

## WILS: Wireless Indoor Localization System using Commercial WiFi Infrastructures with Decimeter Accuracy

Zhu-Chuang Shao, Dong Wang

School of Software Engineering, Shanghai Jiao Tong  
University, Shanghai, China  
E-mail: mrshao@sjtu.edu.cn, wangdong@cs.sjtu.edu.cn

Run Zhao

Computer Science Department, Shanghai Jiao Tong  
University, Shanghai, China  
E-mail: zhaorun@cs.sjtu.edu.cn

Ahsan Jamal Akbar

School of Software Engineering, Shanghai Jiao Tong University, Shanghai, China  
E-mail: ahsanakber512@sjtu.edu.cn

**Abstract**-we present Wireless Indoor Localization System (WILS) that enables commercial WiFi infrastructures to implement localization in the complex indoor environment. Receive Signal Strength Indication (RSSI) is utilized by some systems to get a high accuracy of localization, but they are easily be affected by the environment. Other use modified hardware or firmware, such as increasing the number of antennas to obtain enormous location precision. To solve this problem, WILS uses physical layer information, called Channel State Information (CSI). And it does not require modification in any hardware and firmware. WILS makes four technical contributions for achieving a high accuracy in localization. At first, it proposes a new algorithm, called Modified MUSIC (M- MUSIC), and it can estimate Angle of Arrival (AoA) and Time of Flight (ToF) at the same time from CSI information. Next, it introduces some methods to sanitize the phase, and improve the accuracy of measured phase. Third, it presents an idea of window slides to do CSI Smoothing for CSI measurements. Finally, it utilizes clustering algorithm to identify Line of Sight (LoS) path, and combines the AoA for LoS path and RSSI to locate the target. Our implementation on commodity WiFi cards demonstrates that WILS's accuracy is comparable to advanced localization systems, it can achieve a median accuracy of 60 cm.

**Keywords**-indoor localization; channel state information (CSI); angle of arrival (AoA); time of flight (ToF); modified MUSIC, line of sight(LoS); phase sanitization

### I. INTRODUCTION

Location-Based Service (LBS) plays a vital role in our daily life. LBS can be used to open some APPs, such as UBER, GOOGLE MAP and WECHAT. In an outdoor environment, there are systems such as GPS and BeiDou, which help us to make sure the position of the objects. However, such systems will not work in an indoor environment due to the blocking of the signals, and LoS will not exist. Because these systems rely on signal propagation and will be invalid in a complex and changeful indoor environment. It is a great challenge for indoor localization system.

Many existing systems make use of the RSSI information to implement localization. RSSI based systems [1], [2] are universal and simple but they are not accurate, and can only

achieve a median accuracy of 2-4m. In recent years, some systems utilize AoA to realize the localization, such as ArrayTrack [3] has higher accuracy, but it requires hardware modifications. CUPID [4] is a system proposed by HP, it utilizes physical layer information to get signal strength and AoA of the direct path. It does not need to modify hardware, but it cannot get the higher precision of AoA estimation and location accuracy. There are other methods, such as SAR (Synthetic Aperture Radar) in RFID-based Localization system [5], but these systems require massive calculation to make sure the position of a target, so it cannot satisfy the real-time localization.

Different from the approaches above, WILS proposes a new algorithm to estimate the AoA and ToF of LoS path and only uses existing commercial WiFi infrastructures. It does not require hardware modifications. Besides, WILS achieves decimeter level accuracy for indoor environment comparable to advanced localization systems.

The main augmentation of this paper is to design an indoor localization system, which achieves high accuracy and solves some related technical problems. WILS adopts four techniques to achieve our goal of accuracy and deployment simplicity:

**Modified MUSIC Algorithm:** WILS adopts a new algorithm to estimate AoA and ToF by using CSI information [6], [7], [13]. We know the fact that the number of antennas need be equal or greater than the number of multipath components according to the original MUSIC algorithm [8]. Some prior works solve this problem by increasing the number of antennas. However, WILS adopts the different approach; WILS utilizes the information of each subcarrier to simulate a virtual sensor array, thus overcoming this problem. Our insight is that we can use the CSI across antennas to estimate AoA and across subcarriers to estimate time of flight (ToF). Motivated by this fact, we modify the MUSIC algorithm only estimate AoA, and then generate the M-MUSIC algorithm that jointly estimates AoA and ToF. At last, WILS uses the triangulation approach to achieve accurate localization.

**Phase Sanitization:** Through some researches and experiments [9], [10], [14] found that the phase we obtain generally includes some noise. These noises will make the

performance of M-MUSIC decrease dramatically. This paper will provide some method to eliminate these noises.

