Software Engineering and Applications 软件工程与应用, 2025, 14(1), 73-85
Published Online February 2025 in Hans. https://www.hanspub.org/journal/sea
https://doi.org/10.12677/sea.2025.141008

Hans汉斯

基于同态加密的室内指纹定位隐私保护方案

张　健，乐燕芬

上海理工大学光电信息与计算机工程学院，上海

收稿日期：2025年1月9日；录用日期：2025年2月14日；发布日期：2025年2月26日

摘　要

在室内定位服务中，如何保护用户及位置服务提供商的隐私安全和提高定位实时性一直是一个具有挑战性的问题。已有的方式大都采用欧氏距离结合半同态加密算法来完成，存在定位实时性不高、双方计算开销大等问题，基于此本文提出一种结合Kumar-Hassebrook距离、半同态加密算法及其安全点积性质的定位方案，在提高定位实时性的同时，能实现对用户位置信息和服务商指纹及位置数据隐私的保护。方案采取KH距离来匹配定位用户与指纹库中指纹数据的相似度，得到最近的K个最近邻参考点；在最近邻匹配中引入了半同态加密算法，保护用户和服务商的指纹数据隐私；同时，利用其安全点积性质实现了对服务商的指纹库坐标数据的隐私保护。为降低时间开销，引入了分簇聚类和模糊簇匹配，在提高定位实时性的同时可模糊服务器端对用户所在真实的簇的判断。从理论上对所提方案的安全性，时间开销及定位性能进行了分析，并在公共数据集中进行了性能评估。与同类加密算法比较，在不降低定位性能及安全性的前提下，该方案进一步地降低了时间开销。

关键词

WiFi指纹定位，隐私保护，Kumar-Hassebrook距离，半同态加密

Privacy Preserving Scheme of Indoor Fingerprinting Locatization Based on Homomorphic Encryption

Jian Zhang, Yanfen Le

School of Optical-Electrical and Computer Engineering, Shanghai University of Technology, Shanghai

Received: Jan. 9th, 2025; accepted: Feb. 14th, 2025; published: Feb. 26th, 2025

Abstract

How to protect the privacy security of users and location service providers and improve the real-time location performance has always been a challenging problem in indoor location services.

文章引用: 张健, 乐燕芬. 基于同态加密的室内指纹定位隐私保护方案[J]. 软件工程与应用, 2025, 14(1): 73-85.
DOI: 10.12677/sea.2025.141008

Based on this, this paper proposes a solution that combines Kumar-Hassebrook distance, semi-homomorphic encryption algorithm and its security point product properties to protect the location and fingerprint data privacy of users and service providers and improve the real-time performance of positioning. The KH distance is used to match the similarity between the positioning user and the fingerprint data in the fingerprint database, and the nearest K nearest neighbor reference points are obtained. A semi-homomorphic encryption algorithm is introduced in Nearest Neighbor Matching to protect the privacy of fingerprint data of users and service providers. At the same time, the privacy protection of the fingerprint database coordinate data of the service provider is realized by using its secure dot product nature. In order to reduce the time overhead, clustering and fuzzy cluster matching are introduced, which can improve the real-time positioning and blur the judgment of the real cluster where the user is located on the server. Theoretically, the security, time overhead and positioning performance of the proposed scheme are analyzed, and the performance evaluation is carried out in the public dataset. Compared with similar encryption algorithms, the proposed scheme further reduces the time overhead without reducing the positioning performance and security.

Keywords

WiFi Fingerprint Positioning, Privacy Protection, Kumar-Hassebrook Distance, Semi Homomorphic Encryption

1. 引言

随着移动智能终端的广泛应用，位置数据成为重要的一种数字资源，也使得各类基于位置的服务成为人们生活中不可或缺的一部分。而室内定位服务支撑着室内环境下的众多使用场景，例如室内导航、室内监控等，这些服务具有越来越高的商业价值，并受到学术界的广泛关注[1]。其中基于蓝牙[2]，UWB [3]，WiFi [4]信号的室内定位技术受到广泛关注和研究。因部署在室内的无线网络 WiFi 为室内定位提供了部署成本低，便捷性高及信号易获取等一系列优点，基于 WiFi 信号的指纹定位成为其中最关注的技术之一[5] [6]。该定位技术通常有两个阶段，离线阶段和在线阶段。在离线阶段服务商通过在定位区域内设置的参考点(reference point, RP)上测得接入点(access point, AP)信号值形成指纹信息，并结合该 RP 的物理位置构建离线位置指纹。在在线阶段，用户获取自身指纹信息，构建位置请求并上传至服务器。服务商通过位置估计算法，如利用 KNN 算法或机器学习算法来计算出用户的估计位置。而随着基于室内位置的各类服务的普及，人们在获取高精度的室内定位服务时，对隐私保护有了更高的要求[7]。用户的位置数据中通常包含有个人敏感信息，假如被不可信的服务商或恶意攻击者利用，可能对用户造成经济或名誉损失。同时，恶意用户可能通过多次定位请求获取服务器的指纹库，侵害服务器的数据隐私。因此，在提供定位服务过程中保护用户和服务商的数据隐私，是实现安全室内定位的关键所在。

目前，已经有了很多方法来实现室内定位中的隐私保护，比如：*K* 匿名(K-anonymity) [8] [9]、加密算法[10]-[12]等。这些算法通常用来保护用户的位置隐私和请求隐私，在文献[11]中，作者首次在用户位置隐私的保护中加入 *K* 匿名算法，这使得服务器无法精准定位用户的具体位置，从而保护了用户的位置隐私。Li 等[13]在室内定位中的隐私保护领域提出了创新性地引入了 Paillier 算法，利用 Paillier 算法的加法同态和数乘同态特性实现了用户端与服务器端的指纹数据的隐私保护。虽然 Paillier 算法安全性较高，但

