WiFi fingerprint releasing for indoor localization based on differential privacy

Yujia Zhu[1], Yu Wang[2], Qingyun Liu[1], Yang Liu[1], Peng Zhang[1]
[1] National Engineering Laboratory for Information Security Technologies,
Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
[2]National Computer Network Emergency Response and Coordination Center, Beijing, China
[1] {zhuyujia, liuqingyun, liuyang, pengzhang}@iie.ac.cn
[2]slimzczy@163.com

***Abstract*—WiFi fingerprint-based localization is regarded as one of the most promising techniques for indoor localization. However, this raises serious privacy concerns. Current approaches to mitigate the privacy concerns rely on the encryption with large calculation consumption. In this paper, we propose a data obfuscation mechanism based on the generalized version of differential privacy. We extend the standard definition to the indoor WiFi fingerprint data for spatial counting where the inputs belong to multiple dimensions of numerical data in a limited range. With a given privacy budget, the proposed method generalizes the original dataset, and then specializes it using differential privacy. As the designed novel scheme expand the range for specialization, the data set released by the proposed algorithm can yield better mining results. Furthermore, experimental results give out comparisons between non-uniform and uniform ε selection scheme, and find uniform ε selection scheme can fully use the privacy budget in our situation.**

***Index Terms*—Privacy Preserving, Differential Privacy, Generalizaton, indoor localization, WiFi.**

I. INTRODUCTION

Indoor location-based service has attracted much attention in recent years due to its social and commercial values. WiFi fingerprint-based indoor localization is one of the most popular localization techniques. It has been processed in two steps. A database of RSS (Received Signal Strength) measurements (also called radio map) made at a set of known locations can be initially assembled. The resultant database is used as the training set for a statistical learning model [1-2]. Then, at an online stage, the learning model is used to estimate a location from a given new set of RSS values.

Unlike mature outdoor positioning systems, the natural sensitivity, together with the high maintenance cost for indoor fingerprint database has created a great need for indoor LBS systems to find a method to prevent the disclosure of location information and the user's identity information [3]. Attackers who have the ability to get the user's WiFi signal fingerprints and the fingerprint database, using pattern matching methods are easily to compute the user's real physical location. While the existing privacy-aware methods like location obfuscation [4] and anonymity [5] based on GPS in open area cannot meet the needs in indoor environment.

Due to secure situations of fingerprint method in indoor environment, we have two key problems to solve. One is how to protect the user's RSS profile and the other is how to protect the fingerprint database. To guarantee a non-negligible level of privacy in WiFi fingerprint-based indoor localization, it is urgent need to obfuscate the data with controllable noise.

In this paper, we extend the standard definition to the indoor WiFi fingerprint data releasing for spatial counting based on differential privacy. We propose a differentially private releasing algorithm (inLocDiff) for balancing the requirements of practicality and utility based on fingerprint method. The idea is based on the observation that in the original fingerprint dataset, the numerical attributes varies in a specified range with spatial variation. The solution first partition space into delaunay triangulations according to POI (Position of Interest). Then a novel scheme expanded the range for generalization and specialization. By performing an anonymization algorithm for the non-interactive setting based on the generalization technique, the solution builds a classifier and report the count results of POI requests.

Three issues are to be addressed in this work:

- How to design a differential privacy releasing algorithm based on WiFi fingerprint method? Comparing to interactive setting, non-interactive setting can releasing the entire dataset, which can provide more flexibility for data mining task and is critical to the problem of sharing sensitive indoor signal data.
- We extend the standard definition to indoor WiFi fingerprint data releasing for spatial counting where the inputs belong to multiple dimensions

978-1-5386-3531-5/17/$31.00 ©2017 IEEE

of numerical data in a limited range. Based on the extended definition, we can use the generalization method.

- How to arrange and fully utilize privacy budget in such a releasing algorithm? We compare non-uniform and uniform budget selection scheme for selecting privacy budget in the experiments.

The remainder of this paper is organized as follows. Section II presents the related work. Method inLocDiff is proposed in Section III. Section IV presents the preliminaries. Section V presents the performance evaluation. Finally, section VI gives a conclusion.