**LoS Path Identification:** AoA of Non Line of Sight (NLoS) path is no effective for locating target. So we must acquire the AoA of LoS path. That is, how to identify LoS Path is a key point. WILS utilizes the AoA and ToF estimated by M-MUSIC and K-means algorithm to reach this goal. Meanwhile, WILS defines a likelihood function for each estimated path to be a LoS path.

**Localization:** Next, we will think about how to locate a target by using the result above. In brief, WILS conducts the localization after obtaining the LoS path. WILS mainly uses AoA of LoS path, RSSI and the probability of selected path is LoS path. We will present the details in section III.F.

The rest of this paper is organized as follows: we begin with a detailed description of some related works (section II), followed by details of our implementation, and contains M-MUSIC algorithm, Phase Sanitization, CSI Smoothing, LoS Identification and Localization (section III). Our experiment evaluates WILS's performance and compares with ArrayTrack and CUPID (section IV). WILSs accuracy is comparable to the most advanced localization systems; it can achieve a median accuracy of 60 cm. At last, we make a conclusion about this research in section V.

## II. RELATED WORKS

In this section, we will introduce some systems and related technologies. The existing system can be divided into the following categories.

**Propagation Model Based [1]:** RSSI is related to the distance between transmitter and receiver. We will estimate the distance if we can construct an accurate propagation model. Then we can reach the goal by using trilateration.

**Fingerprinting Based:** There are exist shadowing and multipath effect in an indoor environment, so it is difficult to construct an accurate propagation model and it will result in a major error. Some researchers proposed a method using fingerprinting. The rationality of this method lies in the distribution of wireless signal intensity in space is relatively stable. Compared with the method based on propagation model, this method can obtain higher precision. RADAR [2] system designed by Microsoft is this kind of system.

**SAR Based:** In recent years, some researchers in Germany referred to the radar technology and proposed a new method: Synthetic Aperture Radar (SAR) [11]. This kind of method has higher precision and has been implemented in RFID device and WIFI device, but it needs a lot of calculation.

**Angle of Arrival Based:** Triangulation is commonly used to locate the target. First, we need to measure the Angle of Arrival between referring points and target. ArrayTrack is this kind of system and uses antenna array to estimate AoA. ArrayTrack has higher precision, but it needs some APs with many antennas and is not practical.

**CSI Based [12]:** Different with RSSI, CSI is physical information and can describe the transmission characteristics of the wireless signal. CSI include the information of many subcarriers, it provides the possibility of estimating the AoA but using CSI information.

This paper is based on AoA and ToF of LoS path and CSI to design our system. WILS combines with the advanced technology and proposes some new ideas.

## III. SYSTEM DESIGN

To better understand our system, we introduce the conception of CSI and Multiple Signal Classification (MUSIC) [8]. In wireless communications, channel state information (CSI) refers to known channel properties of a communication link. This information describes how a signal propagates from the transmitter to the receiver and represents the combined effect of, for example, scattering, fading, and power decay with distance. CSI can estimate the channel in each subcarrier. We can obtain CSI with the help of Intel 5300 Wi-Fi NIC and Ubuntu Linux that install 802.11 CSI Tool [6], [7]. As we all know, RSSI only can reflect a each packet, but CSI is a fine-grained value from the PHY layer which describes the amplitude and phase on each subcarrier in the frequency domain [16]. MUSIC is an algorithm used for frequency estimation and emitter location. We will explain it in Section C. By combining the original MUSIC algorithm with CSI values, we proposed a new algorithm, called M-MUSIC, M-MUSIC not only can estimate AoA, but also can use for ToF estimation. The architecture of the WILS is shown in Figure 1.

