is applied to identify a person's location.

The remainder of this work is organized as follows. In Section 2, a channel sounding system based on IEEE 802.11n wireless LANs is presented. Then, we present a indoor location estimation method in section 3. The identification results of the sub-area where a person is present via two measurement campaigns is presented in section 4. It demonstrate the performance of the proposed system and shows its feasibility. Finally, Section 5 concludes the paper.

2. Channel Sounding Using IEEE 802.11n Wireless LANs

2.1. IEEE 802.11n

IEEE 802.11n has been standardized [22] as a MIMO-OFDM based wireless LAN standard, which is the successor to IEEE 802.11a/g. For the backward compatibility of existing standards, it provides the mixed mode [23] to achieve MIMO transmission extending the legacy mode. The frame format of mixed mode is shown in Fig.1. After L-SIG (Legacy Signal Field) that sends wireless LAN frame information such as the total length of the wireless LAN frame, it will send HT-SIG (High Throughput Signal Field) that includes transmission parameters for spatial multiplexing transmission. After then using HT-STF (High Throughput Short Training Field) and HT-LTF (High Throughput Long Training Field) it estimates CSI which is necessary for spatial multiplexing transmission. In IEEE 802.11n, the number of sub-carriers over 20 MHz bandwidth is increased from 54 to 56. Furthermore, it provides 40 MHz mode using channel bonding as optional. In consequence, by 4-stream spatial multiplexing with 40 MHz bandwidth, it can achieve transmission throughput up to 600 Mbps (64QAM, code rate 5/6).

Figure 1. Frame format of IEEE 802.11n hybrid mode

2.2. CSI Acquisition

In this paper, we use an open-source software package called CSI tool [18] to built up our own system. In IEEE 802.11n, the CSI was estimated in the receiver and then feedback to the transmitter, so the transmitter can use the CSI to complete calibration and beamforming. In other words, by setting

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1

the CSI estimate flag in the transmit packet, we can manage the receiver to report the CSI estimation to the transmitter. The CSI tool is a device driver for the Atheros network interface cards (NIC), `ath9k` (kernel module), to extract the CSI using IEEE 802.11n. In the CSI tool, the transmitter repetitively sends the CSI estimate flag packet for CSI acquisition, and the receiver reports the CSI acquisition result to the user space program whenever the CSI is acquired. Here, the CSIs acquired for every combination of transmit and receiver antennas are estimated as complex numbers for 56 sub-carriers, which include transmitter/receiver characteristics, propagation channel characteristics, and antenna characteristics.

The CSI report in one packet transmitted with bandwidth of 20 MHz is obtained in the form of a complex matrix of size $N_R \times N_T \times N_C$ if the number of antennas at the receiver and transmitter is N_R and N_T, respectively. In addition, some information such as the timestamp, channel number, transmission rate, number of antennas (N_R, N_T), number of sub-carriers N_C, noise floor, PHY error, received signal strength indicator (RSSI) of every antenna at the receiver, and length of payload, is added. Since the real and imaginary parts of the CSI are each represented by 10 bits, their amplitude is automatically scaled according to the magnitude of the received power. As well, the phase of CSI fluctuates independently for each packet, but the amount of offset is common to all sub-carriers (common phase offset).

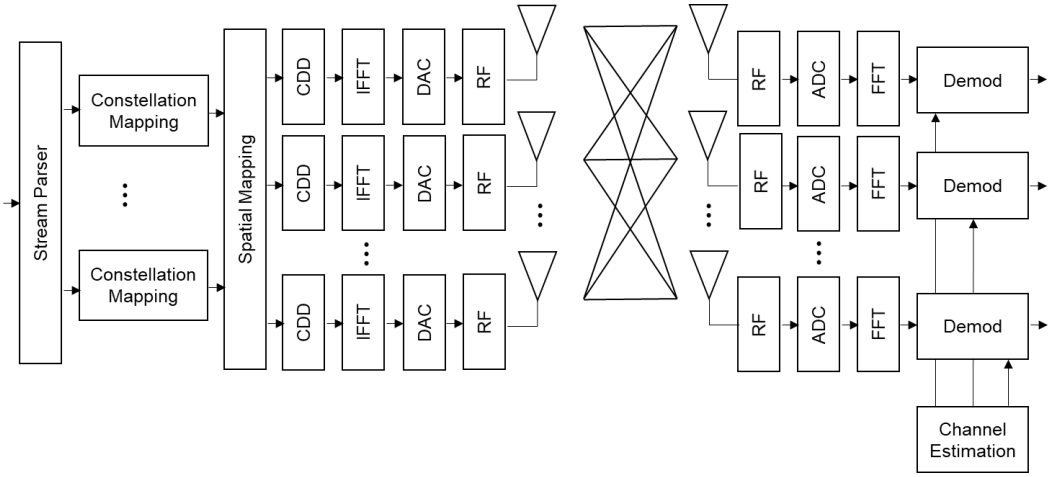


Figure 2. IEEE 802.11n transceiver configuration.

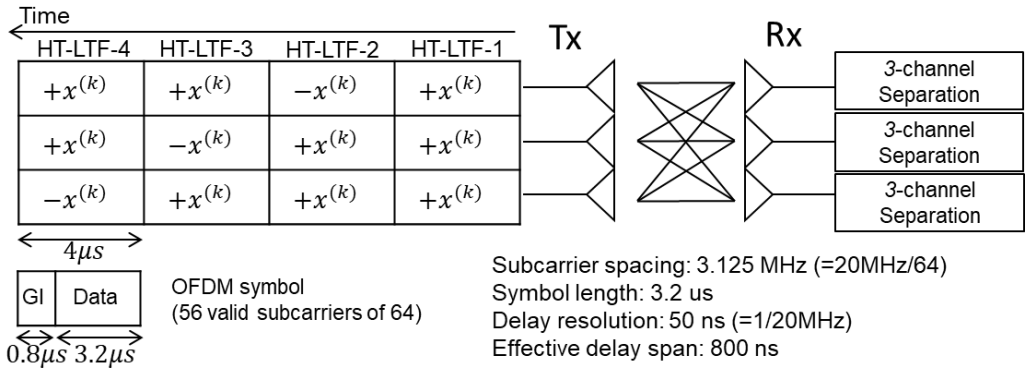


Figure 3. 3×3 MIMO channel sounding.

2.3. MIMO Channel Sounding

Fig. 2 shows the MIMO transceiver configuration of IEEE 802.11n wireless LAN. The stream parser creates N_{st} spatial streams from the encoded bit-stream, and quadrature modulated to symbol

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1

stream by constellation mapping. The transmitter can send N_{st} ($\leq \text{rank}(H) \leq \min(N_\text{T}, N_\text{R})$) parallel streams, where N_T and N_R denote the number of antennas in transmitter and receiver, respectively. Then, the spatial streams is shifted cyclically by cyclic delay diversity (CDD), which can extend the communication area by transmit the same signal at different carrier frequencies. It prevent beamforming that might be made in specific direction when all the antennas are sending the signal with common header information. Spatial mapping (SM) is performed when the number of antennas is larger than the number of spatial streams ($N_T > N_{st}$). The frequency samples of each transmission branch are transformed into time domain signals by inverse Fourier transform and are transmitted simultaneously by all antennas.