II. Related Work

A. Indoor localization privacy mechanisms

Indoor localization privacy mechanisms can be categorized into terminal-based and network-based methods.

Terminal-based solutions calculate the location directly at the user-carried device, which hosts the localization algorithm, using sensor readings produced or downloaded locally. Open systems in this category include in-house Airplace[6] and Redpin.org.

Network-based approaches use location-dependent observations that are either monitored by the network infrastructure or collected by mobile devices in a terminal-assisted fashion. Most major indoor location-based services are currently network-based, including Google (Indoor), Trueposition (formerly Skyhook), Navizon.com, Infsoft.com, Indoo.rs, and Wifarer.com, are all completely network or cloud-based and as such. Several approaches have been proposed to protect location privacy in location-based services (e.g., k-anonymity, and cryptosystem [7]). K-anonymity guarantees that a querying user u is indistinguishable among at least k−1 others [8]. The state-of-the-art k-anonymity approaches [9] [10] mainly rely on historical data and derive sanitized trajectories from a set of real trajectories. Similarly, following an in-house strategy, [11] proposes an integrated platform for applying data mining and privacy-preserving querying over mobility data. [12] proposes a Privacy-Preserving WiFi Fingerprint Localization scheme (PriWFL) that can protect both the client's location privacy and the service provider's data privacy based on Paillier cryptosystem as cryptographic primitive.

B. Differential Privacy

The motivation is to release statistical information without compromising the privacy of the individuals [13]. Because differential privacy provides a rigorous and provable privacy definition, researchers have shown an increased interest on it in many applications [14-16]. The main advantage of this notion is that this mechanism can achieve the privacy preserving without concerning any background knowledge.

In the context of location privacy preserving, Ho et al. [17] proposed an ε-differentially private geographic location pattern mining approach by using a partition-aggregate framework. This framework first utilized the spatial decomposition to limit the number of records within a localized spatial partition and then applied noise-based clustering to discover the interesting patterns in location datasets. Laplace noise was added in both steps with the aim to mask the count of records in a region and the centroid of these records. Dewri [10] incorporated differential privacy to k-anonymity and proposed a mechanism to ensure that the probability of output the same obfuscated location from any set of anonymized k locations is similar.

III. Preliminaries

A. Differential Privacy

Definition 3.1 (ε-Differential Privacy) [18]. A randomized algorithm Ag is differentially private if for all data sets S and S', where their symmetric difference contains at most one record (i.e., $|S \Delta S'| \leq 1$), and for all possible anonymized data sets $\hat{S}$:

$$\Pr[\mathrm{Ag}(S) = \hat{S}] \leq e^{\varepsilon} * \Pr[\mathrm{Ag}(S') = \hat{S}], \quad (1)$$

where the probabilities are over the randomness of Ag.

Intuitively, differential privacy guarantees that no individual tuple can significantly affect the released information: the output distribution generated by Ag is nearly the same, whether or not that tuple is present in the dataset. The most common technique for designing differentially-private algorithms was proposed in [19].

Definition 3.2 (Laplace Mechanism) Let f(S) denote a numeric function over dataset S. An ε-differentially private mechanism for releasing f is to publish L(S) = f(S)+X, where X is a random variable drawn from the Laplace distribution Lap(Δ(f)/ε).

Definition 3.3 (Exponential Mechanism) Let u(S, r) be a utility function of dataset S that measures the score of outputting r. Then an exponential mechanism M is ε-differential privacy if M(S)={return r with the probability exp(εu(S,r) /2Δu)}, where Δu denotes the sensitivity of u.

B. Top-Down Specialization

Let S = $\{s_1, s_2, ..., s_n\}$ be a multiset of records, where each record s_i represents the information of an individual with d attributes $MAC = \{mac_1, mac_2, ..., mac_m\}$. We assume that each attribute mac_i has a finite domain, denoted by Ω(mac_i). The domain of S is defined as Ω(S) = Ω(mac_1)×, … ,×Ω(mac_d). To anonymize a data set S, generalization replaces a value of an attribute with a more general value.

