A Privacy-Preserving Fuzzy Localization Scheme with CSI Fingerprint

Xiaoshan Wang*†, Yao Liu‡, Zhiqiang Shi*, Xiang Lu* and Limin Sun*
*Beijing Key Laboratory of IOT Information Security, Institute of Information Engineering, CAS, China
†University of Chinese Academy of Sciences, China
Email: (wangxiaoshan, shizhiqiang, luxiang, sunlimin)@iie.ac.cn
‡Department of Computer Science and Engineering, University of South Florida, USA
Email: yliu@cse.usf.edu

***Abstract*—CSI fingerprint localization is an advanced and promising technique for indoor localization, which identifies the user's location by mapping his measured CSI against the server's CSI fingerprint database. This approach is highlighted due to its high granularity for location distinction and strong robustness to noise disturbances, but it also causes potential privacy leakage for the three participants in localization process: the user, the server, and the AP. Currently, there has been little research done on this issue, and the existing work often ignores the privacy concern on the AP. To fill the gap, this paper develops a privacy-preserving fuzzy localization scheme with CSI fingerprint. On one hand, it leverages the property of CSI training to guarantee the randomness and independence of the user's measurement in each time of localization, and uses homomorphic encryption to achieve the data transmission and measurement comparison in cipher. These operations enable our scheme to preserve the location privacy of the user and APs as well as the data privacy of the server. On the other hand, the adoption of CSI fingerprint and fuzzy logic enhances the localization accuracy greatly. Through simulation experiments performed on CRAWDAD database, the efficiency of our proposed scheme is validated.**

I. Introduction

Indoor localization is an important and practical service for commercial, public safety, and military applications [1]. Due to the lack of GPS signals indoor and the increasing demand for the high precision, low cost and convenient usage, a vast range of approaches have been proposed based on various signals. Among them, the CSI (Channel State Information) fingerprint localization with access points (APs) is one of the most advanced and promising technologies, which involves three participants in the process of localization: the user, the server, and the AP. It determines the user's location by mapping the CSI measurements from APs against the CSI fingerprints pre-stored on the server.

As its name implies, CSI denotes the characteristic of the transmission links for wireless signals. It includes the signal power (SP), the time-of-arrival (ToA), and the angle-of-arrival (AoA) [2]. Based on a training signal X sent by the transmitter and a received signal Y, the CSI H between the transmitter and receiver can be estimated [3]. In contrast to the traditional fingerprint RSS (Received Signal Strength) which is vulnerable to channel effects (e.g., multipath effect and shadow fading) in the complex indoor environment, CSI takes advantage of the channel effects as valuable signatures to label different positions. By virtue of the abundant information contained in CSI, the indoor locations can be distinguished with more granularity, and less infrastructures (APs) are required to gain the equivalent efficiency compared with RSS. Moreover, the CSI fingerprint localization could be more robust because occasional disturbances would only contaminate the values at one or a small fraction of dimensions in the whole CSI vectors of high dimension, while leaving values at other dimensions unchanged. With the rapid development of commercial off-the-shelf devices and novel techniques for CSI measuring [4], it is more and more convenient to extract the CSI features in practice. Solid evaluations and comparisons have proved that the CSI based method outperforms the RSS based greatly in localization [5]. Therefore, much work has been done in recent years to explore the CSI fingerprint localization methods [6], [7], [8], [9], or incorporate CSI into the existing localization framework [10].

However, the working principles of the CSI fingerprint localization also lead to potential threats on the privacy of localization participants. For the user, his location retrieval request conveying his CSI measurement to the server would leak his location privacy, which indicates his behavior habits, interests, and social relationships [11], [12]. For the server, the violence of its data privacy would cause financial loss to the service provider who has paid much money and time to build the CSI fingerprint database. In particular for APs, the leakage of their CSI data would expose their positions to malicious attackers, thereby incurring physical access and damage. Researchers have discovered that the ratio of signal powers between the first peak and second peak in CSI can be used as a physical layer metric to gauge the distance from the AP [13]. The scheme in [14] depends on the ToA and AoA to achieve localization based on only one AP. Conversely, the similar technique could also be applied to locate the AP. So the AP's privacy is a problem that should not be ignored, especially in the security protection of military applications.