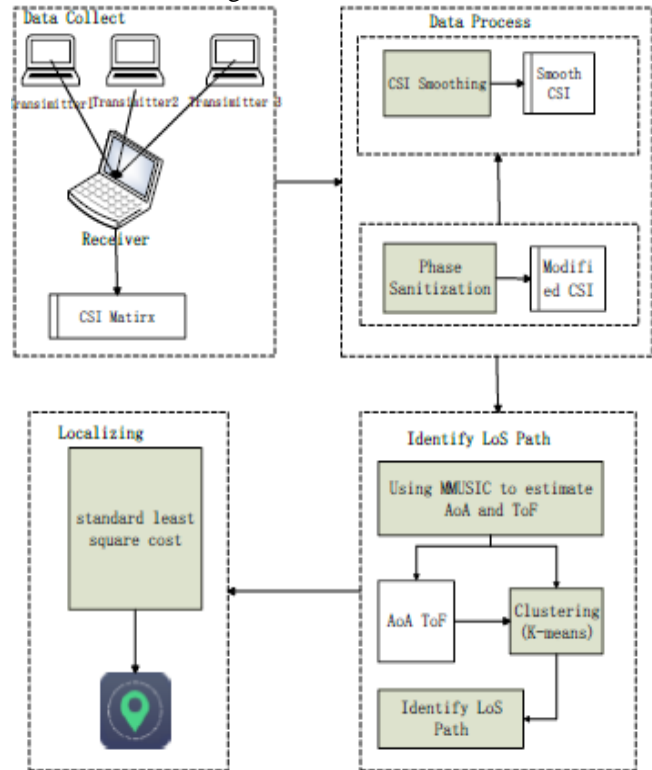


Figure 1. The architecture of the WILS

### A. Angle of Arrival Estimation with MUSIC

Next, we will introduce how to estimate AoA by using MUSIC. There is always serious reflection and multipath effect in the indoor environment. The target and AP will generate different propagation paths; these different

propagation paths have different AoAs. Besides, the AP usually equipped with three or more antennas and formed an antenna array, so the distance of signal propagation for the two nearby antennas is different and the difference will generate a phase shift across the antennas. The phase shift can be expressed by using the AoA and the distance between antennas.

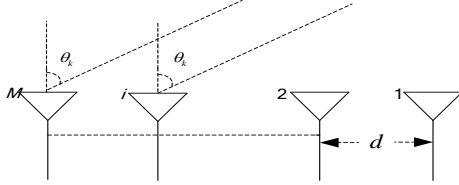


Figure 2. Uniform linear array consisting of M antennas

According to the Figure 2, we define the distance between antennas as  $d$  and AoA as  $\theta$ . The difference of signal propagation distance for the nearby antennas can be expressed as  $d \sin \theta$ . Supposed  $D$  signals arrive at an antenna from direction  $\theta_1, \dots, \theta_D$ . Then the signal arrive at the  $i$ -th antenna from the  $k$ -th signal is,

$$x_{i,k}(t) = a_k S_k(t) e^{-j\omega_0(i-1)d \sin \theta_k/c} \quad (1)$$

where  $S_k(t)$  is the signal of arrival at the first antenna from the  $k$ -th signal.  $a_k$  is the signal attenuation coefficient,  $\omega_0$  is the central frequency,  $c$  is the speed of light.

Allow for measurement noise and all incident signals, then the output signal of  $i$ -th antenna is,

$$x_i(t) = \sum_{k=1}^D a_k S_k(t) e^{-j\omega_0(i-1)d \sin \theta_k/c} + n_i(t) \quad (2)$$

where  $n_i(t)$  is the noise of the  $i$ -th antenna. We assumed the noise of antennas is Gaussian distribution and mean value is zero, the variance is  $\sigma^2$ , namely  $n_i(t) \sim N(0, \sigma^2)$ . Besides, we need make an assumption that the noise of is uncorrelated.

In order to analysis MUSIC well, we transform the expression to vector form, as follows,

$$\mathbf{X}(t) = \mathbf{A}\mathbf{S}(t) + \mathbf{N}(t) \quad (3)$$

where  $\mathbf{X}(t) = [x_1(t) \ x_2(t) \ \dots \ x_M(t)]^T$  is a  $M \times 1$  vector of receive data;  $\mathbf{S}(t) = [S_1(t) \ S_2(t) \ \dots \ S_D(t)]^T$  a  $D \times 1$  vector of signals;  $\mathbf{A} = [\mathbf{a}(\theta_1) \ \mathbf{a}(\theta_2) \ \dots \ \mathbf{a}(\theta_D)]$  is a  $M \times D$  vector;  $\mathbf{a}(\theta_k) = \left[ 1 \ e^{-\frac{j\omega_0 d \sin \theta_k}{c}} \ \dots \ e^{-\frac{j\omega_0 (M-1)d \sin \theta_k}{c}} \right]^T$  is a  $M \times 1$  vector of direction;  $\mathbf{N}(t) = [n_1(t) \ n_2(t) \ \dots \ n_M(t)]^T$  is a  $M \times 1$  vector of noise.