C. Problem Statement

We consider the following problem. Given a WiFi fingerprint dataset S, and user's realtime RSS profile set. Denote as $S=\{\{p_1,r_{11},\ldots,r_{1m}\},\ldots,\{p_n,r_{n1},\ldots,r_{nm}\}\}$ where, r_{ij} is a vector of readings from m MAC (Media Access Control) addresses within query range region p_i (p_i is a label) generated according to POI. POI set can be expressed as $Q=\{(x_1,y_1),(x_2,y_2),...,(x_s,y_s)\}$ .The signal strength from each MAC address is a numerical attribute, which has a finite domain, denoted by Ω(RSS), usually between -45 dbm to -100 dbm. Let $MAC=\{mac_1,mac_2,...,mac_m\}$ denotes m attributes. So a WiFi fingerprint dataset can be expressed as $\begin{Bmatrix}\{p_1,r_{11},\cdots,r_{1m}\}\\ \{p_2,r_{21},\cdots,r_{2m}\}\\ \cdots\\ \{p_n,r_{n1},\cdots,r_{nm}\}\end{Bmatrix}$, while $\begin{Bmatrix}\{r_{11},\cdots,r_{1m}\}\\ \{r_{21},\cdots,r_{2m}\}\\ \cdots\\ \{r_{q1},\cdots,r_{qm}\}\end{Bmatrix}$ can be expressed as user's realtime RSS profile set.

Our aim is to accurately answer count queries within POI query range p_i , while protect the user's RSS profile and the fingerprint database.

IV. PROPOSED METHOD

We propose a data obfuscation mechanism inLocDiff based on the generalized version of differential privacy. Our algorithm is divided into three main parts: (i) query range definition, (ii) generalization and (iii) classification and counts. The query range definition is able to find a proper landmark to hide POI itself using voronoi diagram method. After query range has been defined, inLocDiff will conduct a Top-Down specialization process to find obscure data for WiFi fingerprint sets. At last inLocDiff will tell the results according to the query range generated by voronoi diagram.

A. Overview of inLocDiff

Algorithm 1 gives the pseudocode of how to integrate generalization with differential privacy. The inputs are WiFi fingerprint dataset S with privacy budget ε, POI set Q normalize parameters χ, initial ε_1 , and increment θ_ε . The output $\hat{S}$ is a general dataset with noisy count for classification analysis.

As shown in step 1-4, inLocDiff algorithm first partition the whole spatial area into several query range region according to the POI positions. Each query range region has a label. Step 5 is a normalization procedure with normalize parameters χ. Step 6-10 is the initial Top-Down specialization procedure.

From step 11 to 15, we generate the specialization schemes solution for the entire dataset. Specifically, for each round, we use an exponential mechanism to select a candidate for specialization in each round (step 12). Then determine the split value for each new candidate (step 13). We allocate the selection probability to each scheme according to their utility scores.

Budget Selection Scheme: A natural strategy is to set $\varepsilon_i=\varepsilon/h$, which h stands for the specialization rounds. Prior work that finds counts in trees (e.g. [19]) has used this model. It seems that we should reduce h to reduce the noise. But the height of a tree will improve the matching accuracy. In order to optimize the problem, we consider a non-uniform budget selection scheme. We do not have to use the same ε_i for all nodes in each round. For WiFi fingerprint, the workload of each specification round is known as a priori. As we all know, specialization round can more affect the classification results when the very first rounds the classifications are correct. We analyze it to determine budget in each round and give a non-uniform budget selection scheme (step 14). In each round, we add increment θ_ε to ε_i . Leaves get the largest budget in all specialization rounds. In step 16 to 18, we add the Laplace noise to the counts in each equivalence group. Finally, the sanitized dataset $\hat{S}$ is generated.