Although many works have paid attention to the issue of privacy-preserving in localization now, they might either exclude the security consideration of the AP, or be designed not for the CSI fingerprint based scenario. For example in [15], Hong Li et al. develop a privacy-preserving localization scheme with RSS fingerprint of WiFi signals to protect the user and the server. But it provides no consideration on the protection of APs. In [16], Tao Shu et al. design three protocols for different privacy-preserving levels, among which the third one is of the highest level of security. This scheme could preserve the privacy of APs to a certain degree, but it is not specifically designed for CSI fingerprint localization.

In his paper, we develop a privacy-preserving fuzzy localization scheme with CSI fingerprint to provide protections to the user, server and AP together. Under this scheme, the request signal from the user would be obfuscated, and the location retrieval would be carried out in cipher, such that the user's privacy can be protected. More importantly, based on the characteristic that the CSI is calculated through a training signal, this scheme has APs to use the training signals that are random and secret to the user, such that the user can neither know his CSIs between APs and himself exactly, nor learn the true differences between his own measurement and the measurements derived from the CSI fingerprint database. In this way, the privacy of the server and APs can be protected. Additionally, the employment of the fuzzy algorithm in location retrieval can greatly enhance the localization accuracy while preserving privacy.

The contributions of this paper are three-fold: Firstly, to the best of our knowledge, we propose the first privacy-preserving mechanism for CSI fingerprint localization to protect the privacy of the user, AP and server at meantime. Secondly, we establish a fuzzy localization framework to implement the proposed privacy-preserving mechanism. Thirdly, We carry out simulation experiments on the publicly accessible data set CRAWDAD to validate our scheme.

II. Background

In this section, we will present some background knowledge for the development of our privacy-preserving fuzzy localization scheme with CSI fingerprint. The first one is the basic fingerprint localization method regardless of the fingerprint type. The second one is the fuzzy algorithm used for location retrieval to improve localization accuracy. The last one is the Paillier cryptosystem used to encrypt the interactive data homomorphically in localization process.

A. Basic CSI Fingerprint Localization

A basic fingerprint localization scheme is composed of offline training stage and online service stage. It works mainly as follows.

1) Offline Training Stage: In offline training stage, the service provider would build up a fingerprint database by sampling certain measurements at each reference location, and store the database on the server. No matter whether the fingerprint is RSS or CSI, a single fingerprint is a n-dimensional vector as

$$H_k = [h_{1k}, h_{2k}, \cdots, h_{nk}]^T \in C^n \tag{1}$$

It corresponds to the k-th reference location loc_k, which is expressed in terms of 2-dimensional or 3-dimensional coordinates, $k = 1, 2, \cdots, M$. M is the volume of the database. The fingerprint database can be formulated as

$$\Psi = [H_1, H_2, \cdots, H_k, \cdots, H_M] \in C^{n \times M} \tag{2}$$

2) Online Service Stage: In online service stage, a user to-be-localized samples his own fingerprint as

$$H_0 = [h_{10}, h_{20}, \cdots, h_{n0}]^T \in C^n \tag{3}$$

and transmits this n-dimensional vector to the server. In order to retrieve his location, H_0 would be mapped against all H_ks in the fingerprint database Ψ with some algorithms, such as KNN (K Nearest Neighbor) algorithm [15], [17]. By this algorithm, the server would identify the user's location as the centroid of the K reference locations, which correspond to the K nearest fingerprints to the user's measurement.

B. Fuzzy Logic

Created by Zadeh, the fuzzy logic is a useful tool to deal with the uncertainty and vagueness in computation [18]. As shown in Figure (1), a typical Fuzzy Logic System (FLS) consists of four components: the fuzzy rule, the fuzzifier, the fuzzy inference engine, and the defuzzifier. Take the fingerprint localization for example, the roles that the four components play in a FLS can be summarized as follows.

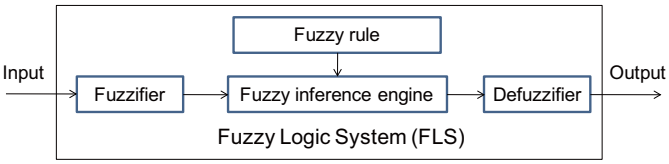

Fig. 1. Diagram of the typical Fuzzy Logic System

1) Fuzzy Rule: The fuzzy rule is the basis of reasoning in a FLS. Under the background of fingerprint localization, the k-th $(k = 1, 2, \cdots, M)$ fingerprint can be regarded as the k-th fuzzy rule in the form of

$$\text{IF } H_0 \text{ is } H_k, \text{ THEN } loc_0 \text{ is } loc_k. \tag{4}$$

The user's measurement H_0 and location loc_0 are the variable vectors of fuzzy input and fuzzy output, while the reference fingerprint H_k and reference location loc_k are the k-th center vectors of fuzzy input and fuzzy output corresponding to the k-th fuzzy rule.