在服务器端的指纹库规模较大时，这个计算开销甚至会达到几十秒，导致定位效率低，无法保证用户请求定位的实时性。

已有的各类隐私安全定位方案大都着力于对用户位置数据和服务商指纹数据的隐私保护，而定位方案中涉及的指纹数据的位置信息表直接交给用户或第三方。保护指纹数据的位置信息可加强服务商的数据安全，目前也有一些文献对此展开了研究。文献[14]提出一种 K 匿名结合 Paillier 的加密方法，并将部分计算外包给云服务提供商来降低服务提供商的计算开销，但该方案使得用户端的计算开销增大了 K 倍，同时用户的位置隐私实质仅是由 K 匿名方法保障。在文献[15]中，作者利用室内定位密码协议提出一种针对恶意用户破解服务商指纹库隐私的解决方案，该方案通过混淆电路，混淆函数等方式来具体实现，在保护用户隐私的前提下完成对定位双方的隐私保护，但是时间开销略大且需要三方中任意两方不能合谋。

为了解决上述问题，本文基于 Kumar-Hassebrook (KH)距离[16]结合半同态加密算法的同态特性提出一种两方室内定位隐私保护方案 KH-KNN。方案兼顾定位服务的安全、效率和精度，所提基于 KH 距离的定位算法保证了用户的定位精度，同时对半同态加密算法安全点积特性的巧妙应用，实现了服务商的坐标数据隐私安全。为进一步提高定位效率，本方案引入了分簇聚类及模糊簇匹配的思想。

2. 室内定位和隐私保护的相关工作

2.1. 基于 WiFi 指纹的室内定位算法

基于 WiFi 指纹信号的室内定位有 KNN 定位，是传统的定位方法之一，其主要定位过程主要分为在线和离线两个阶段。

在离线阶段，服务商通过在一个区域内设置 RP 和 AP 来采集指纹，采集到的指纹组成离线指纹库 D。指纹库中的每一条指纹如公式(1)所示：

$$\varphi_i = \left(P_i, F_i\right) = \left(x_i, y_i, f_{i,1}, f_{i,2}, \cdots, f_{i,M}\right), \forall i = \left(1, 2, \cdots, N\right) \tag{1}$$

式中 $P_i = \left(x_i, y_i\right)$ 表示第 i 个 RP 的物理位置，$F_i = \left(f_{i,1}, f_{i,2}, \cdots, f_{i,M}\right)$ 表示第 i 个 RP 所采集到的指纹信号强度数据，$f_{i,j}$ 表示在 RP 上的第 i 个位置接收到来自第 j 个 AP 的信号强度数值，这个数值经常表示在多次采样中的平均值，通常我们把未采集到指纹信号强度的位置设为 0 或者是−95 dBm。N 表示在此区域内 RP 数量的总数，M 表示在此区域 AP 数量的总数。

在线阶段，用户采集到的自身指纹信号向量可以表示为公式(2)：

$$F_t = \left(f_{t,1}, f_{t,2}, \cdots, f_{t,M}\right) \tag{2}$$

在 KNN 室内定位算法中，服务商通过匹配算法获得与用户指纹最相似的参考点，通常可用欧氏距离来计算指纹信号的距离，计算公式如公式(3)所示：

$$d_{i,t} = \left\|F_i - F_t\right\|^2 = \sum_{j=1}^{M}\left(f_{i,j} - f_{t,j}\right)^2 = \sum_{j=1}^{M} f_{i,j}^2 + \sum_{j=1}^{M}\left(-2 f_{i,j} \cdot f_{t,j}\right) + \sum_{j=1}^{M} f_{t,j}^2 \tag{3}$$

式中 $d_{i,t}$ 表示第 i 个位置的 RP 的指纹强度信号与请求定位的用户的指纹强度信号的空间距离，$\|\bullet\|$ 表示欧氏距离。KNN 算法选取距离最近的 K 个 RP，由这些 RP 点的坐标得到用户的估计位置。计算公式如公式(4)：

$$\left(\dot{x}, \dot{y}\right) = \frac{1}{k}\sum_{i=1}^{k}\left(x_i, y_i\right) \tag{4}$$

式中 $\left(\dot{x}, \dot{y}\right)$ 表示请求定位用户的估计位置，$\left(x_i, y_i\right)$ 表示选择的 K 个 RP 的坐标位置。

目前已有的隐私定位方案大都基于欧式距离展开研究[17] [18]，有学者研究发现，基于 KH 距离的

KNN 匹配算法的成功率和准确率都高于欧氏距离[14]。因此本文基于 KH 距离匹配用户的最近邻。计算公式如公式(5)所示：

$$d_{KH}\left(F_t,F_i\right)=\frac{\sum_{j=1}^{M}\left(f_{t,j}\cdot f_{i,j}\right)}{\sum_{j=1}^{M}f_{t,j}^2+\sum_{j=1}^{M}f_{i,j}^2-\sum_{j=1}^{M}\left(f_{t,j}\cdot f_{i,j}\right)} \tag{5}$$

KH 距离越大，表示用户与参考点的指纹越相似。

2.2. 同态加密算法

在室内定位过程中，常常利用密码学来保护定位双方的数据隐私，同态加密算法作为一种重要的技术手段在这个过程中扮演着重要的角色，如 Paillier、DGK、EIGamal 等。

同态加密有如下性质，[]表示加密过程：

a. 加法同态：

设有两个密文$[m_1]$，$[m_2]$，则明文 m_1，m_2 相加的密文可由这两个密文相乘得到。

$$[m_1]\times[m_2]=[m_1+m_2] \tag{6}$$

b. 标量积同态：

设有明文 m_1 和密文$[m_2]$，则明文 m_1，m_2 相乘的密文可由标量积同态得到。

$$[m_2]^{m_1}=[m_1\times m_2] \tag{7}$$

其中 Paillier 加密技术是加法同态的代表性方案之一，作为一种公钥加密的概率非对称算法，具有加法同态、标量积同态及安全点积等特性，被广泛应用于电子现金交易、安全电子投票，信息安全技术等领域。利用其同态特性，除了完成简单的加密操作，该技术也可以在加密域实现各种复杂计算。在室内定位中，通过使用 Paillier 加密技术，可以使得用户和服务器的隐私数据得到保护。