Algorithm 1. inLocDiff Algorithm
Input: S, ε, Q, χ, θ_ε and initial ε_1

Output: $\hat{S}$
1: Voronoi Polygon VP(Q) = contains(Q);
2: **for each** record in S **do**
3: idx = min(dis(s_i .loc,Q)+Lap(4/ε)), let label p_i be VP(idx);
4: **end for**
5: Normalize all the attributes in S according to χ;
6: Initialize set M with topmost value of every attribute r in S;
7: **for each** m in M **do**
8: $E(S,m)=e^{\frac{\varepsilon_1(S,m)}{2\Delta u}}$
9: **end for**
10: Update set M;
11: **while** ($\sum\varepsilon_i$ <ε/2 && $\varepsilon_i>0$)
12: Select m from M with probability E(S, m);
13: Specialize m with probability E(S, m) and update set M;
14: $\varepsilon_{i+1}=\varepsilon_i+\theta_\varepsilon$; $\varepsilon=\varepsilon-2\varepsilon_i$;
15: **end while**
16: **for each** equivalence group p_i do
17: count (p_i) = count (p_i) + Lap (4/ε);
18: **end for**
19: **return** $\hat{S}=\bigcup p_i$

We next elaborate the key steps of the algorithm.

B. Query Range Definition

We use voronoi diagram to divide indoor space into disjoint polygons where the nearest neighbor of any POI inside a polygon is the generator of the polygon.

Definition 4.1 (Query Range) Given a point set $Q=\{q_1,q_2,...,q_n\}$ as POI positions, a query range for Q is defined as follows:

$$1.VD(S)=\bigcup_{i=1}^{k} VP(Q_i);$$
$$2.\bigcap_{i=1}^{k} VP(Q_i)=\phi;$$
$$3.\forall x \in VP(Q_i), dist(x,Q_i) < dist(x,Q_j) \text{ where } Q_i \neq Q_j. \quad (2)$$

Through the POI query range, we divide the geographical space, and get the classification category to be tested.

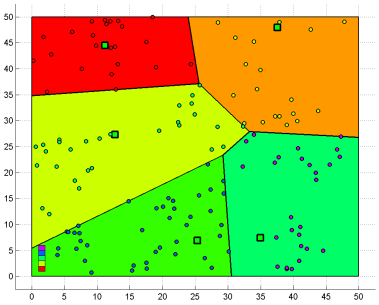


Fig. 1. Example of 5 POI Classification
(The green square with black box is POI points, all the dots represent different geographical classification according to Definition 4.1)

C. Generalization and Specialization

The process of replacing an exact values into a generalized value, is called Generalization. Specialization is the reverse process of Generalization means creating general value with a more specific one.

1) Candidate Selection. We use an exponential mechanism to select a candidate for specialization in each round. We adopt the Max utility function [9], which represent the maximal number of records in the classes.

Let Sv denote the set of records in S generalized to the RSS value r. Let $|S_r^p|$ denote the number of records in Sr having the label value p ∈ Ω(mac_i). Note that $|S_r| = \sum_c |S_c|$, where c ∈ child(r). Then, we get

$$Max(S,r) = \sum_{c \in child(r)} (\max_p |S_c^p|). \quad (3)$$

The Max utility function counts the number of records in each class, so the sensitivity is 1.

2) Split Value. Once a candidate is determined, inLocDiff splits the records into child partitions. We again use an exponential mechanism to determine the split value. The split value is determined by sampling a value uniformly from the interval. Thus, the probability of selecting a value $r_i \in \Omega(mac_i)$ is

$$\frac{\exp(\frac{\varepsilon}{2\Delta u}u(S,r_i))}{\sum_{r\in\Omega(mac_i)}\exp(\frac{\varepsilon}{2\Delta u}u(S,r))} \quad (4)$$

3) Noisy Counts. Publishing the exact counts of these groups does not satisfy differential privacy since for a different data set S', the counts may change. This change can be easily offset by adding noise to the count of each group (**Definition 3.2**). As discussed earlier, the sensitivity of count query is 1; therefore inLocDiff adds Lap(4/ε) noise to each true count of the groups (Line 17). We post-process the noisy counts by rounding each count to the nearest non-negative integer.

D. Privacy Analysis

To analyze the privacy, we apply two widely used composite properties of the privacy budget [19]: the sequential and the parallel composition.

Lemma 4.1 (Sequential composition) Let each M_i provide ε_i-differential privacy. A sequence of $M_i(S)$ over the data set S provides ($\sum \varepsilon_i$)-differential privacy [19].

Lemma 4.2 (Parallel composition) Let each M_i provide ε_i-differential privacy. A sequence of $M_i(S_i)$ over a set of disjoint data sets S_i provides Max(ε_i)-differential privacy [19].