2) Fuzzifier: Before an input vector is fed into a FLS, it needs to be fuzzified firstly by the fuzzifier as a preparation. In the scenario of fingerprint localization discussed here, a probability μ_k will be generated for the input to describe how nearly it is close to the k-th input center vector, i.e., how likely it "belongs to" the k-th fuzzy rule. This probability is called membership, and the function to generate membership is called membership function.

3) Fuzzy Inference Engine and Defuzzifier: The fuzzy inference engine is the core of a FLS. It is responsible to execute fuzzy reasoning according to the fuzzy rules. Based on the memberships of inputs, the inference engine derives the memberships of the output, which describe how nearly the output is close to each output center vector. Since the output of the inference engine is still fuzzy, it is necessary to use the defuzzifier to generate the definite output of the FLS. In our discussion on fingerprint localization, the operations of fuzzy inference and defuzzification can be performed together as

$$loc_0 = \frac{\sum_{k=1}^{M} \mu_k \cdot loc_k}{\sum_{k=1}^{M} \mu_k} \tag{5}$$

C. Paillier Cryptosystem

The construction of our privacy-preserving localization scheme relies on the Paillier cryptosystem. It is an asymmetric encryption method invented by Pascal Paillier based on decisional composite residuosity problem [19]. In this work, we design Paillier encryption scheme as follows.

- Key generation: Choose two large prime numbers p, q, and compute $N = pq$ and $\phi(N) = (p-1)(q-1)$. The public key is N, and the private key is $\langle N, \phi(N)\rangle$.
- Encryption: For a plaintext $m \in \mathbb{Z}_N$, choose a random number $r \in \mathbb{Z}_N^*$, and the ciphertext is given by

$$[\![m]\!] = (1+N)^m \cdot r^N \quad mod\ N^2 \tag{6}$$

- Decryption: Given a ciphertext $c \in \mathbb{Z}_{N^2}$, the plaintext is obtained by

$$m = \frac{[c^{\phi(N)}\ mod\ N^2] - 1}{N} \cdot \phi(N)^{-1} \quad mod\ N \tag{7}$$

The Paillier encryption scheme has the following additively homomorphic property. For m_1, m_2, m, $c \in \mathbb{Z}_N$, there are

$$[\![m_1 + m_2 \quad mod\ N]\!] = [\![m_1]\!] \cdot [\![m_2]\!] \quad mod\ N^2 \tag{8}$$

$$[\![c \cdot m \quad mod\ N]\!] = [\![m]\!]^c \quad mod\ N^2 \tag{9}$$

III. Fuzzy CSI Fingerprint Localization and Its Privacy Threats

In this section, we propose a fuzzy localization framework with CSI fingerprint that is not equipped with any privacy-preserving mechanism. It identifies the user's location with fuzzy logic based on a simple idea that the reference location with less error in fingerprint mapping should contribute more to the final localization result. Under this framework, we discuss the privacy leakage threats on the user, server and AP respectively.

A. Fuzzy CSI Fingerprint Localization

As a fingerprint localization method, the structure of our fuzzy framework with CSI fingerprint is shown in Figure (2). In the training stage, the service provider needs to determine a fuzzy parameter for the future service, besides establishing the CSI fingerprint database. In the service stage, the fuzzy memberships with which the user's location "belongs to" different reference locations would be calculated by the server, and the final estimation of user's location is the sum of all reference locations weighted by the memberships.