Paillier 算法分为三个部分[19]，生成秘钥，数据加密，数据解密。

生成的密钥分成公钥和私钥，公钥用于对数据进行加密，由 n，g 两个参数构成，私钥用于对加密数据进行解密，由 λ，μ 两个参数构成。假设 m 是待加密的明文数据，m 通过公钥加密得到密文$[m]$，具体加密过程可由公式(8)表示：

$$[m]=g^m s^m \bmod n^2, s\in Z_n^* \tag{8}$$

利用私钥对密文$[m]$解密得到明文 m，具体解密过程可由公式(9)表示：

$$m=\left(\left[m\right]^{\lambda} \bmod n^2-1\right)\Big/ n*\mu \bmod n \tag{9}$$

同时 Paillier 算法包含了同态加密的性质，加法同态，标量积同态及安全点积。

安全点积⊙：

设有 M 维向量 v 和 m，及 M 个密文$[m_1]\cdots[m_M]$，则 v 和 m 点积的密文可由安全点积得

$$\begin{aligned}\boldsymbol{v}\odot[\boldsymbol{m}]&:[v_1\times m_1]\times[v_2\times m_2]\times\cdots\times[v_M\times m_M]\\&=[m_1]^{v_1}\times[m_2]^{v_2}\times\cdots\times[m_M]^{v_M}\\&=[v_1\times m_1+v_2\times m_2+\cdots+v_M\times m_M]\\&=[\boldsymbol{v}\cdot\boldsymbol{m}]\end{aligned} \tag{10}$$

3. 基于 Paillier 加密的 KH-KNN 室内定位算法

3.1. 结合 Paillier 加密的 KH-KNN 定位

基于 KH 距离公式(5)可得到三部分数据，分别表示为 $S1=\sum_{j=1}^{M} f_{t,j}^2$，$S2=\sum_{j=1}^{M}\left(f_{t,j}\cdot f_{i,j}\right)$，$S3=\sum_{j=1}^{M} f_{i,j}^2$。此时 KH 距离公式可表示为公式(11)：

$$d_{KH}\left(F_t,F_i\right)=\frac{\sum_{j=1}^{M}\left(f_{t,j}\cdot f_{i,j}\right)}{\sum_{j=1}^{M} f_{t,j}^2+\sum_{j=1}^{M} f_{i,j}^2-\sum_{j=1}^{M}\left(f_{t,j}\cdot f_{i,j}\right)}=\frac{S2}{S1+S3-S2} \tag{11}$$

式中 $S2$ 的计算涉及离线指纹 F_i 与用户定位指纹 F_t，而计算 $S1$ 和 $S3$ 都只涉及一方的指纹数据的运算。本方案中，用户利用 Paillier 算法发送加密后的定位指纹，由服务器完成密文域 $S2$ 的计算并返回给用户。根据 Paillier 同态特性有：

$$[S2]=\left[\sum_{j=1}^{M}\left(f_{t,j}\cdot f_{i,j}\right)\right]=\prod_{j=1}^{M}\left[\left(f_{t,j}\right)\right]^{f_{i,j}} \tag{12}$$

由式(12)可知，用户发送 M 个加密后的指纹强度后，服务器对这些数据与自身指纹强度数据进行模幂运算，最终得到密文[$S2$]。

服务商在离线阶段完成 $S3$ 的计算，在线定位阶段把[$S2$]及 $S3$ 数据发送给用户，由用户解密 KH 距离，由此确定 K 个最近邻。

3.2. 方案描述

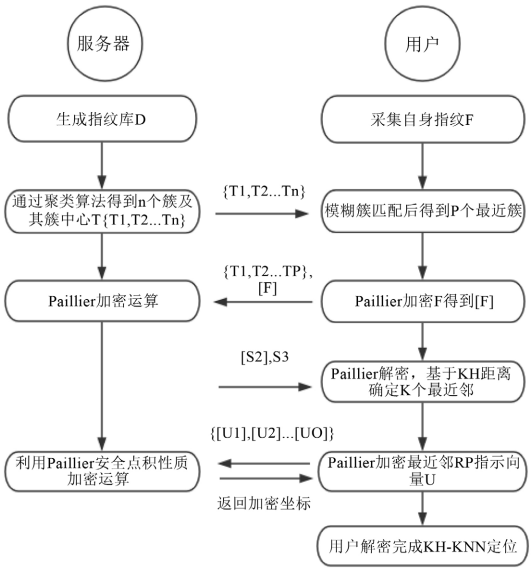

Figure 1. Overall scheme diagram
图 1. 方案整体框图

本文采用 KH 距离方式得到定位用户指纹数据与指纹库中指纹数据的相似度，进而确定 K 个最近邻 RP。首先利用 Paillier 算法完成对用户指纹数据的加密，保护用户位置请求信息，同时服务器利用其同态特性完成 KH 距离中 $S2$ 部分的密文计算。在此之前，本文引入了分簇聚类思想来降低定位过程的时间开

销，同时利用模糊簇匹配模糊服务器端对用户所在真实的簇的判断，完成对连续请求定位用户位置隐私安全的保护。其次，在用户得到最近邻的索引后，考虑到保护服务商坐标数据的隐私安全，利用 Paillier 加密算法的安全点积性质来完成，最终用户得到自身估计位置，完成定位。方案整体框图如图 1 所示。

3.2.1. 分簇及模糊簇匹配

离线阶段，服务器端使用聚类算法和指纹库中的坐标数据将整个定位区域分成若干个小区域，本文把小区域定义为簇。每个簇都有一个簇中心 T_i，簇中心的指纹强度数据可由该簇中所有 RP 的指纹数据强度平均得到，表示为 $f_T=\left(f_{T,1},f_{T,2},\cdots,f_{T,M}\right)$。最后可得簇中心集合 $T=\left\{T_1,T_2,\cdots,T_n\right\}$。