Due to the noise of different antennas and the different signals is uncorrelated, covariance matrix of  $\mathbf{X}(t)$  is,

$$\mathbf{R} = E\{\mathbf{X}(t)\mathbf{X}^H(t)\} \quad (4)$$

where  $\mathbf{X}^H(t)$  is conjugate transpose matrix of  $\mathbf{X}(t)$ ,

$$\mathbf{R} = \mathbf{A}\mathbf{P}\mathbf{A}^H + \sigma^2 \mathbf{I} \quad (5)$$

where  $\mathbf{P}$  is covariance matrix of signals,

$$\mathbf{P} = E\{\mathbf{S}(t)\mathbf{S}^H(t)\} \quad (6)$$

According to the previous assumption,  $\mathbf{A}$  can be expressed as follows,

$$\mathbf{A} = \begin{bmatrix} 1 & \dots & 1 \\ e^{-\frac{j\omega_0 d \sin \theta_1}{c}} & \dots & e^{-\frac{j\omega_0 d \sin \theta_D}{c}} \\ \vdots & \ddots & \vdots \\ e^{-\frac{j\omega_0 (M-1)d \sin \theta_1}{c}} & \dots & e^{-\frac{j\omega_0 (M-1)d \sin \theta_D}{c}} \end{bmatrix} \quad (7)$$

Easy to know that  $\mathbf{A}$  is Vandermonde matrix, the column of  $\mathbf{A}$  is independent of each other if  $\theta_i \neq \theta_j (i \neq j)$ . Then  $\text{rank}(\mathbf{A}\mathbf{P}\mathbf{A}^H) = D$  if  $\mathbf{P}$  is a nonsingular matrix,  $\text{rank}$  is the rank of a matrix. Because  $\mathbf{P}$  is a positive matrix,  $\mathbf{A}\mathbf{P}\mathbf{A}^H$  has  $D$  positive eigenvalue.

The covariance matrix  $\mathbf{R}$  is the full rank matrix,  $\mathbf{R}$  has  $M$  eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_M$  associated with  $M$  eigenvectors  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_M$ . These eigenvectors are orthogonal to each other. If we sort the eigenvectors according to the responding eigenvalues, then  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{M-D}$  is the noise subspace, called  $E_N$  and  $\mathbf{v}_{M-D+1}, \dots, \mathbf{v}_M$  is the signal subspace, called  $E_S$ .

To get AoA estimation, we define the following function:

$$Q(\theta) = \frac{1}{\mathbf{a}^H(\theta)\mathbf{E}_N\mathbf{E}_N^H\mathbf{a}(\theta)} \quad (8)$$

The  $\theta$  correspond to the maximum of a function is the estimation of AoA. As discussed above, the original MUSIC algorithm only can estimate AoA, this algorithm cannot make full use of CSI values. So we will propose the M-MUSIC in part D.

### B. Phase Sanitization

CSI reflect the amplitude and phase on each subcarrier. The phase information will be a key point of our system. We can express CSI as follows [10]:

$$\text{CSI}(f_k) = \|\text{CSI}(f_k)\| e^{j\angle \text{CSI}(f_k)} \quad (9)$$

where  $\text{CSI}(f_k)$  is the CSI value at the subcarrier with central frequency  $f_k$ .  $\|\text{CSI}(f_k)\|$  and  $\angle \text{CSI}(f_k)$  represent the amplitude and phase respectively. Actually, the phase contains some noise. The unprocessed phase  $\hat{\phi}_i$  for the  $i$ -th subcarrier of the  $j$ -th antenna can be expressed as,

$$\hat{\phi}_i^j = \phi_i^j - 2\pi \frac{k_i}{N} \delta + \beta_j + Z \quad (10)$$

where  $\phi_i^j$  is a real phase,  $\delta$  is the offset of time,  $\beta_j$  is related to the  $j$ -th antenna,  $Z$  is some measurement noise,  $k_i$  is the subcarrier index that can get from the standard of IEEE 802.11n.

$\beta_j$  is only related to the antenna, we can eliminate this error by using the CSI values of many subcarriers. We assume  $Z$  is Gaussian distribution and mean value is zero, the variance is  $\sigma^2$ , namely  $Z \sim N(0, \sigma^2)$ , so we can use the nature of Gaussian distribution to eliminate this noise.

Next, we only need to consider how to eliminate timing offset  $\delta$ . The timing offset  $\delta$  is same for every antenna, so we can use the linear fit to remove the effect of timing offset.  $\varphi_i(m, n)$  is the unprocessed phase for the  $i$ -th packet at  $n$ -th subcarrier of the  $m$ -th antenna,  $\delta_i$  is the timing offset for the  $i$ -th packet. The following pseudo code describes the linear fit algorithm.