Theorem 4.1. inLocDiff is ε-differentially private.

Proof. (Sketch) The inLocDiff algorithm contains three private operations: 1) query range definition, 2) generation and specialization, and 3) noisy counts publishing. The privacy budget ε is consequently divided into three pieces, as illustrated in TABLE I.

TABLE I Privacy Budget Allocation in inLocDiff *Algorithm*

Operations	Privacy budget
query range definition	ε/4
generation and specialization	ε/2
noisy counts publishing	ε/4

-The querying range definition operation is performed on the whole dataset with the privacy budget ε/4. According to **Definition 3.2**, this operation preserves ε/4-differential privacy.

-The generalization and specialization operation applies the Exponential Mechanism. This step uses privacy budget ε_i and thus, candidate selection step guarantees ε_i-differential privacy for each iteration. Then the algorithm determines the split value for each new candidate. All records in the same partition have the same generalized values; therefore, each partition can only contain at most one candidate. Thus, determining the split value for the new candidates requires ε_i privacy budget for each iteration due to the parallel composition property. Totally, it applies $\sum \varepsilon_i = 2h\varepsilon_1 + h(h-1)\theta_\varepsilon (\leq \frac{\varepsilon}{2})$-differential privacy.

-The noisy counts publishing adopts the Laplace Mechanism. The privacy budget for the Laplace distribution is ε/4. According to the definition of Laplace Mechanism (**Definition 3.2**), this step satisfies ε/4-differential privacy.

Consequently, we can conclude that the proposed inLocDiff algorithm preserves ε-differential privacy (**Lemma 4.1**).

V. EXPERIMENTAL EVALUATION

In this section, we conduct a set of experiments to evaluate the performance of inLocDiff.

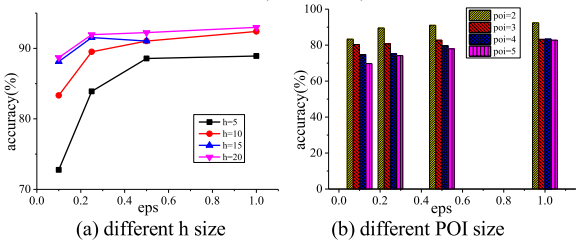
- How does the privacy budget affect the performance of inLocDiff? We examine the trade-off between the utility and the privacy of inLocDiff on benchmark datasets by varying it in a wide range. The performance is evaluated in term of classification accuracy.
- How does the specialization iterations affect the hierarchy tree construction? We compare the performance of Max utility function in various iteration times and analyze the most suitable specialization rounds and the accuracy trends for different POI sizes.
- How does the inLocDiff perform comparing with other related algorithm? We compare inLocDiff with DiffGen [19], the well known differentially private algorithm for building classifiers. Though DiffGen is not designed for multiple dimensions of numerical data. We did some adaption.

To the best of our knowledge, no real large-scale indoor localization data sets have been publicly released to date. Therefore, in our experiments, we generate the indoor localization objects using MATLAB. By default, the test area is 50m*50m. The number of WiFi AP (Access Point) is 6. Algorithms are implemented in MATLAB R2013a and run on a Lenovo X220 laptop with an Intel Core i7 processor and 8 GB main memory.

As a training dataset for building classifiers, the simulated data has 10000 fingerprint records with 6 numerical attributes. The record label column is a two dimensional cartesian coordinates. In addition, 5000 records are extracted as the test dataset. The classification test processes are executed on Weka with the classifier J48. Each experiment is executed for 10 rounds and the results are averaged over these.

A. Impact of Privacy Budget

To get a comprehensive examination of inLocDiff, we evaluate its performance under various privacy preserving levels by vary $\varepsilon \in \{0.02, 0.05, 0.1, 0.5\}$. In order to test the impact of specialization rounds, $\varepsilon_i = \varepsilon / 4h(\theta_\varepsilon = 0)$ is set to be uniform. Specialization rounds are set to $h \in \{5, 10, 15, 20\}$. We consider the number of POI are set to {2, 3, 4, 5}.