在线阶段，用户在申请定位服务时，首先从服务器获取各簇中心的指纹强度 $f_T=\left(f_{T,1},f_{T,2},\cdots,f_{T,M}\right)$，用户将自身指纹强度 $F_t=\left(f_{t,1},f_{t,2},\cdots,f_{t,M}\right)$ 与各个簇中心的指纹强度作比较，选出最匹配的簇。文中采用指纹强度的欧式距离大小来完成簇匹配。计算过程如下：

$$d=\sqrt{\sum_{i=1}^{M}\left(f_T-f_t\right)^2} \tag{13}$$

通常用户会位于最匹配簇内的 RP 所在区域内，对于连续请求定位的用户，服务器在不知道用户确切位置的情况下，也可能由最匹配簇推测用户轨迹。为保护用户的定位隐私，尤其当用户发起连续定位请求时，服务商可能通过多次定位结果的最匹配簇来推断出用户的运动轨迹，本文采用匹配 P 个最近簇来混淆服务器对用户所在真实簇的判断，同时通过模糊簇匹配方式，保证用户在该 P 个簇内，而不是总在最匹配的簇中。

具体方案如下，在簇匹配时采用部分 AP 信号，也即部分指纹信息 $F_t'=\left(f_{t,1},f_{t,2},\cdots,f_{t,L},L<M\right)$ 来完成簇匹配过程。L 个 AP 信号是由用户在采集指纹后，为全部采集到 AP 信号的及若干个未采集到 AP 信号的指纹强度组成，即 AP 信号，计算过程如下：

$$d'=\sqrt{\sum_{i=1}^{L}\left(f_T-f_t\right)^2} \tag{14}$$

这样使得请求定位用户的真实位置落在 P 个最近簇内的概率是接近的，即保护了用户的定位隐私。

3.2.2. 基于 KH 距离的 K 近邻获取

根据 3.1 所提的定位过程，可知用户需对自身指纹数据进行加密，以便服务器完成[$S2$]的计算。此后用户端将 P 个最近簇的簇号、公钥、加密指纹[F]上传至服务器端。

根据公式(5)，当服务器端收到用户端发送的数据以后，通过 Paillier 加密算法开始计算 $S2$ 部分，计算过程如下：

$$\begin{aligned}[S2]&=\left[\sum_{j=1}^{M}\left(f_{t,j}\cdot f_{i,j}\right)\right]=\left[f_{t,1}\cdot f_{i,1}+f_{t,2}\cdot f_{i,2}+\cdots+f_{t,M}\cdot f_{i,M}\right]\\&=\left[f_{t,1}\cdot f_{i,1}\right]\cdot\left[f_{t,2}\cdot f_{i,2}\right]\cdots\left[f_{t,M}\cdot f_{i,M}\right]\\&=\left[f_{t,1}\right]^{f_{i,1}}\cdot\left[f_{t,2}\right]^{f_{i,2}}\cdots\left[f_{t,M}\right]^{f_{i,M}}\\&=[S2]\end{aligned} \tag{15}$$

根据 KH 距离公式，d 的计算涉及 $S1$，$S2$ 和 $S3$ 部分。为保证定位过程的实时性，服务器可在离线阶段就完成对 $S3=\sum_{j=1}^{M}f_{i,j}^2$ 的计算，并在用户请求定位时发送给用户，服务器端后将密文[$S2$]数据及 $S3$ 数据

发送至用户端。

用户端利用私钥对密文[S2]进行解密，依据公式(11)进行 KH 距离的计算，由此得到 K 个最近的 KH 距离及对应 RP 点的索引。

3.2.3. 利用安全点积完成位置估计

为避免服务商参考点位置信息泄露，同时要避免服务器获取用户位置有关的 K 近邻信息，本方案利用 Paillier 算法的安全点积性质完成自身的位置估计。位置估算过程具体如下：

1) 用户生成最近邻 RP 的指示向量 $\boldsymbol{U}=(U_1,U_2,\cdots,U_O)=(0,0,\cdots,0)$，$\boldsymbol{U}$ 的维数等于参与运算 RP 点的个数，假设 RP 点的个数为 O，$\boldsymbol{U}$ 各元素的初始值为 0，当该 RP 点为最近邻时，对应的初始值变更为 1。

2) 用户利用 Paillier 加密算法加密指示向量 $\boldsymbol{U}$ 的每个元素，共得到 O 个密文，即 $[\boldsymbol{U}]=([U_1],[U_2],\cdots,[U_O])$，后将该密文数据全部上传至服务器端。

3) 服务商利用 Paillier 加密算法安全点积性质计算 $[X^*]=X\odot[U],[Y^*]=Y\odot[U]$，其中 $\boldsymbol{X}=(x_1,x_2,\cdots,x_O)$，$\boldsymbol{Y}=(y_1,y_2,\cdots,y_O)$，$\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 坐标向量包含参与运算的 RP 点所有的 x 和 y 坐标。$[X^*]$ 是服务商给用户发送加密后的用户估计 x 坐标，$[Y^*]$是服务商给用户发送加密后的用户估计 y 坐标。假设最近邻的索引是 1, 3, 4，那么 $[X^*]=[x_1+x_3+x_4],[Y^*]=[y_1+y_3+y_4]$，$x$ 坐标推导过程见公式(16)，y 坐标同理，此处不再赘述。

$$\begin{aligned}[X^*]&=X\odot[U]=(x_1,x_2,\cdots,x_O)\odot[U]\\&=[x_1\cdot U_1]\cdot[x_2\cdot U_2]\cdots[x_O\cdot U_O]=[U_1]^{x_1}[U_2]^{x_2}\cdots[U_O]^{x_O}\\&=[U_1\cdot x_1+U_2\cdot x_2+\cdots+U_O\cdot x_O]\\&=[x_1+x_3+x_4]\end{aligned}\tag{16}$$

计算完成后服务商将[X^*]、[Y^*]发送至用户。

4) 用户收到数据后利用私钥解密，完成自身估计位置的定位，公式如下：

$$(\dot{x},\dot{y})=\frac{1}{K}(X^*,Y^*)\tag{17}$$

3.3. 安全性分析

本文设定用户和服务商都遵循半诚实，即诚实但好奇(curious-but-honest)的安全模型，在该模型中用户和服务商都诚实的遵守规定的协议，但都可能泄露对方的私人信息。在此模型下，本文利用公共数据集[20]中 1 楼的数据，分为 661 个 RP，50 个测试点，其中每个 RP 和测试点都包含 32 个 AP。本文将从用户位置信息的安全性及服务商数据的安全性两个方面展开分析。