1. Using the procession of linear fit for unprocessed CSI phase
$$\hat{\delta}_i = \underset{\rho}{\operatorname{argmin}} \sum_{m,n=1}^{M,N} (\varphi_i(m, n) + 2\pi f(n-1)\rho + \beta)^2$$
2. Using  $\hat{\delta}_i$  get from step 1 to modify the CSI phase
$$\varphi_i(m, n) = \varphi_i(m, n) + 2\pi f(n-1)\hat{\delta}_i$$

By using the above procession, we can get pure phase. It will build a solid foundation for M-MUSIC.

### C. CSI Smoothing

The analysis above is based on an assumption that the different signals are uncorrelated, but this assumption is generally invalid and the covariance matrix of a signal is probably singular [3], [9]. In this case, the performance of MUSIC and M-MUSIC will decrease dramatically. We need to find some method to solve the correlation of different signals.

To address this problem, our system introduces the idea of window slides. We called it CSI Smoothing. We can transform the CSI matrix to the following form, as shown in Equation (11). We will verify the performance of CSI Smoothing in Part D.

$$\text{CSI} = \begin{bmatrix} \text{csi}_{1,1} & \cdots & \text{csi}_{1,11} & \text{csi}_{2,1} & \cdots & \text{csi}_{2,11} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \text{csi}_{1,20} & \cdots & \text{csi}_{1,30} & \text{csi}_{2,20} & \cdots & \text{csi}_{2,30} \\ \text{csi}_{2,1} & \cdots & \text{csi}_{2,11} & \text{csi}_{3,1} & \cdots & \text{csi}_{3,11} \\ \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ \text{csi}_{2,20} & \cdots & \text{csi}_{2,30} & \text{csi}_{3,20} & \cdots & \text{csi}_{3,30} \end{bmatrix} \quad (11)$$

### D. Modified MUSIC Algorithm (M-MUSIC)

Due to the MUSIC algorithm has a constraint that the number of antennas is required to be larger than the number of signal paths. But this constraint cannot be satisfied in an indoor environment. Then we modify the MUSIC algorithm, and propose M-MUSIC for the joint estimation of AoA and ToF with only three antennas [9]. In our system, we can obtain the CSI of 30 subcarriers at three antennas. As shown in Equation (12), the CSI can be expressed as:

$$\text{CSI} = \begin{bmatrix} \text{csi}_{1,1} & \text{csi}_{1,2} & \cdots & \text{csi}_{1,30} \\ \text{csi}_{2,1} & \text{csi}_{2,2} & \cdots & \text{csi}_{2,30} \\ \text{csi}_{3,1} & \text{csi}_{3,2} & \cdots & \text{csi}_{3,30} \end{bmatrix} \quad (12)$$

Where  $\text{csi}_{m,n}$  is the value of the  $n$ -th subcarriers at the  $m$ -th antennas. Same as Equation (3), the signal arrive at antennas is,

$$X(t) = AS(t) + N(t) \quad (13)$$

$$X(t) =$$

$[x_{1,1}(t) \cdots x_{1,30}(t) \ x_{2,1}(t) \cdots x_{2,30}(t) \ x_{3,1}(t) \cdots x_{3,30}(t)]^T$ ,  $x_{m,n}(t)$  is the signal of the  $n$ -th subcarrier received at the  $m$ -th antennas. The definition of  $(t)$ ,  $n_{m,n}(t)$  is similar with  $X(t)$  and  $x_{m,n}(t)$ .  $S(t) = [S_1(t) \ S_2(t) \ \cdots \ S_D(t)]^T$  is a vector of signals.  $A = [a(\theta_1, \tau_1) \ a(\theta_2, \tau_2) \ \cdots \ a(\theta_D, \tau_D)]$  is the direction matrix.  $a(\theta_k, \tau_k)$  is a direction vector.

$$a(\theta_k, \tau_k) = [a_1(\theta_k, \tau_k) \ a_2(\theta_k, \tau_k) \ a_3(\theta_k, \tau_k)]^T \quad (14)$$

$$a_m(\theta_k, \tau_k) = [a_{m,1}(\theta_k, \tau_k) \ \cdots \ a_{m,30}(\theta_k, \tau_k)]^T \quad (15)$$

$$a_{m,i}(\theta_k, \tau_k) = e^{(-j2\pi[(i-1)\Delta f\tau_k + (m-1)f\frac{d \sin \theta_k}{c}]})} \quad (16)$$