(a) different h size (b) different POI size

Fig. 2. Impact of Privacy Budget

Fig. 2 shows the results that we use Max as utility function [20]. It is observed that given a fixed number of specializations iteration h, the classification accuracy increases with the increasing of the privacy budget ε. It shows that a higher privacy budget ensures better specialization schemes can be selected.

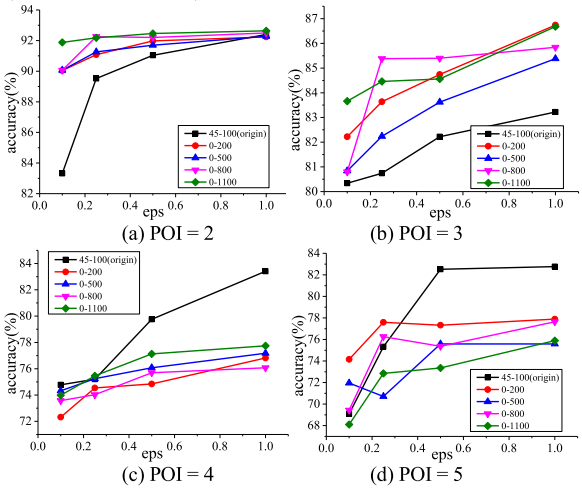
Fig. 2(a) shows the results when h takes other values. It can be observed that when we fix the privacy budget epsilon, a higher value of h results in better classification accuracy. This is because when we perform multiple specialization iteration on dataset, the records will be more specific, which enhance the classification accuracy.

But when h achieves a threshold, the accuracy is decreasing. It is because the number of equivalence groups increases rapidly with the increasing of h, a higher magnitude of noise is added to the counts of each group. The benefit from specialization will be offset by large noise. So we should set a threshold for h. In this dataset, we find that the accuracy achieve their peaks when h = 10.

Similar trends are also observed in Fig. 2(b) that given a fixed number of POI counts, the classification accuracy increases with the increasing of the privacy budget ε. And the accuracy decreased when POI counts increased.

B. Impact of Dilusion parameter

In this section, we examine the dilusion parameterχaffect the classification accuracy. The dilution range is set to [0, 200], [0, 500], [0, 800] and [0, 1100]. And the origin [45,100] is the adaption of DiffGen[19] algorithm. The privacy budget is still set to {0.1, 0.25, 0.5, 1.0}.



(a) POI = 2 (b) POI = 3

(c) POI = 4 (d) POI = 5

Fig. 3. Impact of Normalization Parameter

The experimental results are shown as Fig. 3(a)-(d). It can be observed that when POI counts are set to be 2 and 3, a higher value of normalization range result in better classification accuracy. This is because when we add noises on a larger data, the less influence will affect the data itself, which enhance the classification accuracy.

But when POI counts are set to be 4 and 5, the accuracy is definitely decreasing according to the expansion section. It is because the number of equivalence groups increases rapidly with the increasing of POI counts, a higher magnitude of noise is added to the counts of each group. The benefit from expansion will be offset by large noise.

C. *Impact of Different ε Selection Scheme*

We examine the different ε selection schemes as 1) Non-uniform ε selection scheme. We divide this scheme into ε increasing scheme and ε decreasing scheme; 2) Uniform ε selection scheme. We average privacy budget in each round.

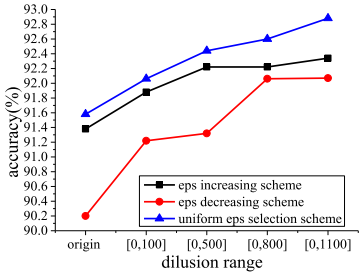


Fig. 4. Impact of Different ε (expressed as eps) Selection Scheme

As seen in Fig. 4 when the POI is set to 2, ε set to 1, classification accuracy is different under three different privacy budget selection policy. On one hand, ε decreasing scheme and uniform ε selection scheme is obviously better than ε increasing scheme. And ε decreasing scheme is slightly better than uniform ε selection scheme. This can be seen the importance of the initial division is stronger than the leaves division. On the other hand, we see that dilution method can improve classification accuracy.