3.3.1. 用户位置信息的安全性

a. 分簇及模糊簇匹配

本节安全分析着重分析模糊簇匹配时使用的 AP 个数 L 是否能让服务器端对用户所在真实的簇的判断起到混淆作用，模糊簇匹配时使用的 AP 个数 L 指的是在单个指纹中检测到所有有 AP 信号的和随机抽取出来若干个无 AP 信号的 AP 个数总和。当用户返回给服务器端最近簇的簇号信息后，匹配到的最近簇和用户所在的真实簇不是同一个簇时，则起到了混淆作用。如图 2 所示，当参与模糊簇匹配的 AP 个数 L 取 15 个时，实验所用到的 50 个测试点的定位结果落在三个簇的概率是接近的，此时的混淆作用最好，起到了保护连续请求定位用户隐私的作用。

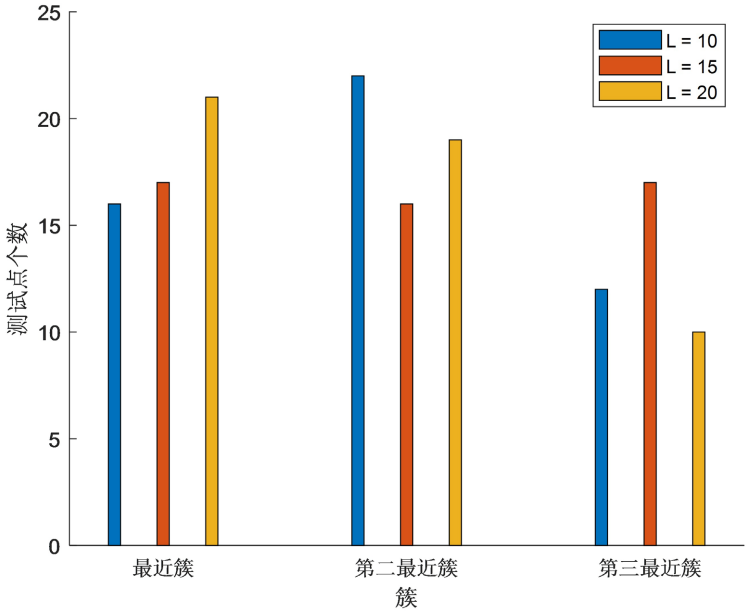

Figure 2. The distribution of test points under different APs *L*
图 2. 在不同的 AP 个数 *L* 下测试点的分布情况图

b. 基于 KH 距离的 *K* 近邻获取阶段

首先用户发给服务商的指纹数据是用户利用 Paillier 加密算法加密自身指纹后的数据，这步保护了用户指纹的隐私安全，其次明文状态下的 *S*1 数据只在用户端计算，并不发送至服务器端，即使发送给服务器端，服务器端也不能通过一个数据才推测出用户全部的指纹数据。

c. 利用安全点积保护服务商坐标信息阶段

用户通过最近簇内的参考点个数生成 RP 指示向量 $\boldsymbol{U}$，在利用 Paillier 加密算法加密后向服务器上传的是若干个密文数据，故服务器并不知道用户最近邻的索引，这一步保护了请求定位用户的隐私安全。

3.3.2. 服务商位置信息的安全性

a. 基于 KH 距离的 *K* 近邻获取阶段

首先计算加密后的 *S*2 数据这一过程是在服务器端进行的，服务商在收到用户发来的加密指纹数据后，进行 *S*2 数据的计算，并不会将自身指纹库的指纹数据发送给用户，这一步保护了服务商的指纹库的指纹信息隐私安全。其次，*S*3 数据也是服务商自身计算的，之后发送给用户也只是一个数据，同理用户并不能根据这一个数据推测出服务商一条指纹全部的指纹数据。在计算完成后，服务商发送给用户加密后的 *S*2 数据。

b. 利用安全点积保护服务商坐标信息阶段

服务商在收到用户发送的若干个密文数据以后，将生成 $\boldsymbol{X}$ 坐标向量及 $\boldsymbol{Y}$ 坐标向量，利用 Paillier 的安全点积性质进行 *X* 坐标向量及 *Y* 坐标向量与若干个密文数据的计算，计算完成后，得到最近邻的 *x* 坐标之和的密文数据及最近邻的 *y* 坐标之和的密文数据。用户解密该密文数据得到自身估计位置，服务商在不告诉用户最近邻的坐标信息的前提下，完成了定位，这一步有效保护了服务商坐标信息不被泄露。

4. 方案性能分析

4.1. 理论分析

为论证本文所提方案性能，现从以下两个方面进行分析：

a. 获取最近邻的时间开销分析

M 代表 AP 个数，N 代表 RP 个数，N'代表最近簇 P 内的参考点个数，k 代表 k 组指纹，T_E 表示加密一个数据计算开销，$T_{E,M}$ 表示一次模幂及模乘运算的时间开销，T_M 表示一次模乘运算的时间开销，T_D 表示解密一个数据计算开销。虽然在 KH 距离计算中涉及明文计算，但这部分计算开销相比较加解密运算很小，可以忽略不计，故本表格不做分析。同时每个密文长度用 L_E 表示。由于明文与密文相差较大，本表格只对密文分析。

表 1 对几种算法的计算开销和通信进行了理论对比。其中无粗定位 KH-KNN 算法表示不进行模糊簇匹配，在整个指纹库内进行 K 近邻搜索。

Table 1. Comparison of the KH-KNN algorithm with the time cost theory of various algorithms
表 1. KH-KNN 算法与各类算法时间开销理论对比

	用户端计算开销	服务商计算开销	通信开销
文献[13]方案	$MT_E + NT_D$	$MN(T_E + T_{E,M}) + NT_E$	$ML_E + MNL_E$
文献[14]方案	kMT_E	$kMNT_M + 2kNT_D$	kML_E
无粗定位 KH-KNN	$MT_E + NT_D$	$MNT_{E,M}$	$ML_E + MNL_E$
KH-KNN	$MT_E + N'T_D$	$MN'T_{E,M}$	$ML_E + MN'L_E$