Where  $\Delta f$  is the frequency difference.  $\tau_k$  is the ToF and  $\theta_k$  is the AoA of the  $k$ -th signal path. CSI can be expressed as,

$$\text{CSI} = A\beta + n \quad (17)$$

$$\text{CSI}(t) =$$

$[csi_{1,1}(t) \cdots csi_{1,30}(t) \ csi_{2,1}(t) \cdots csi_{2,30}(t) \ csi_{3,1}(t) \cdots csi_{3,30}(t)]^T$  is a vector of CSI collected from 30 subcarriers at 3 antennas.  $\beta = [\beta_1 \ \beta_2 \ \cdots \ \beta_D]^T$  is an amplitude attenuation vector. So the pseudo-spectrum can be calculated by

$$Q(\theta, \tau) = \frac{1}{a^H(\theta, \tau) E_N E_N^H a(\theta, \tau)} \quad (18)$$

By considering the AoA and ToF, we can get the above objective function. We will get the estimation of AoA and ToF by searching the peak of an objective function.

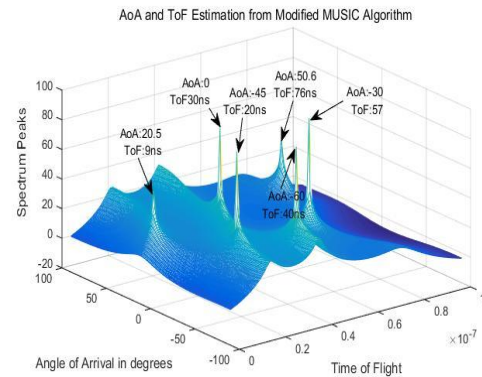


Figure 3. AoA and ToF estimation using M-MUSIC



Next, we conduct a simulation to evaluate the performance of M-MUSIC and CSI Smoothing. We use one AP with three antennas in the simulation. We simulate 30 signals that range from 5.18GHz to 5.21GHz (nearly 5.20GHz) and each signal differs by 1 MHz from the adjacent signal. Meanwhile, we also add noise to the signal. Figure 3 shows the result of the simulation with AoAs -60 degrees, -45 degrees, -30 degrees, 0 degrees, 20.5 degrees, 50.6 degree and ToFs 40 ns, 20 ns, 57 ns, 30 ns, 9 ns, 76ns. We can find the result of the simulation is very close to the real AoA and ToF.

AoA and ToF Estimation from Modified MUSIC Algorithm Without CSI Smoothing

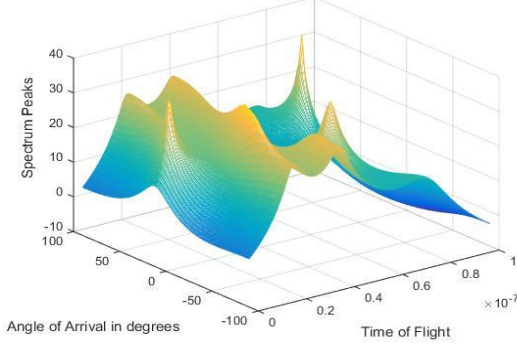


Figure 4. AoA and ToF estimation using M-MUSIC without CSI Smoothing

In order to evaluate the performance of CSI Smoothing, we conduct another simulation. We still use the AoA and ToF of the last simulation, but without CSI Smoothing. The result of the simulation is shown in Fig. 4. The performance of M-MUSIC will decrease dramatically if we don't use CSI Smoothing for CSI measurements.

#### E. LoS Path Identification

In this section, we will introduce how to identify the LoS path. Due to existing reflections that generate NLoS path and NLoS cannot be used in localization [15]. So it is a key point that how to identify the LoS path. By adopting M-MUSIC, we can estimate AoA and ToF for the paths between the APs and target. We can easily treat the path has the lowest ToF as the LoS path. But there is no time synchronization between the transmitter and the receiver, so this method is not reasonable.

We proposed a new idea combined with Machine Learning. After obtaining the AoA and ToF estimation, we treat AoA and ToF as X-value and Y-value in the two-dimensional coordinate system. Based on the assumption that AoA and ToF of the affinity propagation is similar. Then we apply the clustering algorithm to cluster the different packets. We adopt K-means to finish this function.