VI. CONCLUSION

Differential privacy can be applied in indoor fingerprint matching analysis showed in inLocDiff with the following contributions: 1) design a non-interactive releasing algorithm for spatial counting in indoor environment using differential privacy, 2) present a novel dilution operation to deal with the numerical attributes, and 3) the algorithm gives out a comparison between non-uniform and uniform ε selection scheme, and find uniform ε selection scheme can fully use the privacy budget in our situation.

ACKNOWLEDGMENT

This work was supported by National Key R&D Program 2016 (Grant No. 2016YFB0801304), National Natural Science Foundation of China (No. 61402464), and Youth Innovation Promotion Association CAS.

REFERENCES

[1] Brunato M, Battiti R. Statistical learning theory for location fingerprinting in wireless LANs[J]. Computer Networks, 2002, 47(6):825–845.

[2] Kaemarungsi, Kamol, Krishnamurthy, Prashant. Modeling of indoor positioning systems based on location fingerprinting[J]. Proceedings - IEEE INFOCOM, 2010, 2(2):1012-1022 vol.2.

[3] Zhu Y, Zhai L. Location Privacy in Buildings: A 3-Dimensional K-Anonymity Model[C]// International Conference on Mobile Ad-Hoc and Sensor Networks. IEEE, 2014:195-200.

[4] Duckham M, Kulik L, "A Formal Model of Obfuscation and Negotiation for Location Privacy," Pervasive Computing. 2005, 3468: pp. 152-170.

[5] Shin, K. G., Xiaoen, Ju, Zhigang, Chen and Xin, Hu, "Privacy protection for users of location-based services," Wireless Communications, IEEE, 2012, 19(1): pp. 30-39.

[6] Laoudias C, Constantinou G, Constantinides M, et al. The Airplace Indoor Positioning Platform for Android Smartphones[C]// IEEE International Conference on Mobile Data Management. 2012:312-315.

[7] Samarati P. Protecting Respondents' Identities in Microdata Release[J]. IEEE Transactions on Knowledge & Data Engineering, 2001, 13(6):1010-1027.

[8] Abul O, Bonchi F, Nanni M. Anonymization of moving objects databases by clustering and perturbation[J]. Information Systems, 2010, 35(8):884-910.

[9] LATANYA SWEENEY. k-ANONYMITY: A MODEL FOR PROTECTING PRIVACY[J]. IEEE Security and Privacy Magazine, 2002, 10(5):1-14.

[10] Dewri R, Ray I, Ray I, et al. On the Formation of Historically k-Anonymous Anonymity Sets in a Continuous LBS[M]// Security and Privacy in Communication Networks. 2009:71-88.

[11] Pan X, Xu J, Meng X. Protecting Location Privacy against Location-Dependent Attacks in Mobile Services[J]. IEEE Transactions on Knowledge & Data Engineering, 2011, 24(8):1506-1519.

[12] Li H, Sun L, Zhu H, et al. Achieving privacy preservation in WiFi fingerprint-based localization[C]//IEEE INFOCOM 2014-IEEE Conference on Computer Communications. IEEE, 2014: 2337-2345.

[13] Dwork C. Differential Privacy[J]. Icalp, 2006, 26(2):1-12.

[14] Mcsherry F, Mahajan R. Differentially-private network trace analysis[J]. Acm Sigcomm Computer Communication Review, 2011, 41(4):123-134.

[15] Xiong P, Zhu T, Niu W, et al. A differentially private algorithm for location data release[J]. Knowledge & Information Systems, 2015:1-23.

[16] Dwork C, Mcsherry F, Nissim K. Calibrating Noise to Sensitivity in Private Data Analysis[C]// Conference on Theory of Cryptography. Springer-Verlag, 2006:265-284.

[17] Ho S S, Ruan S. Preserving Privacy for Interesting Location Pattern Mining from Trajectory Data[J]. Transactions on Data Privacy, 2013, 6(1):87-106.

[18] Dewri R. Local Differential Perturbations: Location Privacy under Approximate Knowledge Attackers[J]. IEEE Transactions on Mobile Computing, 2013, 12(12):2360-2372.

[19] Mohammed N, Chen R, Fung B C M, et al. Differentially private data release for data mining[C]// ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, Ca, Usa, August. 2011:493-501.