由表 1 可知，本文所提方案与两篇文献所提方案同样采用了 Paillier 加密算法来保护定位过程中双方的隐私信息，但是本文方案具有更高的定位效率。同时文献[13]算法在定位过程中，服务商需加密全部参与运算的 RP 及得到的欧氏距离，这极大地增加了时间开销。在文献[14]中，用户需要加密 k 组指纹，服务商需要解密 k 组数据才能得到最后的 k 组估计位置，如此大的计算开销带来的只是 k 匿名的安全强度。在无粗定位的情况下，本文方案并不需要加密全部参与运算的 RP 点且不需要加密 KH 距离，加入粗定位后，时间开销进一步降低。

b. 保护服务商坐标信息的时间开销分析

为实现对服务商坐标信息的保护，用户在完成 K 近邻检索后，还需发送加密后的最近邻 RP 指示向量 U 给服务商，这会引入一定的计算和通信开销。表 2 给出了无粗定位 KH-KNN 算法与 KH-KNN 算法在这一阶段的开销。

Table 2. Theoretical comparison of time overhead between the KH-KNN algorithm without coarse positioning and the KH-KNN algorithm
表 2. 无粗定位 KH-KNN 算法与 KH-KNN 算法时间开销理论对比

	用户端计算开销	服务商计算开销	通信开销
无粗定位 KH-KNN	$NT_E + 2T_D$	$2NT_{E,M}$	$NL_E + 2L_E$
KH-KNN	$N'T_E + 2T_D$	$2N'T_{E,M}$	$N'L_E + 2L_E$

由表 2 可知，在加入粗定位后，无论是用户端计算开销还是服务器端计算开销，亦或通信开销，都有明显的降低效果，同时本文利用 Paillier 算法的安全点积性质巧妙地保护了服务商坐标信息，文献[13][14]所提方案并未对服务商坐标信息进行保护，故本表格不作分析。

4.2. 实验验证

为验证本文所提方案在实际室内定位的可靠性及定位性能，采用公共数据集[20]中一楼的数据进行

相关实验，该公共数据集所搭建的实验平台大小约 30 m*50 m，数据集中的数据经处理后，实验环境布局如图 3 所示，图中黑色点代表参考点，共计 661 个，红色点代表测试点，共计 50 个，在整个实验环境中共能采集到 32 个 AP 信号，当测试点未能采集到 AP 信号时，该位置的 RSSI 值设为 0。

Paillier 算法采用了基于 MATLAB 平台工具箱实现加解密运算[21]，定位算法同样基于 MATLAB 平台。所运行的 PC 环境如下：DELL G7、Intel (R) Core (TM) i7-8750H CPU @ 2.20 GHz-2.21 GHz、16 GB RAM、64 bit 操作系统。

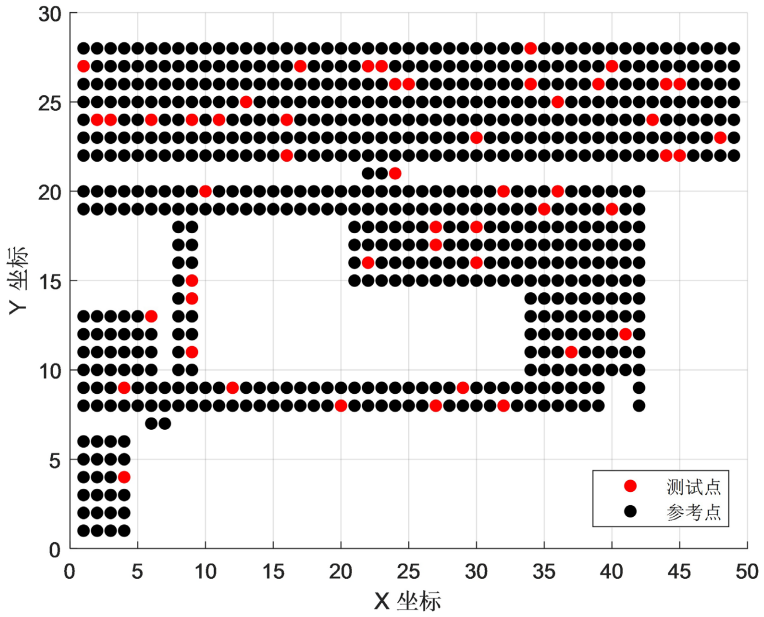

Figure 3. Experimental environment layout diagram
图 3. 实验环境布局图

4.2.1. 时间开销分析

为确保实验结果可靠性，在本实验条件下进行多次实验后取平均结果，单个用户完成一次定位的时间开销如表 3 所示：

Table 3. Comparison of the time cost of the KH-KNN algorithm with the algorithm in Ref [13]
表 3. KH-KNN 算法与文献[13]算法时间开销对比

	用户端计算开销(s)	服务商计算开销(s)	时间开销(s)
文献[13]方案	3.0	10.4	13.2
无粗定位 KH-KNN	5.9	11.0	17.1
KH-KNN	1.5	2.4	3.9

由上表可知：由于加入对服务商参考点坐标信息保护的计算，无粗定位的 KH-KNN 算法在用户端计算开销显著提高，略微提高服务端计算开销，导致整体定位时间开销增大。加入粗定位以后，很明显用户端和服务器端的时间开销明显降低，提高了单次定位的实时性。由于文献[14]是基于三方的定位方案，故本表格不作分析。

4.2.2. 定位精度分析

a. 粗定位对精度的影响分析

所提方案中，通过模糊簇匹配把用户定位到局部区域，因此匹配的最近簇个数 P 直接决定了定位精度及定位的时间开销。结合簇模糊匹配的定位误差如下图 4 所示，当匹配的最近簇个数增加时，定位精度有提升趋势，但不明显；定位时间开销如图 5 所示，当匹配的最近簇个数增加时，局部区域包含更多的 RP，使得服务器需要计算更多的[*S*2]，同时用户需解密的[*S*2]数量也增加，这使得定位时间显著增加。故匹配的最近簇个数最终取为 3。

