ISSN 1004-9037,CODEN SCYCE4
Journal of Data Acquisition and Processing Vol. 39,No. 2,Mar. 2024,pp. 456－470
DOI:10. 16337/j. 1004-9037. 2024. 02. 018

http://sjcj. nuaa. edu. cn
E-mail:sjcj@nuaa. edu. cn
Tel/Fax:＋86-025-84892742

基于密文KNN检索的室内定位隐私保护算法

欧锦添，乐燕芬，施伟斌

（上海理工大学光电信息与计算机工程学院，上海 200093）

摘　要：在定位请求服务中，如何保护用户的位置隐私和位置服务提供商（Localization service provider，LSP）的数据隐私是关系到WiFi指纹定位应用的一个具有挑战性的问题。基于密文域的K-近邻（K-nearest neighbors，KNN）检索，本文提出了一种适用于三方的定位隐私保护算法，能有效提升对LSP指纹信息隐私的保护强度并降低计算开销。服务器和用户分别完成对指纹信息和定位请求的加密，而第三方则基于加密指纹库和加密定位请求，在隐私状态下完成对用户的位置估计。所提算法把各参考点的位置信息随机嵌入指纹，可避免恶意用户获取各参考点的具体位置；进一步利用布隆滤波器在隐藏接入点信息的情况下，第三方可完成参考点的在线匹配，实现对用户隐私状态下的粗定位，可与定位算法结合降低计算开销。在公共数据集和实验室数据集中，对两种算法的安全、开销和定位性能进行了全面的评估。与同类加密算法比较，在不降低定位精度的情况下，进一步增强了对数据隐私的保护。

关键词：隐私保护；指纹定位；密文K-近邻检索；布隆滤波器；WiFi
中图分类号：TP391　　**文献标志码**：A

Indoor Location Privacy Protection Algorithm Based on Ciphertext KNN Retrieval

OU Jintian，LE Yanfen，SHI Weibin

（School of Optical-Electrical and Computer Engineering，University of Shanghai for Science and Technology，Shanghai 200093，China）

Abstract：In the location request service，how to protect the user's location privacy and the data privacy of the location service provider（LSP）is a challenging issue related to WiFi fingerprinting applications. Based on the K-nearest neighbors（KNN）retrieval of the ciphertext，this paper proposes a positioning privacy protection algorithm suitable for the three party，which can effectively improve the protection intensity of the privacy of LSP fingerprint information and reduce calculation overhead. The positioning algorithm is completed by a third party based on the encrypted fingerprint database and encrypted positioning request，which is completed in the state of privacy. Through the random embedding of the location information in the fingerprint，the algorithm can avoid the physical location of the reference point（RP）in the fingerprint database. The Bloom filter（BF）is further used to complete the online matching of the reference point when hiding the access point information，which achieves rough positioning in the privacy of the user，and reduces the calculation overhead with the positioning algorithm. In the data set of public datasets and laboratory data，the security，expense and positioning performance of the two algorithms have been

基金项目：国家自然科学基金（51705324）。
收稿日期：2023-05-18；**修订日期**：2024-01-16

comprehensively evaluated. Compared with similar encryption algorithms, without reducing positioning accuracy, it further enhances the protection of data privacy.

Key words: privacy protection; fingerprinting localization; ciphertext K-nearest neighbors (KNN) retrieval; Bloom filter; WiFi

引　言

近年来,基于蓝牙[1]、超宽带(Ultra wide band, UWB)[2]、WiFi[3-5]信号的室内定位技术受到广泛关注和研究。因WiFi部署成本低、便捷性好及信号易获得等优势,基于WiFi无线信号的指纹定位成为其中最受关注的技术之一[6-8],并形成了一系列相应的商业应用,如Google、Skyhook Wireless、WifiSLAM和Wifarer等厂商提供的基于室内WiFi的定位和导航服务。

WiFi的指纹定位技术通常包含离线和在线两个阶段。在离线阶段,服务提供商在定位区域中按照一定的布局设定一系列的参考点,并测量对应位置上的WiFi接入点(Access point, AP)的接收信号强度(Received signal strength, RSS)。经过多次测量和采取均值等处理方式,将位置数据和与其对应的信号值存入指纹库中,形成离线的无线地图。在在线阶段,用户端扫描获得各个AP的RSS信号,并采用K近邻(K-nearest neighbors,KNN)算法,找出指纹库与其RSS信号值最接近的K个指纹,并利用对应参考点(Reference point,RP)的物理位置来估计用户自身的位置。