Next, we can define a likelihood function for each estimated path to be a LoS path. WILS is based on the fact that it is very close for AoA and ToF of the same path. The likelihood for the k-th path to be the LoS path is calculated as

$$\text{likelihood}_k = e^{w_c C_k - w_\theta \sigma_{\theta_k} - w_\tau \sigma_{\tau_k} - w_s \tau_k} \quad (19)$$

Where  $\text{likelihood}_k$  is the likelihood that the k-th path is the LoS path,  $C_k$  is the number of points in the cluster,  $\sigma_{\theta_k}$  and  $\sigma_{\tau_k}$  are the variances of the AoA and ToF respectively, and  $\tau_k$  is the mean of ToF in the cluster. The main idea of this likelihood is that the cluster with the most points and the least variance (tightest clustering) should be selected since the LoS path should have less variance in its measurement. The term with the mean ToF for a cluster is included since the signal with shorter ToF should be more likely to be the LoS path. The weighting factors  $w_c, w_\theta, w_\tau, w_s$  are weight coefficient of the corresponding variable. To compute the weights for each of the likelihood terms, I used the collected CSI again and then fed cluster information labeled with the correct cluster choice into a linear support vector machine. The SVM learned appropriate weights from the data.

WILS treats the path with the highest likelihood as the LoS path, and use the AoA and ToF of this path to locate a target. We will introduce the method to locate a target in next section.

#### F. Localization

Finally, WILS utilizes the AoA of LoS path and a likelihood value of the path to locating the target. Further, WILS also considers RSSI to obtain higher accuracy.

We define the following objective function and make sure the location of a target by minimizing the objective function.

$$\sum_{i=1}^N l_i [(\bar{p}_i - p_i)^2 + (\bar{\theta}_i - \theta_i)^2] \quad (20)$$

Where  $N$  is the number of AP,  $\theta_i$  is the AoA of LoS path from the i-th AP,  $p_i$  is the RSSI value from the i-th AP.  $\bar{\theta}_i$  and  $\bar{p}_i$  are unknown variables, and represent the AoA and RSSI,  $l_i$  is the highest likelihood value of paths from the i-th AP.

### IV. EXPERIMENT

In our experiments, we use Intel 5300 Wi-Fi NIC and Ubuntu Linux that install 802.11 CSI Tool [6], [7]. We use one computer in a monitor mode to collect CSI information. Transmitters and Receivers are work at 5.2GHz and the distance of two nearby antennas is set as half-wavelength. After obtaining the CSI information, we use MATLAB to running the algorithm to infer the target location. In order to simulate the environment well, we choose our lab as the testbed. Because the lab is a naturally dynamic environment and can test our system better.

#### A. AoA Measurement Estimation

In this section, we aim to demonstrate that the M-MUSIC is able to accurately estimate the AoA, and compare the performance of M-MUSIC with MUSIC in LoS and NLoS environment. The Cumulative Distribution Functions of errors in AoA estimation by the M-MUSIC and MUSIC are compared in Figure 5. We can find that by using the M-MUSIC in LoS environment, the mean of errors is about 3, which are much smaller than the ones by using MUSIC. In NLoS environment, M-MUSIC still shows up an improved performance.

## B. Location Estimation

We conduct our experiment by comparing state-of-the-art systems. WILS uses three APs with known locations and a computer act as a target. Three APs act as transmitter and target act as a receiver. The target can collect the CSI from different APs. First, WILS runs the M-MUSIC to obtain AoA and ToF. Second, WILS runs K-means to identify the LoS path. At last, we will locate the target. WILS compares with two existing indoor localization systems, the ArrayTrack [3] system and CUPID [4] system in LoS and NLoS environment, as shown in Fig. 6.

For LoS, We repeat our experiments several times to locate a target and shows that WILS have achieved a median accuracy of 0.6m as compared with ArrayTrack have achieved a median accuracy of 1.7 by increasing the number of antennas, and CUPID which uses propagation model based on CSI have achieved 3.2m median accuracy.

In order to evaluate our system, we further test our system under the NLoS environment. For NLoS, the performance of WILS is much better than the performance of ArrayTrack and CUPID, even better than ArrayTrack and CUPID in LoS environment. The Root Mean Square Error of WILS is about 1.5m in an NLoS environment.

In conclusion, whether in a LoS environment or NLoS environment, the performance of WILS is much better than ArrayTrack and CUPID.