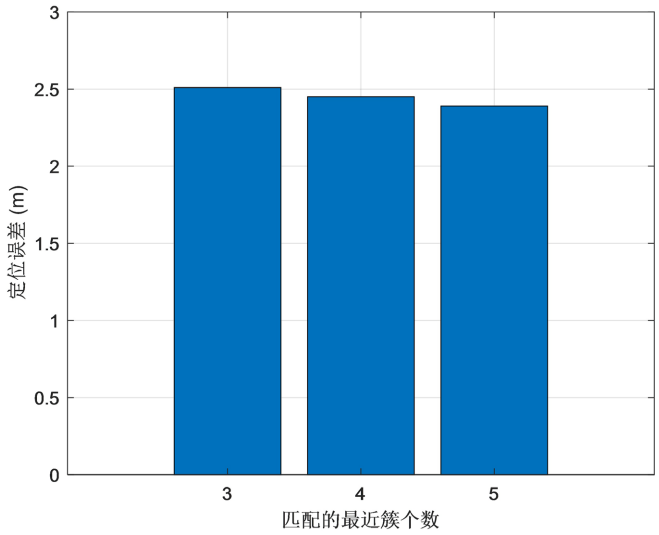

Figure 4. The number of nearest clusters affects the positioning error
图 4. 最近簇的个数对定位误差影响图

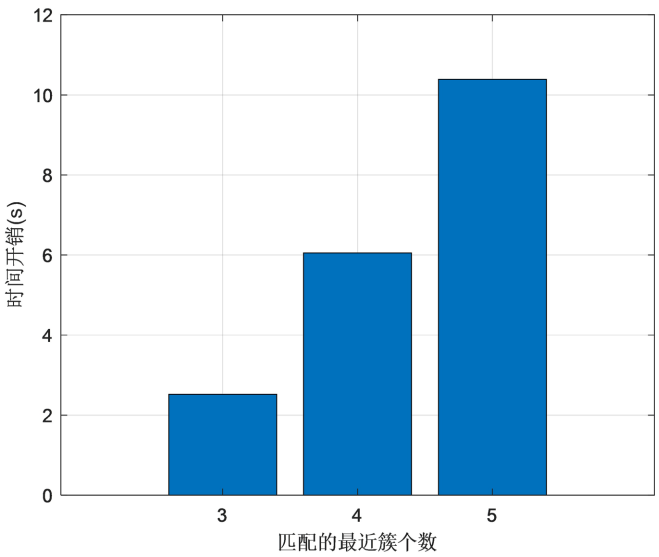

Figure 5. The impact of the number of recent clusters on time overheamd
图 5. 最近簇的个数对时间开销影响图

b. 与不同定位算法作比较

本文方案中，采用 KH 距离大小选择 K 个最近邻，完成用户位置估计。目前已有的隐私保护定位算法[13]采用了欧式距离，其中也有文献[14]采用了 Sφrensen 距离。图 6 对基于 3 种不同距离的 KNN 近邻算法的定位性能进行了比较，其中欧式距离的平均误差为 2.71 m，Sφrensen 距离的平均误差为 2.52 m，

KH 距离的平均误差为 2.45 m。由图可知，无论是否加入粗定位，本文所提算法在定位性能上都要稍优于其余两种算法，考虑到降低定位时的时间开销，故加入粗定位来提高实时性。

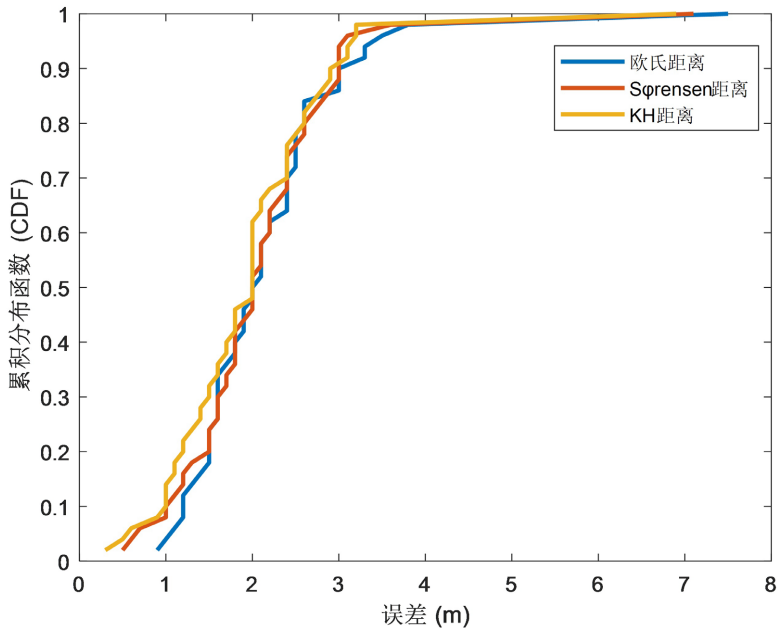

Figure 6. CDF diagram based on different distance matching and positioning algorithms
图 6. 基于不同距离匹配定位算法的 CDF 图

5. 结束语

针对室内定位服务中，用户位置隐私和服务器数据资源隐私的安全问题，本文基于 KH 距离和 Paillier 算法提出一种兼顾定位实时性、定位精度和隐私安全的室内定位方案。方案采取 KH 距离匹配定位用户的指纹数据与指纹库中参考点的指纹数据完成对最近邻的判定。同时方案将 Paillier 算法与 KH 距离计算结合，实现用户和服务商的指纹数据安全，同时，利用此算法的安全点积性质实现了对服务商的指纹库坐标数据安全保护。为了降低时间开销，方案引入了分簇聚类思想，利用模糊簇匹配混淆服务器端对用户真实位置的判断，在提高定位实时性的同时实现用户位置的隐私安全。研究结果表明了所提方案的安全性、实时性和定位性能。

参考文献

[1] Li, S., Hedley, M., Bengston, K., Humphrey, D., Johnson, M. and Ni, W. (2019) Passive Localization of Standard Wifi Devices. *IEEE Systems Journal*, **13**, 3929-3932. https://doi.org/10.1109/jsyst.2019.2903278