The MIMO propagation channel matrix $\boldsymbol{H}$ is estimated by using the training signal (HT-LTF), which is contained in the frame format as shown in Fig. 1. Fig. 3 shows the block diagram of the 3×3 MIMO channel sounding system. Assume the number of spatial streams and the number of transmit antennas are the same ($N_\text{T} = N_\text{st} = 3$), the received signal vector at the kth ($k = 1, \ldots, N_c = 56$) sub-carrier is expressed as

$$\boldsymbol{y}^{(k)} = \boldsymbol{G}_\text{agc}\boldsymbol{H}^{(k)}\boldsymbol{\Phi}^{(k)}\boldsymbol{d}x^{(k)} + \boldsymbol{n}^{(k)} \in \mathbb{C}^{3\times 1} \tag{2.1}$$

where $\boldsymbol{G}_\text{agc}$, $\boldsymbol{H}^{(k)}$, and $\boldsymbol{\Phi}^{(k)}$ denote the gain of automatic gain control (AGC), the MIMO channel matrix, and the CDD matrix, respectively. They are expressed as

$$\boldsymbol{G}_\text{agc} = \text{diag}\left(\begin{bmatrix} \sqrt{G_{\text{agc},1}} & \sqrt{G_{\text{agc},2}} & \sqrt{G_{\text{agc},3}} \end{bmatrix}^T\right), \tag{2.2}$$

$$\boldsymbol{H}^{(k)} = \begin{bmatrix} H_{11}^{(k)} & H_{12}^{(k)} & H_{13}^{(k)} \\ H_{21}^{(k)} & H_{22}^{(k)} & H_{23}^{(k)} \\ H_{31}^{(k)} & H_{32}^{(k)} & H_{33}^{(k)} \end{bmatrix}, \tag{2.3}$$

$$\boldsymbol{\Phi}^{(k)} = \text{diag}\left(\begin{bmatrix} 1 & e^{-\text{j}\frac{2\pi}{N_\text{f}}\delta_2 k} & e^{-\text{j}\frac{2\pi}{N_\text{f}}\delta_3 k} \end{bmatrix}^T\right), \tag{2.4}$$

where $x^{(k)}$ is the transmitted signal and $\boldsymbol{n}^{(k)} = \begin{bmatrix} n_1^{(k)} & n_2^{(k)} & n_3^{(k)} \end{bmatrix}^T$ is the noise vector. Here, δ_2 and δ_3 are specified as $\delta_2 = 8$ (400 ns) and $\delta_3 = 4$ (200 ns), and N_f is the number of FFT points and set to be 64.

In order to separate the transmitted signals from multiple transmitting antennas, the signal $x^{(k)}$ is transmitted 4 times with each transmitting antenna by set the $\boldsymbol{d}$ as $\boldsymbol{d}[1] = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$, $\boldsymbol{d}[2] = \begin{bmatrix} -1 & 1 & 1 \end{bmatrix}^T$, $\boldsymbol{d}[3] = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}^T$, $\boldsymbol{d}[4] = \begin{bmatrix} 1 & 1 & -1 \end{bmatrix}^T$ in Eq. (2.1), then the matrix of the received signal can be obtained as

$$\boldsymbol{Y}^{(k)} = \boldsymbol{G}_\text{agc}\boldsymbol{H}^{(k)}\boldsymbol{\Phi}^{(k)}\boldsymbol{D}x^{(k)} + \boldsymbol{N}^{(k)} \in \mathbb{C}^{3\times 4}, \tag{2.5}$$

where

$$\boldsymbol{Y}^{(k)} = \begin{bmatrix} \boldsymbol{y}^{(k)}[1] & \boldsymbol{y}^{(k)}[2] & \boldsymbol{y}^{(k)}[3] & \boldsymbol{y}^{(k)}[4] \end{bmatrix}, \tag{2.6}$$

$$\boldsymbol{N}^{(k)} = \begin{bmatrix} \boldsymbol{n}^{(k)}[1] & \boldsymbol{n}^{(k)}[2] & \boldsymbol{n}^{(k)}[3] & \boldsymbol{n}^{(k)}[4] \end{bmatrix}, \tag{2.7}$$

and

$$\boldsymbol{D} = \begin{bmatrix} 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}. \tag{2.8}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1

Then the CSI is obtained as follows based on Eq. (2.5).

$$\hat{\boldsymbol{H}}_{\mathrm{CSI}}^{(k)} = \boldsymbol{Y}^{(k)}\boldsymbol{D}^{-1}/x^{(k)}. \tag{2.9}$$

It is necessary to remove the effects of the amplifier and CDD in order to obtained the MIMO channel matrix, which expressed as

$$\hat{\boldsymbol{H}}^{(k)} = \boldsymbol{G}_{\mathrm{agc}}^{-1}\hat{\boldsymbol{H}}_{\mathrm{CSI}}^{(k)}\bar{\boldsymbol{\Phi}}^{(k)}. \tag{2.10}$$

However, due to nonlinear distortions of transmitter and receiver circuits contained in CSI, further calibration to Eq. (2.10) is required.

2.4. Developed System

The bandwidth of IEEE 802.11n is basically 20 MHz with an optional 40 MHz mode. Therefore the delay time resolution is limited to 50 ns (= 1/20 MHz) theoretically, which is not sufficient in order to apply WiFi to sensing with fine variation of the propagation path. The bandwidth expansion with channel bonding has been proposed [17] and the toolkit has been published as open-sources [20]. Utilizing this toolkit, the channel bandwidth could be extended to approximately 67.8125 MHz (delay resolution: approximately 15 ns, distance resolution: approximately 4.5 m) by concatenating the CSIs taken at six consecutive WiFi channels as shown in Fig. 4.

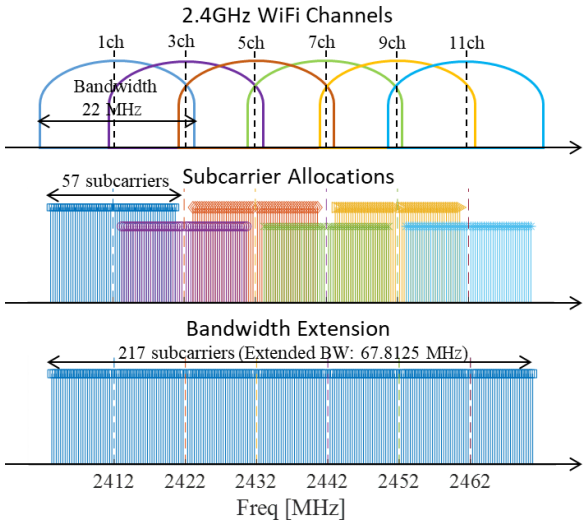
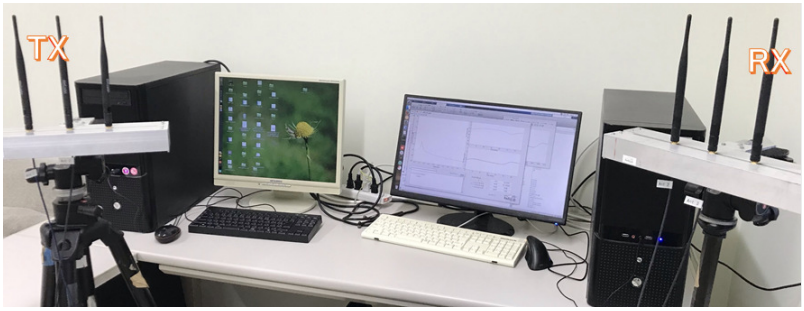