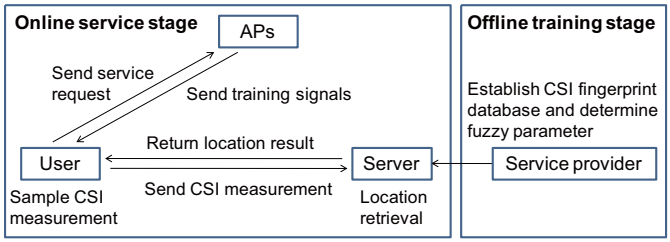

Fig. 2. Framework of fuzzy CSI fingerprint localization

1) Offline Training Stage: Assume there are B APs deployed and M reference locations. In offline training stage, the service provider samples CSI data for m_k times at the k-th reference location loc_k to get the measurement $L_k(j) = [L_{1k}(j)^T, L_{2k}(j)^T, \cdots, L_{Bk}(j)^T]^T$, $j = 1, 2, \cdots, m_k$, $k = 1, 2, \cdots, M$. $L_{bk}(j)$ in $L_k(j)$ is the n_b-dimensional CSI vector contributed by the b-th AP, $b = 1, 2, \cdots, B$. The total dimension of $L_k(j)$ is $n = \sum_{b=1}^{B} n_b$. Let $H_{bk} = \frac{1}{m_k}\sum_{j=1}^{m_k} L_{bk}(j)$. The service provider regards the following vector as the CSI fingerprint for loc_k.

$$H_k = [H_{1k}^T, H_{2k}^T, \cdots, H_{bk}^T, \cdots, H_{Bk}^T]^T \in C^n \tag{10}$$

For the convenience of writing, it can be reformulated as (1). The whole CSI fingerprint database is

$$\Psi = [H_1, H_2, \cdots, H_k, \cdots, H_M] \in C^{n \times M} \tag{11}$$

For the convenience of writing, it can be reformulated as (2).

In our fuzzy localization scheme, we adopt Guassian fuzzy membership function as

$$\mu_k = exp(-\frac{D_k}{\sigma^2}) \tag{12}$$

where $D_k = \|H_0 - H_k\|_2^2$, H_0 is the user's CSI measurement, and σ^2 is the fuzzy parameter. According to fuzzy theory, σ^2 is related to the "width" of Guassian membership function. It can be selected by the service provider based on practical experience or expert knowledge, such as

$$\sigma^2 = \frac{1}{M}\sum_{k=1}^{M}(\frac{1}{m_k}\sum_{j=1}^{m_k}\|L_k(j) - H_k\|_2^2) \tag{13}$$

The service provider stores both Ψ and σ^2 on the server.

2) Online Service Stage: In online service stage, a user can be aware of his location by interacting with the server and APs. The server would figure out the user's location with a fuzzy method. The detailed process is as follows.

Firstly, the user broadcasts service request to APs. For $b = 1, 2, \cdots, B$, the b-th AP replies training signal X_b that is known to the user beforehand. Correspondingly, the user receives signal Y_{b0}. Based on X_b and Y_{b0}, the user can solve the CSI vector H_{b0} as demonstrated in [3]. To avoid the interference among different training signals, APs send signals in the order how H_{bk}s are arranged in H_k ($k = 1, 2, \cdots, M$), and in each time slot there is only one AP sending. All H_{b0}s lined up together compose the entire CSI measurement of the user as

$$H_0 = [H_{10}^T, H_{20}^T, \cdots, H_{b0}^T, \cdots, H_{B0}^T]^T \in C^n \tag{14}$$

For the convenience of writing, it can be reformulated as (3). Then the user sends H_0 to the server for fuzzy location retrieval.

After receiving user's message, the server computes the localization result according to (12) and (5), and returns this result to the user finally.

B. Privacy Threats in Fuzzy CSI Fingerprint Localization

The use of CSI fingerprint and fuzzy algorithm can enhance the localization accuracy, but does nothing more to protect the privacy compared with other fingerprint based schemes. In this work, we assume that all the participants involved in

localization (the user, the server, and the AP) are all honest but curious. This means that they comply with the specified localization process and do not collaborate with attackers. Next, we will summarize the privacy threats on the user, server and AP respectively.

1) For the user: In localization, the threats on the user's privacy are three-fold. Firstly, the user may expose his CSI to attackers when he requests localization service, leading his location to be inferred unwantedly. Secondly, the user's location can be learned by the server according to his CSI measurement reported. Thirdly, an attacker can capture the CSI measurement sent by the user, and use it to retrieve the user's location on the server.

2) For the server: In localization, an attacker can fabricate a large number of artificial CSI measurements in a fuzzing like way, or sample real CSI measurements at a large number of spots. Then he uses these measurements to ask the server for localization service. By recording all pairs of the CSI and its corresponding location, he can establish his own fingerprint database which is similar to the original one on the server.

3) For the AP: When APs send training signals to the user in service stage, they might leak out their CSI data, which can be used to infer their locations with the technologies mentioned in Introduction. It might lead to the threats of access and damage physically.