[2] Yang, X., Wu, Z. and Zhang, Q. (2022) Bluetooth Indoor Localization with Gaussian-Bernoulli Restricted Boltzmann Machine Plus Liquid State Machine. *IEEE Transactions on Instrumentation and Measurement*, **71**, 1-8. https://doi.org/10.1109/tim.2021.3135344

[3] Ibnatta, Y., Khaldoun, M. and Sadik, M. (2022) Indoor Localization System Based on Mobile Access Point Model MAPM Using RSS with UWB-OFDM. *IEEE Access*, **10**, 46043-46056. https://doi.org/10.1109/access.2022.3168677

[4] Lee, S., Kim, J. and Moon, N. (2019) Random Forest and Wifi Fingerprint-Based Indoor Location Recognition System Using Smart Watch. *Human-Centric Computing and Information Sciences*, **9**, Article No. 6. https://doi.org/10.1186/s13673-019-0168-7

[5] Leitch, S.G., Ahmed, Q.Z., Abbas, W.B., Hafeez, M., Laziridis, P.I., Sureephong, P., *et al*. (2023) On Indoor Localization Using Wifi, BLE, UWB, and IMU Technologies. *Sensors*, **23**, Article 8598. https://doi.org/10.3390/s23208598

[6] 王慧强, 高凯旋, 吕宏武. 高精度室内定位研究评述及未来演进展望[J]. 通信学报, 2021, 42(7): 198-210.

[7] Jiang, H., Zhao, P. and Wang, C. (2018) Roblop: Towards Robust Privacy Preserving against Location Dependent Attacks in Continuous LBS Queries. *IEEE/ACM Transactions on Networking*, **26**, 1018-1032. https://doi.org/10.1109/tnet.2018.2812851

[8] Zhai, F., Liang, X., Qin, Y., Li, B., Shen, L. and Xie, J. (2024) Privacy-Preserving Method for Sensitive Partitions of Electricity Consumption Data Based on Hybrid Differential Privacy and K-Anonymity. *Journal of Physics*: *Conference Series*, **2806**, Article 012010. https://doi.org/10.1088/1742-6596/2806/1/012010

[9] Yazdanjue, N., Yazdanjouei, H., Karimianghadim, R. and Gandomi, A.H. (2024) An Enhanced Discrete Particle Swarm Optimization for Structural K-Anonymity in Social Networks. *Information Sciences*, **670**, Article 120631. https://doi.org/10.1016/j.ins.2024.120631

[10] 张志武, 雷若兰, 乐燕芬. 移动对象室内定位中的隐私保护方案[J]. 数据采集与处理, 2024, 39(3): 761-774.

[11] Kiarashi, Y., Saghafi, S., Das, B., Hegde, C., Madala, V.S.K., Nakum, A., *et al*. (2023) Graph Trilateration for Indoor Localization in Sparsely Distributed Edge Computing Devices in Complex Environments Using Bluetooth Technology. *Sensors*, **23**, Article 9517. https://doi.org/10.3390/s23239517

[12] Xie, S., Yu, X., Guo, Z., Zhu, M. and Han, Y. (2023) Multi-Output Regression Indoor Localization Algorithm Based on Hybrid Grey Wolf Particle Swarm Optimization. *Applied Sciences*, **13**, Article 12167. https://doi.org/10.3390/app132212167

[13] Li, H., Sun, L., Zhu, H., Lu, X. and Cheng, X. (2014). Achieving Privacy Preservation in Wifi Fingerprint-Based Localization. *IEEE INFOCOM* 2014—*IEEE Conference on Computer Communications*, Toronto, 27 April-2 May 2014, 2337-2445. https://doi.org/10.1109/infocom.2014.6848178

[14] 张应辉, 张思睿, 赵秋霞, 等. 基于 Wi-Fi 指纹且计算外包的室内定位隐私保护方案[J]. 通信学报, 2024, 45(2): 31-39.

[15] Eshun, S.N. and Palmieri, P. (2024) A Cryptographic Protocol for Efficient Mutual Location Privacy through Outsourcing in Indoor Wifi Localization. *IEEE Transactions on Information Forensics and Security*, **19**, 4086-4099. https://doi.org/10.1109/tifs.2024.3372805

[16] Jarvinen, K., Leppakoski, H., Lohan, E., Richter, P., Schneider, T., Tkachenko, O., *et al*. (2019) PILOT: Practical Privacy-Preserving Indoor Localization Using Outsourcing. 2019 *IEEE European Symposium on Security and Privacy* (*EuroS&P*), Stockholm, 17-19 June 2019, 448-463. https://doi.org/10.1109/eurosp.2019.00040

[17] Bundak, C.E.A., Abd Rahman, M.A., Abdul Karim, M.K. and Osman, N.H. (2022) Fuzzy Rank Cluster Top K Euclidean Distance and Triangle Based Algorithm for Magnetic Field Indoor Positioning System. *Alexandria Engineering Journal*, **61**, 3645-3655. https://doi.org/10.1016/j.aej.2021.08.073

[18] Chen, J., Song, S., Gu, Y. and Zhang, S. (2022) A Multisensor Fusion Algorithm of Indoor Localization Using Derivative Euclidean Distance and the Weighted Extended Kalman Filter. *Sensor Review*, **42**, 669-681. https://doi.org/10.1108/sr-10-2021-0337

[19] Damgård, I. and Jurik, M. (2001) A Generalisation, a Simplification and Some Applications of Paillier's Probabilistic Public-Key System. In: *Lecture Notes in Computer Science*, Springer, 119-136. https://doi.org/10.1007/3-540-44586-2_9

[20] Toth, Z. and Tamas, J. (2016) Miskolc IIS Hybrid IPS: Dataset for Hybrid Indoor Positioning. 2016 26*th International Conference Radioelektronika* (*RADIOELEKTRONIKA*), Kosice, 19-20 April 2016, 408-412. https://doi.org/10.1109/radioelek.2016.7477348

[21] GitHub (2022) MATLAB Class-Based Toolbox for Paillier Crypto System. https://github.com/martin-kaluz/PaillierCrypto-matlab