Figure 4. Bandwidth extension by 6 channel boding

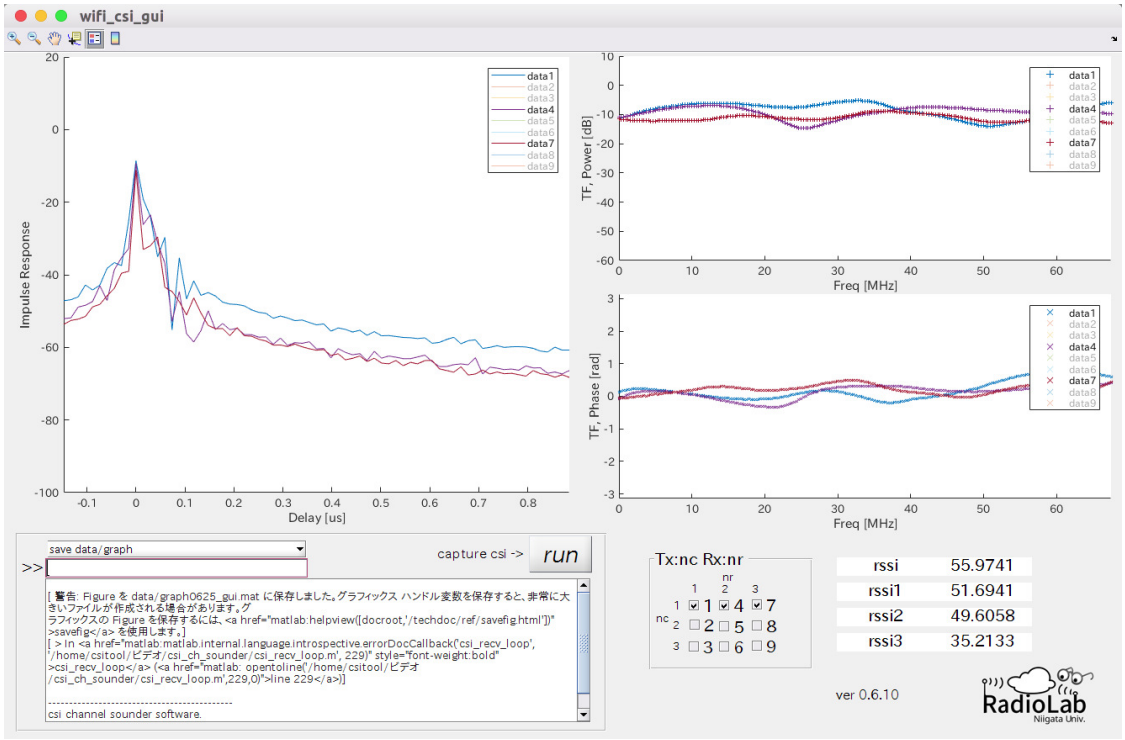
Table 1. System setup

Item	Values
Hardware	Linux (kernel 4.1 Modified version) PC × 2 (equipped with Rabortw AR9380 WiFi card)
Software	AP(transmitter) : hostapd, hostapd_cli STA(receiver) : WPA_supplicant
WiFi channels	1 ch (2412 MHz), 3 ch (2422 MHz), 5 ch (2432 MHz), 7 ch (2442 MHz), 9 ch (2452 MHz), 11 ch (2462 MHz)
Sub-carriers	217
Extend bandwidth	67.8125 MHz (bonded by 6 channel)
Beacon interval	15 ms (usually, 100 ms)
Packet transmission method	Send 50 packets in a row and wait for 50 ms Repeat while switching channels

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1



(a) hardware setup



(b) software

Figure 5. Developed channel sounding system.

2.4.1. System Configuration

The system specification is as shown in Table 1. Two Linux PCs equipped with an Atheros AR9380 WLAN card are used for a transmitter and a receiver as shown in Fig. 5. In order to extend the bandwidth, 50 packets are transmitted and received continuously at every channel, while sweeping the channel from 1 to 11 at two channel interval. Separately averaging the amplitude and phase of the CSI value at each sub-carrier filters out abrupt fluctuation and improves the signal-to-noise power ratio (SNR). Here, the beacon interval was set to the minimum value of 15 ms for fast channel switching.

2.4.2. Back-To-Back Calibration

The phase variation of CSI occurs due to hardware imperfection such as nonlinear distortion of amplifiers, power control uncertainty, phase ambiguity, etc [16,17], which should be removed to obtain the genuine propagation channel characteristics. However, since the six CSIs taken at consecutive WiFi channel are concatenated to extend the bandwidth, it is necessary to calibrate the inter-channel discontinuity in amplitude and phase values.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1

In order to remove the hardware imperfection contained in CSI, we measure the system characteristics including the hardware imperfection in advance by back-to-back measurement that is conducted connecting the transmit and receive antenna ports directly via a RF cable with known characteristics, then obtain the channel transfer function dividing the CSI by the system characteristics, which is called back-to-back calibration (B2B calibration). Here, the CSI $\hat{H}^{(k)}_{\mathrm{CSI,B2B},j,i}$ is obtained by connecting the ith transmit and jth receive antenna ports directly with a cable, and is performed for every antenna combinations. The MIMO channel matrix can be obtained by

$$\hat{H}^{(k)}_{j,i} = \frac{\hat{H}^{(k)}_{\mathrm{CSI},j,i} / \sqrt{G_{\mathrm{agc},j}}}{\hat{H}^{(k)}_{\mathrm{CSI,B2B},j,i} / \sqrt{G_{\mathrm{agc,B2B},j}}} e^{\mathrm{j}\frac{2\pi}{N_{\mathrm{f}}}\delta_i k}, \tag{2.11}$$

where $G_{\mathrm{agc},j}$ denotes the gain of the AGC amplifier of the jth receive branch. In commodity WiFi devices, the AGC amplifier operates automatically according to received signal level to keep the signal power at a certain level, thus the CSI values are scaled as in Eq. (2.11). The gain of the AGC amplifier can be defined referring to the amplitude of the CSI and the RSSI as

$$G_{\mathrm{agc},j} = \frac{1}{P_j} 10^{\frac{\mathrm{RSSI}_j}{10}}, \tag{2.12}$$

where $P_j = E\left[|y^{(k)}_j|^2\right]$.

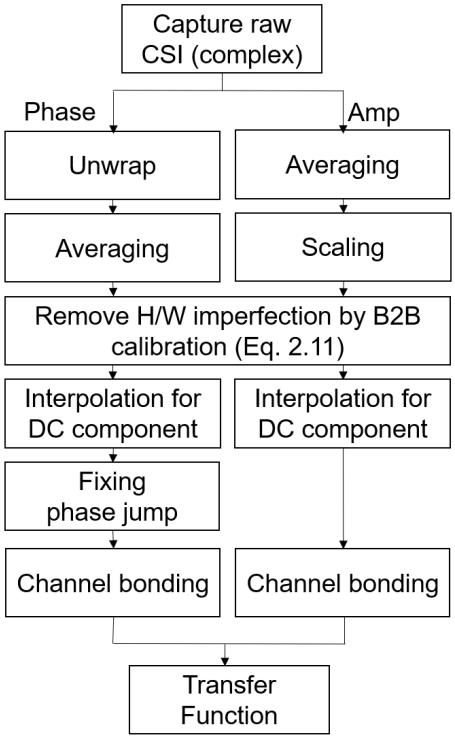
The B2B calibration allows the CSIs acquired on multiple channels to be connected continuously because the phase and amplitude should become zero and unit after calibration. However, phase jump occurs between adjacent channels due to independent random offset of CSI phase characteristics in every packet. It can be removed by calculating the inter-channel phase difference using overlapped sub-carriers in adjacent channels. The amplitude of each channel is also fluctuated due to the AGC gain uncertainty. Using the averaged values of the overlapped sub-carriers, the amplitudes are smoothly concatenated. The calibration procedure is shown in Fig. 6(a) and Fig.6(b).