IV. Privacy-Preserving in Fuzzy CSI Fingerprint Localization

As discussed before, the privacy threats in CSI fingerprint localization are mainly attributed to the direct service request from the user, and the fingerprint comparison in plaintext. So in this section, we will present a privacy-preserving fuzzy localization scheme with CSI fingerprint to deal with the two problems above. Moreover, we will analyze its protective effects on the user, server and AP respectively.

A. Privacy-Preserving Localization Scheme

Based on the framework of fuzzy CSI fingerprint localization, our privacy-preserving scheme is additionally equipped with a set of protective mechanism to address the privacy threats in localization. Firstly, an obfuscating operation would be performed by the user when he requests training signals from APs. Secondly, a homomorphic encryption method would be used for the data transmission and measurement comparison in cipher. Lastly and the most importantly, random secret training signals would be leveraged by APs to break the correlation of the user's measurement with the CSIs of APs and the CSI fingerprints on the server. The work flow of our privacy-preserving fuzzy scheme is shown in Figure (3). According to the different parties and their operations in localization, the detailed localization process can be divided into the following five steps.

Step 1: Preparation

In offline training stage, the service provider establishes the CSI fingerprint database Ψ as (2) and determines the fuzzy parameter σ^2 as (13). Then he stores Ψ on the server, but releases σ^2 and each reference location loc_k $(k = 1, 2, \cdots, M)$ to the public. It is worth noting that the fuzzy parameter σ^2 here should be different from that in the original fuzzy scheme without privacy-preserving. The new selection strategy will be given in Step 3.

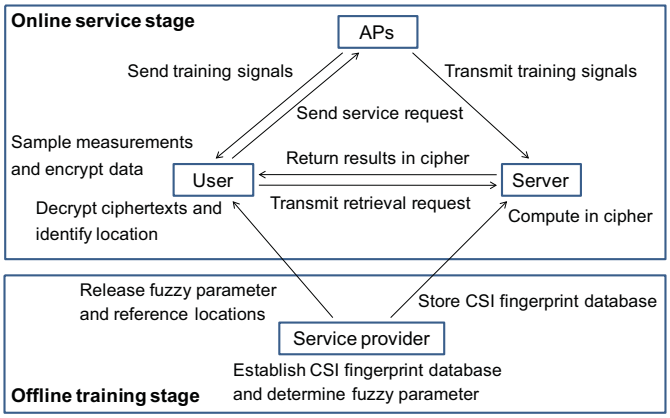

Fig. 3. Privacy-preserving fuzzy localization scheme with CSI fingerprint

Step 2: Service request

When requesting localization service from APs, the user obfuscates his request signal with some methods. As discussed in [3], if the CSI fingerprint is defined in time domain, a delay-and-sum mechanism will be required, and if the CSI fingerprint is defined in frequency domain, a scaling modulation will be implemented for each subcarrier instead.

Step 3: CSI measuring

As a reply to the user's request, the b-th $(b = 1, 2, \cdots, B)$ AP sends him a random training signal X_b, which is secret to the user and independent on any CSI data. At the same time, the value of X_b is transmitted to the server.

On the server side, supposing that X_b is sent over an indoor wireless channel with CSI H_{bk}, the server can obtain the imagined output Y_{bk} for $k = 1, 2, \cdots, M$. The server arranges all Y_{bk}s in the same order as H_{bk}s in Ψ, thus yielding a new measurement database as

$$\Psi^{'} = [Y_1, Y_2, \cdots, Y_k, \cdots, Y_M] \in C^{n \times M} \tag{15}$$

where $Y_k = [Y_{1k}^T, Y_{2k}^T, \cdots, Y_{bk}^T, \cdots, Y_{Bk}^T]^T$ is the k-th reference measurement. Like Ψ before, $\Psi^{'}$ can be reformulated for the convenience of writing as

$$\Psi^{'} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1k} & \cdots y_{1M} \\ y_{21} & y_{22} & \cdots & y_{2k} & \cdots y_{2M} \\ \vdots & \vdots & & \vdots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nk} & \cdots y_{nM} \end{bmatrix} \in C^{n \times M} \tag{16}$$

To adapt to the new measurement database above, the selection of fuzzy parameter σ^2 in Step 1 should be modified as

$$\sigma^2 = \frac{1}{M} \sum_{k=1}^{M} \left(\frac{1}{m_k} \sum_{j=1}^{m_k} \sum_{b=1}^{B} \| Y_{bk}^L(j) - Y_{bk} \|_2^2 \right) \tag{17}$$

where $Y_{bk}^L(j)$ is the imagined output supposing X_b is sent over an indoor wireless channel with CSI $L_{bk}(j)$.