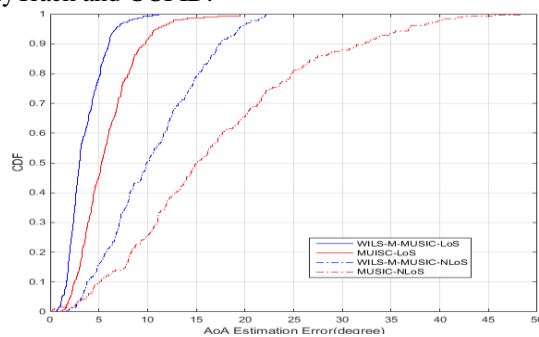


Figure 5. AoA estimation error

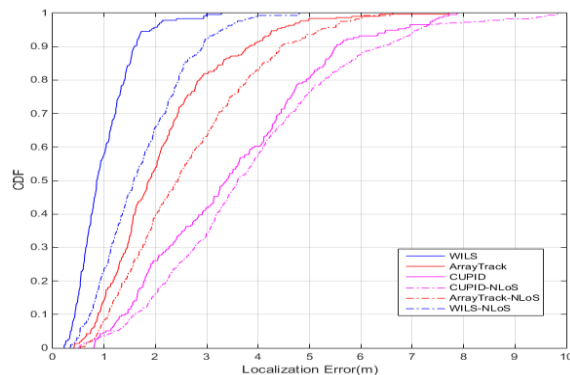


Figure 6. Localization error

## V. CONCLUSION

In this paper, we propose Wireless Indoor Localization System (WILS) that enables commercial WiFi

infrastructures to implement localization in the complex indoor environment. It proposes a new algorithm, called Modified MUSIC (M- MUSIC), that can estimate Angle of Arrival (AoA) and Time of Flight (ToF) at the same time from CSI information. Meanwhile, it introduces some methods of sanitizing the phase and CSI Smoothing for CSI measurements. Finally, it utilizes the clustering algorithm to identify LoS path. WILS's accuracy and AoA measurement estimation are comparable to state-of-the-art localization systems, and it can achieve a median accuracy of 60 cm.

## REFERENCES

- [1] Seidel S Y, Rappaport T S. 914 MHz path loss prediction models for indoor wireless communications in multifloored buildings[J]. IEEE Transactions on Antennas & Propagation, 1992, 40(2):207-217.
- [2] Bahl P, Padmanabhan V N. RADAR: An In-Building RF-based User Location and Tracking System[J]. Institute of Electrical & Electronics Engineers Inc, 2000, 2:775--784.
- [3] Xiong J, Jamieson K. ArrayTrack: a fine-grained indoor location system[C]// Usenix Conference on Networked Systems Design and Implementation. 2013:71-84.
- [4] Sen S, Lee J, Kim K H, et al. Avoiding multipath to revive inbuilding WiFi localization[C]// Proceeding of the, International Conference on Mobile Systems, Applications, and Services. ACM, 2013:249-262.
- [5] Parr A, Miesen R, Vossiek M. Inverse SAR approach for localization of moving RFID tags[C]// IEEE International Conference on Rfid. 2013:104-109.
- [6] Halperin D, Hu W, Sheth A, et al. Predictable 802.11 packet delivery from wireless channel measurements [J]. ACM SIGCOMM Computer Communication Review, 2011, 41(4): 159-170.
- [7] Halperin D, Hu W, Sheth A, et al. Tool release: Gathering 802.11 n traces with channel state information [J]. ACM SIGCOMM Computer Communication Review, 2011, 41(1): 53-53.
- [8] Schmidt R. Multiple emitter location and signal parameter estimation[J]. IEEE Transactions on Antennas & Propagation, 1986, 34(3):276-280.
- [9] Kotaru M, Joshi K, Bharadia D, et al. Spotfi: Decimeter level localization using wifi[C]//Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication. ACM, 2015: 269-282.
- [10] Qian K, Wu C, Yang Z, et al. Tuning by turning: Enabling phased array signal processing for WiFi with inertial sensors[C]. International Conference on Computer Communications, 2016.
- [11] Kumar S, Gil S, Katabi D, et al. Accurate indoor localization with zero start-up cost[C]//Proceedings of the 20th annual international conference on Mobile computing and networking. ACM, 2014: 483-494.
- [12] Wu K, Xiao J, Yi Y, et al. CSI-based indoor localization [J]. Parallel and Distributed Systems, IEEE Transactions on, 2013, 24(7): 1300-1309.
- [13] Yang Z, Zhou Z, Liu Y. From RSSI to CSI: Indoor localization via channel response [J]. ACM Computing Surveys (CSUR), 2013, 46(2)
- [14] Gjengset J, Xiong J, Mcphillips G, et al. Phaser: enabling phased array signal processing on commodity WiFi access points[C]// International Conference on Mobile Computing and NETWORKING. ACM, 2014:153-164.
- [15] Wu C, Yang Z, Zhou Z, et al. PhaseU: Real-time LOS identification with WiFi[C]// IEEE International Conference on Computer Communications. 2015.
- [16] Wu K, Xiao J, Yi Y, et al. FILA: Fine-grained indoor localization[J]. Proceedings - IEEE INFOCOM, 2012, 131(5):2210-2218.