2.5. Evaluation

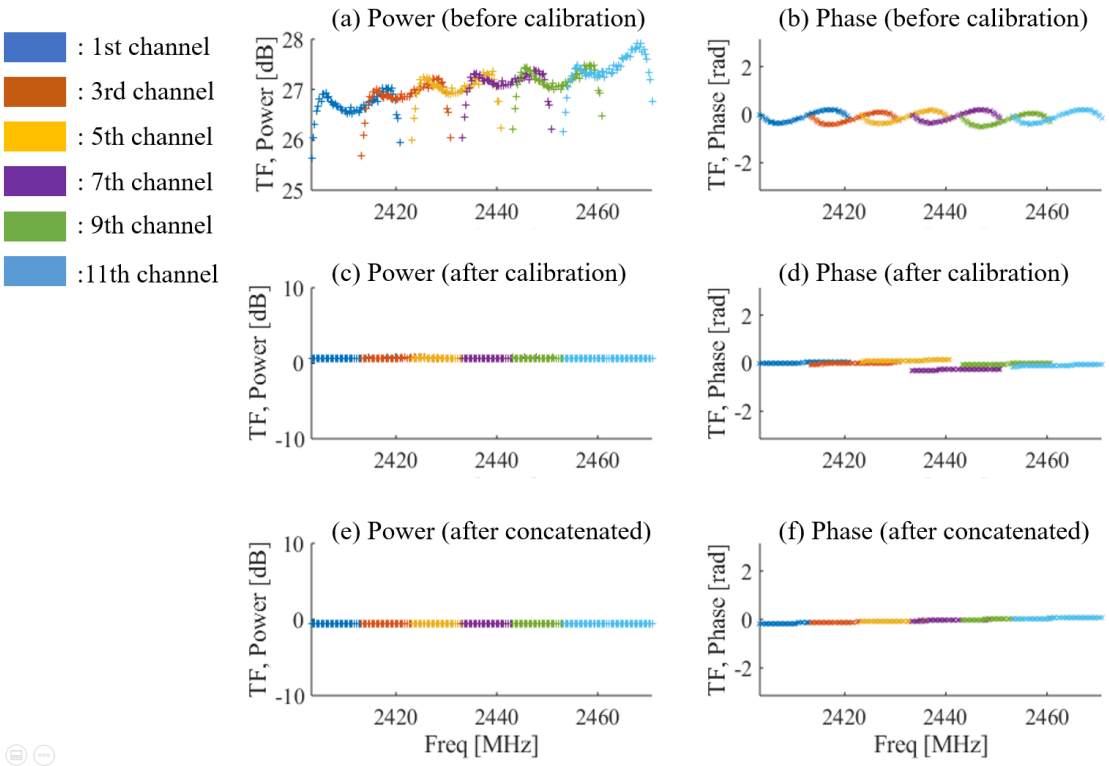
The channel impulse response test measurement is conducted for system validation purpose. The unmatched T-junction (BL41-6203-00, Orient Microwave Corp.) [24] was used in the measurement. Fig. 7(a) shows the measurement configuration. The transmit antenna port 1 (`Tx1`) was connected to the power splitter. The port 1 of the splitter was connected to the receive antenna port 1 (`Rx1`) which is used for the reference measurement. The port 2 of the splitter was connected to the port A of the T-junction, and the signal from the port B was input to the receive antenna port 2 (`Rx2`). Here, the port C is connected with a 6 m length cable left the end opened, which generates reflected waves at the cable end. The attenuation per 1 m of the used cable (Enviroflex_316, Huber Shuner, Corp.) [24] is approximately 1.62 dB/m. The measurement results are shown in Fig. 7(b). It is seen that only direct waves are observed in `Tx1-Rx1`, and the first reflected wave from the cable end also can be observed in `Tx1-DUT-Rx2`, which is well agreed with the calculated value. Here, it notes that it was calculated supposing to the fact that the propagation speed of electromagnetic waves in coaxial cables is about 77 % of that in a vacuum, and the path length and attenuation rate of the cable (delay time: $\frac{12}{3\times10^8\times0.77} = 52$ ns, power attenuation: $12 \times 1.65 = 20$ dB).

In order to evaluate the inter-channel phase stability, the Tx1 signal was distributed to `Rx1`, `Rx2`, and `Rx3` using the splitter. After B2B calibration, 210 channel transfer functions were obtained. The variation of $\angle\hat{H}_{2,1}$ and $\angle\hat{H}_{3,1}$ based on the reference value of $\hat{H}_{1,1}$ is shown in Fig. 8. Here, the left figure shows the relative phase difference of 217 sub-carriers along 210 time samples, and the right figure shows the phase difference of all sub-carriers acquired at 210 time samples. It is seen that the phase variation is less than approximately 20 degrees.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1



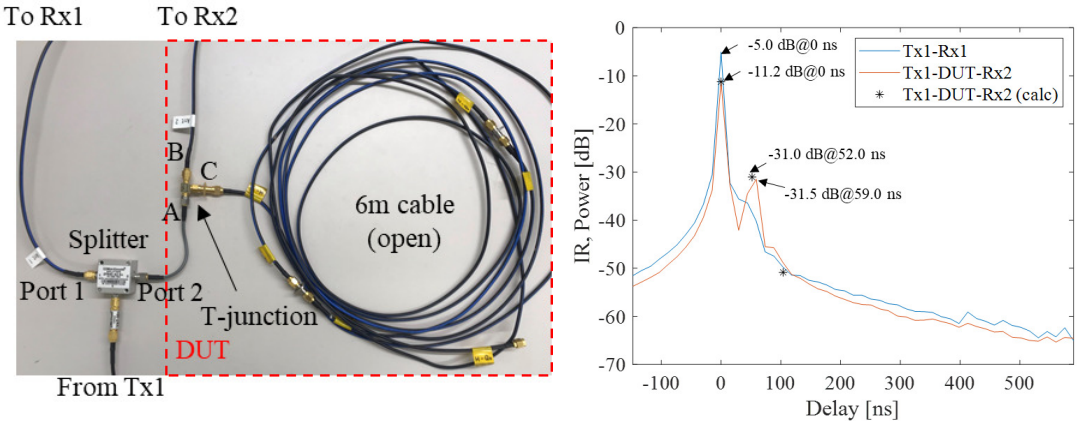
(a) Calibration procedure



(b) Calibration result

Figure 6. B2B calibration

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1



(a) DUT test setup (b) DUT test result

Figure 7. DUT test

3. DF Indoor Location Estimation

The DF indoor location estimation system was developed by using the WiFi sounder that described above. The receiving antennas are arranged linearly at about 6.3 cm intervals (half-wavelength of the minimum frequency 2.403 GHz). The bandwidth broadened channel impulse responses are obtained by concatenating the CSIs taken at six consecutive WiFi channels to extend the bandwidth to approximately 68 MHz (delay resolution: approximately 15 ns, distance resolution: approximately 4.5 m).

3.1. Experiment Scenario

The experiment was conducted in a small office and a medium sized conference room. The measurement specifications are shown in Table 2. Note that the WiFi communication is performed between the access point (AP) and station (STA) fixed in the environment, and a target person moving in the environment does not equip with any device. Here, the CSI is acquired 10 times for each target position and averaged to reduce the effects of noise.