On the user side, he cannot solve his CSI vector due to the unknown X_b for $b = 1, 2, \cdots, B$. Instead, he puts each received signal Y_{b0} together in a line as

$$Y_0 = [Y_{10}^T, Y_{20}^T, \cdots, Y_{b0}^T, \cdots, Y_{B0}^T]^T \in C^m \tag{18}$$

Like H_0 before, the new measurement Y_0 can be reformulated for the convenience of writing as

$$Y_0 = [y_{10}, y_{20}, \ldots, y_{n0}]^T \in C^n \tag{19}$$

Note that

$$\begin{aligned} D_k^{'} &= \|Y_0 - Y_k\|_2^2 = \sum_{i=1}^{n}(y_{i0} - y_{ik})^2 \\ &= \underbrace{\sum_{i=1}^{n}(y_{i0})^2}_{G_1} + \underbrace{\sum_{i=1}^{n}(-2y_{i0} \cdot y_{ik})}_{G_2} + \underbrace{\sum_{i=1}^{n}(y_{ik})^2}_{G_3} \end{aligned} \tag{20}$$

The user generates a public key K_p and a private key K_s using Paillier cryptosystem. With K_p, he computes the following ciphertexts

$$S_1 = [\![G_1]\!] = [\![\sum_{i=1}^{n}(y_{i0})^2]\!] \tag{21}$$

$$S_2 = \{[\![-2y_{10}]\!], [\![-2y_{20}]\!], \cdots, [\![-2y_{n0}]\!]\} \tag{22}$$

Then he transmits $\{S_1, S_2, K_p\}$ as the location retrieval request to the server.

Step 4: Fingerprint comparison in cipher

With the data from user, the server computes the ciphertext of $D_k^{'}$ as follows.

$$[\![G_2]\!] = [\![\sum_{i=1}^{n}(-2y_{i0} \cdot y_{ik})]\!] = \prod_{i=1}^{n}([\![-2y_{i0}]\!])^{y_{ik}} \tag{23}$$

$$[\![G_3]\!] = [\![(y_{ik})^2]\!] \tag{24}$$

$$[\![D_k^{'}]\!] = [\![G_1 + G_2 + G_3]\!] = [\![G_1]\!] \cdot [\![G_2]\!] \cdot [\![G_3]\!] \tag{25}$$

Then, the server returns encrypted $[\![D_k^{'}]\!]$ $(k = 1, 2, \ldots, M)$ to the user.

Step 5: Location identification

The user decrypts every $D_k^{'}$ from $[\![D_k^{'}]\!]$ using K_s. In the similar fuzzy way that the server does in Section III, the user identifies his location finally via (12) and (5) by replacing D_k by $D_k^{'}$.

B. Analysis

In our privacy-preserving CSI fingerprint localization scheme, the confidentiality of the encrypted data is guaranteed by Paillier cryptosystem. Combining with the interactions among the user, server and APs, we will analyze the privacy-preserving effects on them respectively.

1) For the user: Firstly, because of the obfuscation in service requesting, the user masks his true CSI towards untrusted APs and other potential attackers in the environment. Secondly, as the measurements transmitted to the server have been encrypted asymmetrically, it is unfeasible for the server to know the user's location from his location retrieval request. Thirdly, by the same reason of encryption, attackers cannot learn the user's location from the ciphertexts returned by the server. To sum up, the location privacy threat on the user can be addressed.

2) For the server: For each time of localization, the training signals are chosen by APs randomly and kept secret to the user. There is no latent consistency in the measurements sampled at the same location. Consequently, the attacker cannot steal the CSI fingerprint on the server by requesting localization service for multiple times. Although the reference locations are released to the public, they are irrelevant to the concrete values of CSI fingerprints in the database. So the data privacy threat on the server can be addressed.

3) For the AP: Since the training signals sent by APs are random and secret, the user and other potential attackers cannot learn their true CSIs when sampling measurements. So the locations of APs would not be inferred through CSI data. The location privacy threat on the AP can be addressed.

V. Simulation Experiments

As analyzed before, our proposed privacy-preserving mechanism is capable to protect the location privacy of the user and APs as well as the data privacy of the server in localization. So in this section, we will focus on evaluating the accuracy of the localization scheme. All our simulation experiments are carried out on the publicly accessible data set CRAWDAD [20], which contains over 9300 temporal CSI data measured in an real indoor environment as shown in Figure (4). For (almost) all pairs of locations, the CSI is measured for 5 times. We choose the former 4 measurements for fingerprint training, i.e. $m_k = 4$, and leave the last one for testing. The dimension of the CSI vector contributed by each AP is set to be $n_b = 6$. Considering that the accuracy enhancement of our localization scheme lies in the adoption of fuzzy algorithm and CSI fingerprints, so under the same privacy-preserving mechanism aforementioned, we will make comparisons with other schemes from two perspectives as follows.

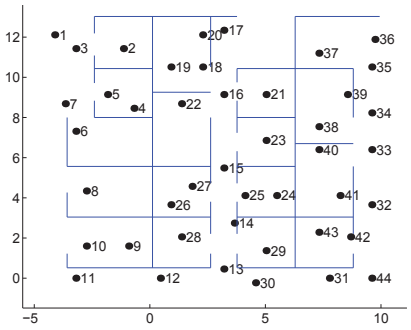

Fig. 4. The measuring environment of CRAWDAD [6]

A. Compared with KNN Algorithm

Given the same CSI fingerprint database, the localization errors of schemes with fuzzy algorithm and KNN algorithm are illustrated in Figure (5). For the KNN based scheme, we only show the condition when $K = 2$, because the performance of schemes with other values of K is even worse for this case. In the left subfigure, the APs' placements are selected intuitively, while in the right subfigure, the APs' placements are optimized by traversing all possible options for both of the schemes. The x-axis represents the number of APs employed for localization

(denoted as B), and the y-axix represents the localization error. From the figure, it can be seen that the localization error of our scheme with fuzzy algorithm is dramatically less than that of the scheme with KNN algorithm, and its decreasing rate is much faster than the counterpart. In the intuitive scenario, when $B = 2$, the mean error of our fuzzy scheme is 0.0223m, which can be negligible in practice. While for the KNN based scheme, even when $B = 5$, its mean error is still 2.5266m, which is 4 times more than the mean error of fuzzy scheme with only one AP. In the optimized scenario, the mean error of our fuzzy scheme is less than the KNN based scheme for over 2 orders of magnitude when $B \geq 2$.

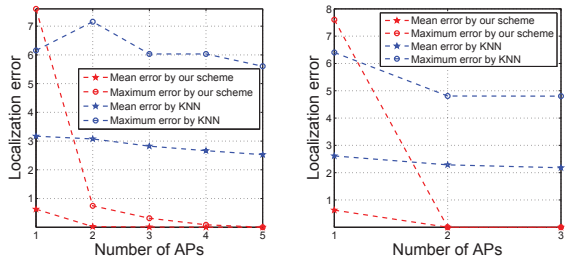

Fig. 5. Compared with the scheme with KNN algorithm

B. Compared with RSS Fingerprint

With the same fuzzy algorithm, the logarithmized localization errors of schemes with CSI fingerprint and RSS fingerprint are illustrated in Figure (6). Like the above, in the left subfigure, the APs' placements are selected intuitively, while in the right subfigure, the APs' placements are optimized for both of the schemes. The x-axis represents the number of APs B, and the y-axix represents the logarithmized localization error. From the figure, it can be seen that our scheme with CSI fingerprint surpasses the scheme with RSS fingerprint greatly in localization accuracy. In the intuitive scenario, from $B = 1$ to $B = 5$, the difference between the logarithmized mean errors of the RSS based scheme and our CSI based scheme increases from 0.7467 to 3.5421. In the optimized scenario, when $B = 2, 3$, both the mean and maximum errors of our CSI based scheme is less than those of the RSS based scheme for over 2 orders of magnitude.

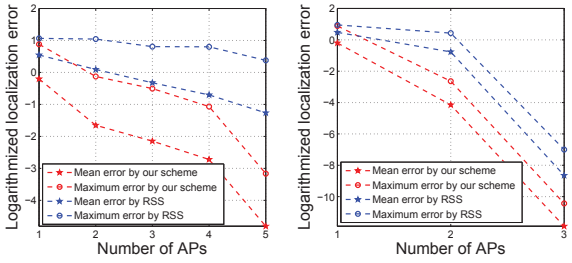

Fig. 6. Compared with the scheme with RSS fingerprint

VI. Conclusion

In this paper, we develop a privacy-preserving fuzzy localization scheme with CSI fingerprint using homomorphic encryption and fuzzy logic. It cannot only achieve better localization accuracy, but also protect the location privacy of the user and APs as well as the data privacy of the server. Through simulation experiments performed on CRAWDAD database, it is comfirmed that the proposed scheme is efficient.