Table 2. Measurement Specifications

Room type	Small office	Conference room
Size	4.0×4.6 (m^2)	7.8×8.8 (m^2)
Bandwidth	67.813 MHz (6 channel bonding)	
Transmitting antennas	1	2
Receiving antennas	3-element linear array	
Element pattern	Omni-directional	

3.1.1. Small Office

The room model of small office and the position of antenna is as shown in Fig. 9(a) and Fig. 9(b). The receiving and transmitting antennas are placed at a height of 0.9 m and 1.9 m, respectively. For this environment, 30 CSIs were captured while a person moved randomly in a sub-area, and thus, 180 samples were captured in total.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1

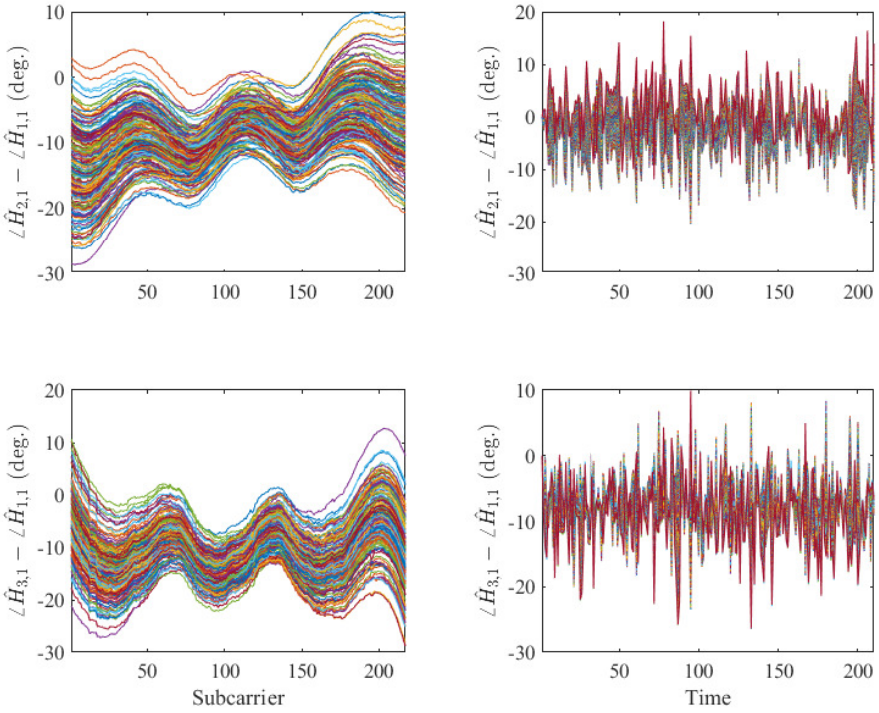
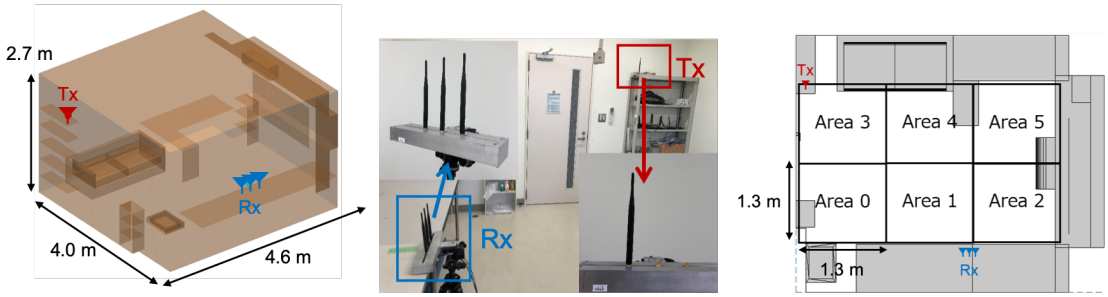


Figure 8. Phase stability



(a) Room model (b) Antenna position (c) Defined area

Figure 9. Small office condition

3.1.2. Conference Room

The room model of conference room and the position of each antenna is as shown in Fig. 10(a) and Fig. 10(b). The receiving antenna is a 3-element linear array and placed at a height of 2.0 m. One of the transmitting antenna (Tx1) is placed at a height of 0.9 m inside the room and antenna (Tx2) is placed at a height of 1.6 m outside the room. For this environment, the CSI was captured at each position during a person moves along the predetermined route. Here, the person moved at 0.5 m interval with random orientation and the 151 CSIs were captured in total.

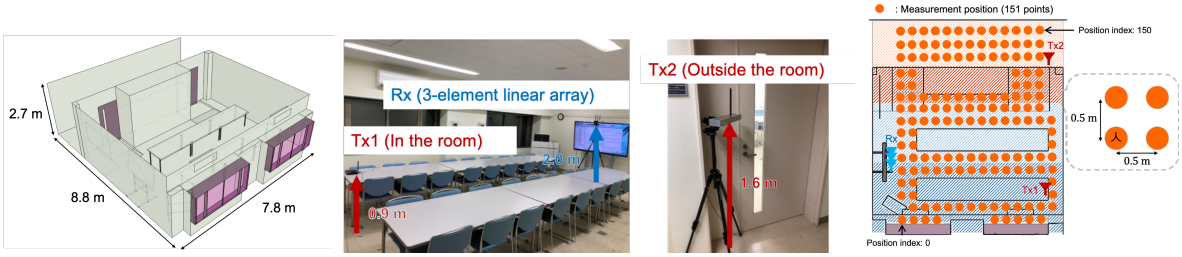
3.2. Signal Processing

The array output $y(t)$ of K-element linear array is expressed by

$$y(t) = \boldsymbol{w}^H \boldsymbol{x}(t), \tag{3.1}$$

$$\boldsymbol{x}(t) = [x_1(t), x_2(t), \cdots, x_K(t)]^T, \tag{3.2}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1



(a) Room model (b) Antenna position (c) Human position

Figure 10. Conference room condition

$$\boldsymbol{w} = [w_1, w_2, \cdots, w_K]^T. \tag{3.3}$$

Where $[\cdot]^H$ denotes the complex conjugate transpose, $\boldsymbol{x}(t)$ and $\boldsymbol{w}$ denote the input signal and the weight of each antenna, respectively. Then, the output power can be expressed as

$$P_{out} = \frac{1}{2}E\left[|y(t)|^2\right] = \frac{1}{2}\boldsymbol{w}^H\boldsymbol{R}_{xx}\boldsymbol{w}. \tag{3.4}$$

where $\boldsymbol{R}_{xx}$ denotes the correlation matrix of the received signal and is expressed as

$$\boldsymbol{R}_{xx} = E\left[\boldsymbol{x}(t)\boldsymbol{x}^H(t)\right]. \tag{3.5}$$

To direct the main beam of the array antenna to an arbitrary angle ϕ, each weight expressed in the Eq. (3.3) can be set as

$$w_k = \exp\left(-\mathrm{j}\frac{2\pi}{\lambda}d_k\sin\phi\right), \tag{3.6}$$

then the weight vector can be expressed as

$$\boldsymbol{w} = \left[\exp\left(-\mathrm{j}\frac{2\pi}{\lambda}d_1\sin\phi\right), \cdots, \exp\left(-\mathrm{j}\frac{2\pi}{\lambda}d_K\sin\phi\right)\right]^T \equiv \boldsymbol{a}(\phi). \tag{3.7}$$

Here, $\boldsymbol{a}(\phi)$ denotes the mode vector and angle ϕ is a variable. Then the output power is expressed as

$$P_{out} = \frac{1}{2}\boldsymbol{a}^H(\phi)\boldsymbol{R}_{xx}\boldsymbol{a}(\phi). \tag{3.8}$$

3.2.1. Existing method

This subsection described a existing method that using spatial features. Assuming there are L waves is arriving from angles $\theta_1, \cdots, \theta_L$. According to Eq. (3.2), the array signal can be expressed as

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t), \tag{3.9}$$

$$\boldsymbol{A} = [\boldsymbol{a}(\theta_1), \cdots, \boldsymbol{a}(\theta_L)], \tag{3.10}$$

where $\boldsymbol{A}$ denotes the direction matrix and $\boldsymbol{a}(\phi)$ denotes the mode vector that defined by Eq. (3.7), $s(t)$ and $n(t)$ denote the signal vector and the internal noise, respectively. Assume internal noise component is independent in each antenna with 0 mean and σ^2 variance. Then the correlation matrix of the input signal is expressed as

$$\boldsymbol{R}_{xx} = E\left[\boldsymbol{x}(t)\boldsymbol{x}^H(t)\right] = \boldsymbol{A}\boldsymbol{R}_{ss}\boldsymbol{A}^H + \sigma^2\boldsymbol{I}. \tag{3.11}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1

The eigenvalue λ_k and eigenvector $\boldsymbol{e}_k$ $(k = 1, 2, \cdots, K)$ can be obtained via the eigenvalue decomposition (EVD) of the correlation matrix $\boldsymbol{R}_{xx}$, they satisfy the following equation:

$$\boldsymbol{R}_{xx}\boldsymbol{e}_k = \lambda_k \boldsymbol{e}_k \qquad (k = 1, 2, \cdots, K). \tag{3.12}$$

Then, the Eq. (3.11) can be expressed as

$$\boldsymbol{R}_{xx} = \boldsymbol{E}\boldsymbol{\Lambda}\boldsymbol{E}^H = \sum_{k=1}^{K} \lambda_k \boldsymbol{e}_k \boldsymbol{e}_k^H, \tag{3.13}$$

where

$$\boldsymbol{E} := [\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_K], \tag{3.14}$$

$$\boldsymbol{\Lambda} := \mathrm{diag}\{\breve{}_1, \breve{}_2, \cdots, \breve{}_\mathrm{K}\}, \tag{3.15}$$

since $\boldsymbol{R}_{xx}$ is a Hermitian matrix, the eigenvalues can be sorted as

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L \geq \lambda_{L+1} = \cdots = \lambda_K = \sigma^2, \tag{3.16}$$

because rank $\left[\boldsymbol{A}\boldsymbol{R}_{\mathrm{ss}}\boldsymbol{A}^{\mathrm{H}}\right] = 1$, the eigenvalue distribution is

$$\lambda_1 > \lambda_2 = \cdots = \lambda_L > \lambda_{L+1} = \cdots = \lambda_K = \sigma^2. \tag{3.17}$$

Thus, the space spanning by the eigenvector matrix $\boldsymbol{E}$ can be separate into signal subspace and noise subspace.

The first eigenvalue separated by the process above and the corresponding vector (first eigenvector) are used in machine learning. The feature value for machine learning [14] is defined by

$$P = |\boldsymbol{v}_0^H \cdot \boldsymbol{v}|, \tag{3.18}$$

$$Q = 1 - \frac{\lambda' - \lambda_0'}{\lambda_0'}, \tag{3.19}$$

$$\boldsymbol{v} := \boldsymbol{e}_1 \propto \boldsymbol{a}(\phi), \tag{3.20}$$

$$\lambda' := \lambda_1, \tag{3.21}$$

where $\boldsymbol{v}_0$ and λ_0' denote the first eigenvector and first eigenvalue obtained in advance, the reference. The first eigenvector is used as the optimal weight vector for maximum ratio combining to maximizes the output SNR. Therefore it is strongly influenced by the direct wave and there is a problem that the features hardly fluctuated when blocking the multi-path component with low power.

3.2.2. Proposed Method

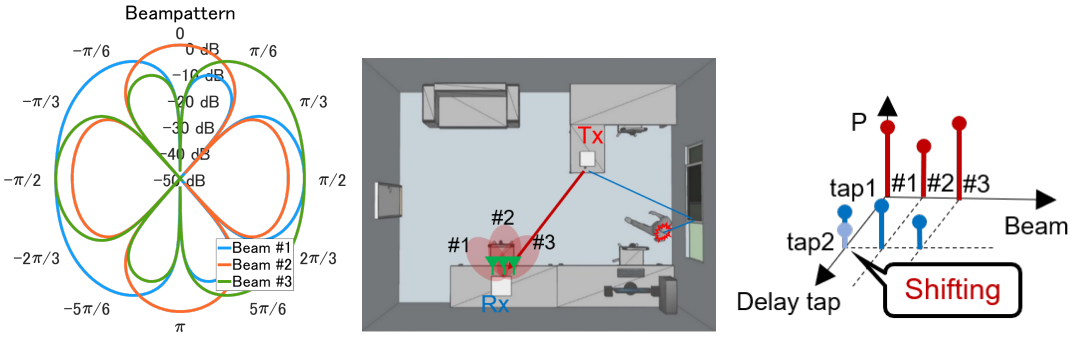
To cope with the problem that mentioned above, we proposed a method that treat the multiple-path components separately in both delay time domain and beam space. Here, beamforming with three orthogonal beams is used to separate the multi-paths in 3 beamspaces. The output signal obtained at multiple delay taps and multiple beamspaces is expressed as

$$y_{m,n}(t) = \boldsymbol{w}_m^H \boldsymbol{x}_n(t) = \gamma_{m,n}(t) e^{\mathrm{j}\theta_{m,n}}, \tag{3.22}$$

where γ and θ denote the amplitude and phase, respectively, and m and n denote the multi-beam index and delay tap index, respectively. Fig. 11(a) shows the beam pattern, the target space is divided into three parts to receive the multi-waves. Fig. 11(b) shows the condition when the single-bounce reflection wave is blocked by human body. Here, since the multiple-path components have been

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1

treated separately and received, the fluctuate of it can be detected when it's been blocked as can be seen in Fig. 11(c). Finally, the amplitude and phase of received signal obtained at multiple delay taps and multiple beamspaces are used as spatio-temporal feature values for machine learning.

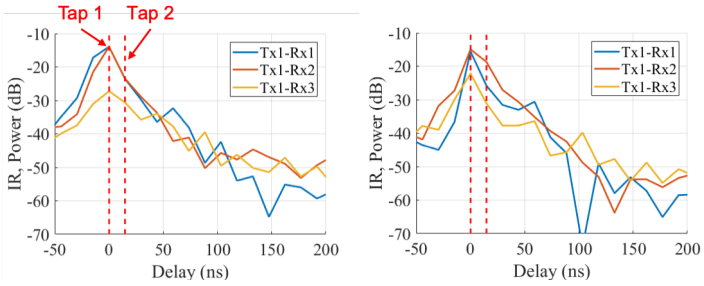


(a) Beam pattern (b) Event detection (c) Multi-path separation

Figure 11. Concept of the proposed method

4. Result and Discussion

The machine learning used to create the classification model is support vector machine (SVM) [26,27] and Random forest [28]. The implementation used one of the Python's [29] libraries scikit-learn [30], and the hyperparameters for the model are determined by a frame work call Optuna [31]. Fig. 12 shows the impulse response captured by Wi-Fi channel sounder at different condition (With and without a person). It is seen that the channel bonding extends the delay resolution to about 15 ns. Therefore the first delay tap was set to 0 ns and the second delay tap was set to about 15 ns considering the dimension of the room.