References

[1] K. Pahlavan, X. Li, and J. Makela, "Indoor geolocation science and technology," *IEEE Communications Magazine*, vol. 40, no. 2, pp. 112-118, 2002.

[2] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Computing Surveys*, vol. 46, no. 2, pp. 25-32, 2013.

[3] S. Fang, Y. Liu, W. Shen, and H. Zhu, "Where are rou from confusing location distinction using virtual multipath camouflage," in *Proc. of ACM MobiCom '14*, pp. 225-236.

[4] G. Li, J. Teng, F. Yang, A. C. Champion, D. Xuan, H. Luan, and Y. Zheng, "EV-Sounding: A visual assisted electronic channel sounding system," in *Proc. of IEEE INFOCOM '14*, pp. 1483-1491.

[5] N. A. Khanbashi, N. Alsindi, S. Al-Araji, N. Ali, and J. Aweya, "Performance evaluation of CIR based location fingerprinting," in *Proc. of IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 2466-2471, 2012.

[6] N. Patwari, and S. K. Kasera, "Robust location distinction using temporal link signatures," in *Proc. of ACM MobiCom '07*, pp. 111-122, 2007.

[7] J. Zhang, M. H. Firoozz, N. Patwariz, and S. K. Kasera, "Advancing wireless link signatures for location distinction," in *Proc. of ACM MobiCom '08*, pp. 26-37, 2008.

[8] H. Liu, Y. Gan, J. Yang, S. Sidhom, Y. Wang, Y. Chen, and F. Ye, "Push the limit of WiFi based localization for smartphones," in *Proc. of ACM MobiCom '12*, pp. 305C316.

[9] S. Bai, and T. Wu, "Analysis of K-Means algorithm on fingerprint based indoor localization system," in *Proc. of IEEE 5th International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications (MAPE)*, pp. 44-48, 2013.

[10] Y. Chen, D. Lymberopoulosz, J. Liuz, and B. Priyanthaz, "FM-based indoor localization," in *Proc. of ACM MobiSys '12*, pp. 169-182.

[11] L. Ma, A. Y. Teymorian, and X. Cheng, "A hybrid access point protection framework for commodity wi-fi networks," in *Proc. of IEEE INFOCOM '08*, pp. 1894-1902.

[12] R. Shokri, G. Theodorakopoulos, J. Le Boudec, and J. Hubaux, "Quantifying location privacy," in *IEEE Symposium on Security and Privacy, 2011*, pp. 247-262.

[13] T. Wang, and Y. Liu, "Secure distance indicator leveraging wireless link signatures," in *Proc. of 2014 IEEE MILCOM*, pp. 222-227.

[14] A. Mariakakis, S. Sen, J. Lee, and K. H. Kim, "SAIL: Single access point-based indoor localization," in *Proc. of ACM MobiSys '14*, pp. 315-328.

[15] H. Li, L. Sun, H. Zhu, X. Lu, and X. Cheng, "Achieving privacy preservation in WiFi fingerprint-based localization," in *Proc. of IEEE INFOCOM '14*, pp. 2337-2345.

[16] T. Shu, Y. Chen, J. Yang, and A. Williams. Multi-lateral privacy-preserving localization in pervasive environments. in *Proc. of IEEE INFOCOM '14*, pp. 2319-2327.

[17] L. Li, G. Shen, C. Zhao, T. Moscibroda, J. H. Lin, and F. Zhao, "Experiencing and handling the diversity in data density and environmental locality in an indoor positioning service," in *Proc. of ACM MobiCom '14*, pp. 459-470.

[18] L. A. Zadeh, "Fuzzy Sets," *Information and Control*, vol. 8, pp. 338-353, 1965.

[19] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in *Proc. of ACM EUROCRYPT*, 1999.

[20] SPAN, "Measured channel impulse response data set," http://span.ece.utah.edu/pmwiki/pmwiki.php?n=Main.MeasuredCIRDataSet.