(a) Without human (b) Human exists

Figure 12. Fluctuation of impulse response

4.1. Small office

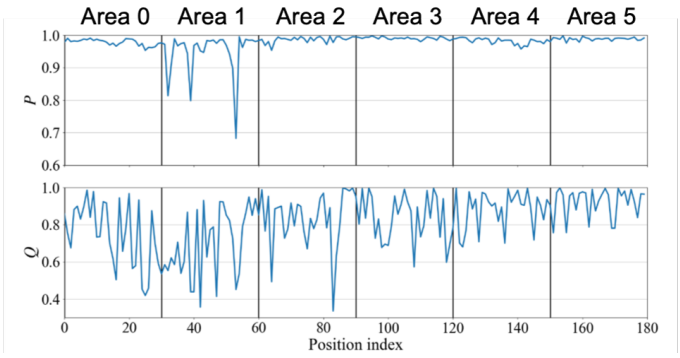
Fig. 13(a) shows the features obtained by the existing method in samll office. Here, the first eigenvector and the first eigenvalue obtained in the environment without person are used as reference data to calculate the features. One of the features, P, which is the correlation of eigenvoctor obtained both in environment with person exists and without person, fluctuates greatly in three locations in Area 1. This is considered due to the effect of shielding the direct wave. However, the Area 3 and Area 0 is relatively stationary compare to Area 1 even the direct wave also passed these areas. This may be due to the height of transmitting antenna, which is higher that the subject, thus the direct wave might not been shielded in Area 3, and not completely shielded depending on the orientation of the person's body in Area 0.

Fig. 13(b) shows the features obtained by the proposed method in small office. The solid line and dotted line are denoted as the delay tap 1 and delay tap 2, respectively. The different color blue,

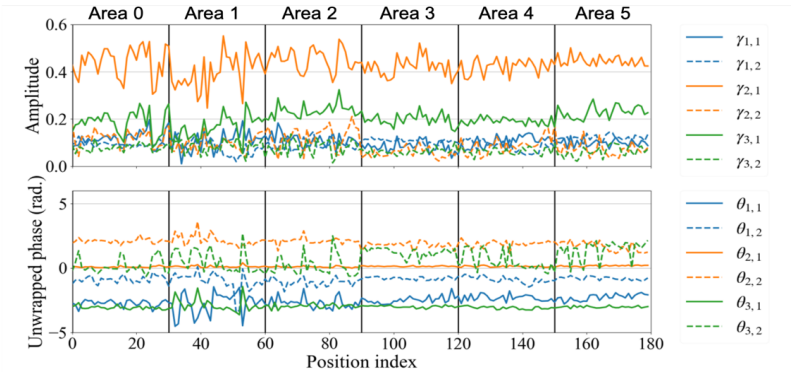
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1

orange, and green represent beam 1, beam 2, and beam 3, respectively. There are relatively large phase fluctuations can be seen on the first delay tap, beam 1, due to shielding the direct wave. In addition, there are many fluctuations also can be confirmed at second delay tap, this is considered as the result of shielding the single-bounce reflection wave. Which shows the events can be detected even when the multi-waves with a small power is blocked.

The classification result of the existing method and proposed method are shown in Fig. 14. Classification was performed by support vector machine. For the existing method, the success rate was 35.8 % and the average F-score was 0.31, which is not correctly classified. This may be due to not enough features and one of the features P is relatively stationary. For the proposed method, the success rate and average F-score are improved to 64.5 % and 0.63. However, the misclassification has occurred and considered to be due to the effect of similar propagation channel conditions. It is expected to be improved by defining the specific area of the room, like near the sofa, table, corridor, etc, instead of simply dividing the room equally.



(a) Existing method



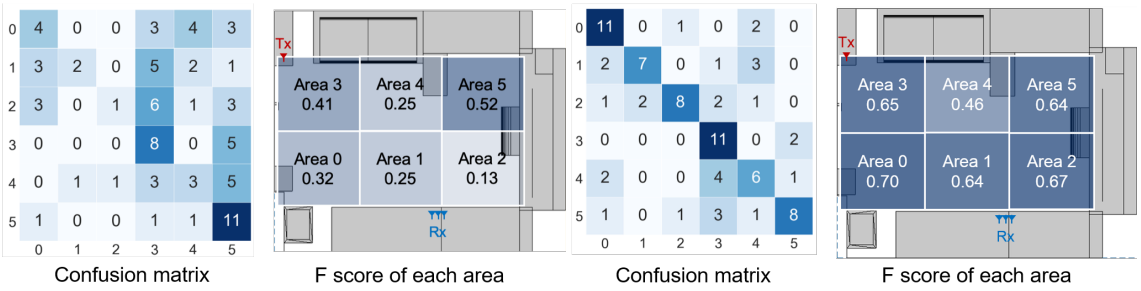
(b) Proposed method

Figure 13. Features of small office

4.2. Conference Room

Fig. 15 shows the features obtained by the proposed method in conference room. It is seen that the amplitude of the Tx2 fluctuates greatly in 15(b). This is considered due to the position of Tx2, which is placed outside the room and greatly affected by the noise. The classification results are shown in Fig. 16(a). The success rate was 88.2 and the average F value for each area was 0.84. Here, features of 76 positions are used to build the machine learning model, and rest are used as test data. Notice the positions of each area are randomly selected and the ratio of positions used for training to those used for prediction should be the same. From Fig. 16(b), it is seen that most of the misclassified positions are at nearby subarea boundary.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1



(a) Existing method (b) Proposed method

Figure 14. Classification result

5. Conclusions

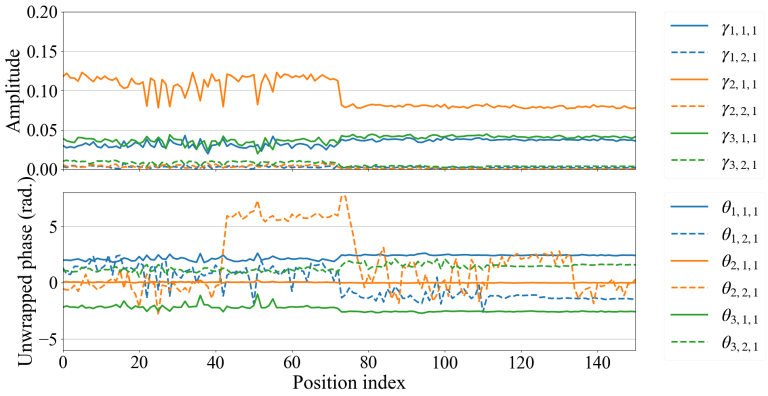
In this paper, a 3-by-3 MIMO channel sounder was built, and the CSIs taken at six consecutive WiFi channels was concatenated in order to extend the bandwidth and improve the delay time resolution. Based on this, a practical indoor location estimation system was proposed to cope with the problem that the features are significantly influenced by the strong line-of-sight path. Then we conducted the experiment in both small office and conference room, and the results of person position's estimation are presented by machine learning. The classification result was shown that rough classification of the subarea in a room where a target person is present is possible by using a commodity WiFi device.

Author Contributions: Formal analysis, Satoru Yasukawa; Funding acquisition, Minseok Kim; Methodology, Minseok Kim and Satoru Yasukawa; Project administration, Minseok Kim; Software, Hideaki Momose; Supervision, Minseok Kim; Validation, Yuan Zhou; Visualization, Hideaki Momose; Writing – original draft, Yuan Zhou.

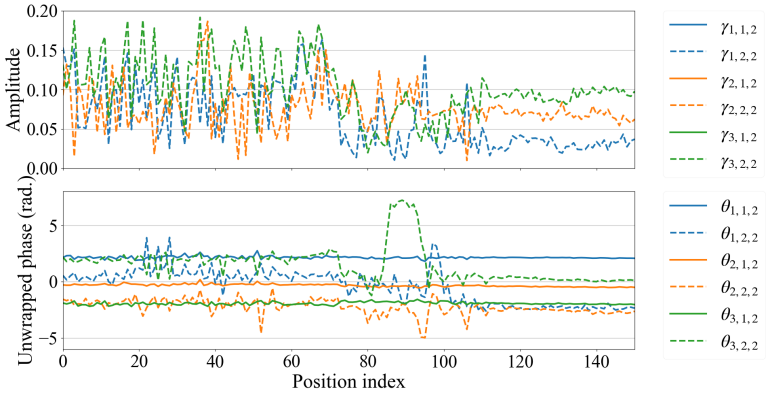
Acknowledgments: In this section you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

Conflicts of Interest: Declare conflicts of interest or state "The authors declare no conflict of interest." Authors must identify and declare any personal circumstances or interest that may be perceived as inappropriately influencing the representation or interpretation of reported research results. Any role of the funders in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript, or in the decision to publish the results must be declared in this section. If there is no role, please state "The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results".

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1



(a) Features of Tx1



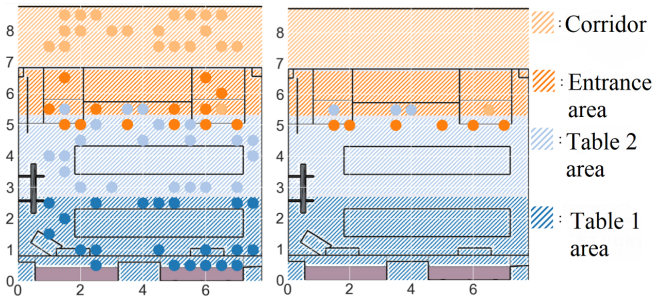
(b) Features of Tx2

Figure 15. Features of conference Room

References

1. M. Rainer, “Indoor positioning technologies” *ETH Zurich Research Collection*, pp.11–14, 2012.
2. J. Nash, “Global sales of video surveillance equipment projected to surpass 20 billion this year,” Biometric Update. Jan.10 2020.
3. E. Kaplan, C. Hegarty, “Understanding GPS Principles and Applications,” 2nd ed.; Artech House: Norwood, MA, USA, 2005.
4. S. Haykin, “Cognitive radio: brain-empowered wireless communications,” *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 20–220, 2005.
5. S. Woo, S. Jeong, E. Mok, L. Xia, C. Choi, M. Pyeon, “Application of WiFi-based indoor positioning system for labor tracking at construction sites: A case study in Guangzhou MTR” *Automation in Construction*, Vol. 20, Issue 1, pp. 3–13, Jan 2011.
6. S. Chan, G. Shon, “Indoor Localization Using Wi-Fi Based Fingerprinting and Trilateration Techiques for Lbs Applications,” *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* Volume XXXVIII-4/C26, pp. 1–5, 2012.
7. K. Kaemarungsi, P. Krishnamurthy, “Analysis of WLAN’s received signal strength indication for indoor location fingerprinting," *Pervasive and Mobile Computing*,Vol. 8, Issue 2, pp 292–316, 2012.
8. R. F. Brena, J. Pablo García-Vázquez, C. E. Galván-Tejada, D. Muñoz-Rodriguez, C. Vargas-Rosales, J. Fangmeyer, “Evolution of Indoor Positioning Technologies: A Survey," *Journal of Sensors*, vol. 2017, 2017.
9. S. Abdul Samadh, Q. Liu, X. Liu, N. Ghourchian and M. Allegue, "Indoor Localization Based on Channel State Information," *2019 IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNet)*, Orlando, FL, USA, pp. 1-4, 2019.
10. K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo and L. Ni, “CSI-Based Indoor Localization," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 7, pp. 1300–1309, July 2013.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 March 2021 doi:10.20944/preprints202103.0434.v1



(a) Classified position (b) Misclassified position

Figure 16. Classification result

11. X. Wang, L. Gao, S. Mao and S. Pandey, "CSI-Based Fingerprinting for Indoor Localization: A Deep Learning Approach," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 763–776, Jan. 2017.
12. B. Altintas, T. Serif, "Improving RSS-Based Indoor Positioning Algorithm via K-Means Clustering," *17th European Wireless 2011 - Sustainable Wireless Technologies*, Vienna, Austria, pp. 1–5, 2011.
13. Y. Zhang, Q. Cui, P. Zhang, X. Tao, "Channel State Information Feedback with Zero-Overhead in Closed-Loop MIMO System" *IEICE Transactions on Communications*, E93.Vol. B, No. 1, pp. 9–15, 2010.
14. T. Ohtsuki and J. Hong, "Activity Recognition Based on Array Sensor," Proc. APSIPA ASC 2011 Xi'an, Oct. 2011.
15. S. Yasukawa and M. Kim, "Intruder Detection Using Radio Wave Propagation Characteristics," Proc. ICCE-Asia 2018, Jun 2018.
16. N. Tadayon, M. Rahman, S. Han, S. Valaee, W. Yu, "Decimeter Ranging with Channel State Information," *IEEE Trans. Wireless Commun.*, Vol. 18, No. 7, pp. 3453–3468, July 2019.
17. H. Zhu, Y. Zhuo, Q. Liu, S. Chang, "π-Splicer: Perceiving Accurate CSI Phases with Commodity WiFi Devices," *IEEE Trans. Mobile Comput.*, Vol.17, No.9, pp. 2155–2165, Sept. 2018.
18. Y. Xie, "Atheros CSI Tool," https://wands.sg/research/wifi/AtherosCSI/
19. S. Sen, J. Lee, K. Kim, P. Congdon, "Avoiding Multipath to Revive Inbuilding WiFi Localization," Proc. *ACM Int. Conf. Mobile Syst., Appl., Services* (MobiSys' 13), pp. 249–262, Jun. 2013.
20. Lab of Innovation on Networking, "π-Splicer," http://lion.sjtu.edu.cn/project/projectDetail?id=5
21. Linux 802.11n CSI Tool, https://dhalperi.github.io/linux-80211n-csitool/
22. "IEEE standard for information technology -local and metropolitan area networks -specific requirements- part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications amendment 5: Enhancements for higher throughput," *IEEE Standard 802.11n*, pp. 1–565, Oct. 2009.
23. T. Yasushi, A. Yusuke, I. Yasuhiko, "MIMO Technologies in Wireless LAN Systems," *The journal of the Institute of Image Information and Television Engineers*, Vol.51, no.1, pp.35–40, 2016.
24. T junction, BL41-6203-00, Orient Microwave Corp, http://www.orient-microwave.com/index.html
25. Coaxial cable Enviroflex_316, Huber+Shuner AG Corp, http://hubersuhner.com/en/
26. V. N. Vapnik and A. Ya. Chervonenkis, "Pattern recognition using generalized portrait method," Automation and Remote Control, Vol.24, pp.774-780, 1963.
27. B. E. Boser et al., "A training algorithm for optimal margin classifiers," Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory, pp.144–152, 1992.
28. L. Breiman, "Random Forests," Machine Learning, Vol.45, pp.5–32, 2001.
29. Python Software Foundation, "Welcome to Python.org," https://www.python.org/, ref. Jan. 2020.
30. F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, Vol.12, pp. 2825-2830, 2011.
31. Preferred Networks, Inc., "Optuna – A hyper parameter optimization framework," https://optuna.org, ref. Jau. 2020.
32. T. Hastie, R. Tibshirani, J. Friedman, "The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition," Springer, 2009.