新兴的计算模式,如数据库服务外包和云计算等方式,给企业将其数据库管理系统等计算业务迁移到服务提供商提供了可能。但用户数据若泄漏给未经授权的一方,就会导致隐私泄漏等问题[9]。在室内定位中,若该项服务由第三方提供,则通过对定位数据的挖掘,与位置服务相关的用户的个人数据,如用户的日常习惯、社交网络、个人兴趣和宗教信仰等可能被追踪并加以利用,最终可能造成用户的经济损失和名誉受损。而服务提供商构建数据库不仅需耗费大量的人力和时间成本,同时数据库的泄露使得更容易对用户定位信息进行攻击。所以,对用户和服务器提供商的数据同时提供隐私保护是安全定位系统的关键所在。

目前具有用户位置隐私保护的算法,主要有各类空间K匿名算法[10-11]、混合区技术[12-13]和加密算法[14-16]。这类算法旨在保护用户的位置隐私和请求隐私,避免特定用户被识别、或避免由请求内容泄露用户其他的隐私。但是,K-匿名算法容易受到同性质的攻击;混合技术中用户在混合区内无法正常获取LBS服务;而Paillier等加密算法具有良好的安全性,但为保证安全强度,算法的计算开销较大,秒级的时间开销限制了其在某些实时性要求高的目标追踪等中的应用。

也有学者将基于保距变换(Distance-preserving transformation,DPT)的数据加密用于密文域的KNN查找[17],但这种转换安全性不强。如果攻击者可以访问DPT加密的数据库(Ecrypted database, E(DB)),掌握了部分数据的明文和相应的密文,就可以完全恢复DB[18],也即该加密方式不能承受已知明文攻击(Known plaintext attack, KPA)。文献[19]提出了一种非对称标量积(Asymmetric scalar-product-preserving encryption, ASPE)的加密定位算法,第三方攻击者无法恢复数据库中的指纹信息,但是用户解密可获得DB,泄露了指纹信息。

为保护服务器指纹数据库和用户的个人数据隐私,基于文献[19],本文提出了一种融合ASPE并结合匿名分簇的密文KNN定位算法,达到既保护用户的请求隐私和位置隐私,又有效保护服务提供商的指纹数据不受恶意窃取。考虑到在实际应用中经常需要将用户粗定位在一个局部区域内[20-21],算法引入了布隆滤波器(Bloom filter,BF)完成基于匿名AP序列的参考点匹配,实现隐私状态下对用户的局部定位。

1 研究背景

1.1 WiFi指纹定位算法

在目前众多的室内定位技术方案中,基于WiFi的指纹定位算法是广泛研究的技术之一。这类算法的定位过程,分为离线阶段和在线阶段。在离线阶段中,服务器端在目标区域进行数据的采集,完成指纹库的构建$\boldsymbol{p}_i=(p_{i1},p_{i2},\cdots,p_{id})$。采集的数据包括参考点RP的位置坐标$l_i=(x_i,y_i)$,以及在该点收集到的RSS信号$p_{ij}$,该值表示在$\mathrm{RP}_i$位置上接收到的来自$\mathrm{AP}_j$的指纹信号,$d$则表示为整个定位区域AP的数量。在在线阶段,用户端扫描可获得的RSS信号值,然后将RSS数据发送至服务器端。服务器端使用定位算法,基于用户发送的RSS信号计算定位结果,并将结果发回用户端。

针对定位精度要求不同,用户端可以采用不同的定位算法。在对定位要求比较高的场合,服务器会采用分类算法、神经网络和基于核的回归算法等机器学习算法[22-24]。这些算法利用指纹库中的位置数据和RSS信号值进行离线模型训练,得到两类数据之间的映射关系。但是,这些算法较为复杂且耗费的模型训练资源较多。相比之下,KNN算法是一种综合定位性能较优的算法。如式(1)所示,KNN算法通过计算出用户指纹与服务器数据库中指纹之间的欧式距离,找到最近的K个指纹,并通过这K个最近邻指纹的坐标估计目标位置。具体来说,KNN算法的估计位置为

$$d_i=\left\|\boldsymbol{q}-\boldsymbol{p}_i\right\|^2=\sum_{j=1}^{d}(q_j-p_{ij})^2 \tag{1}$$

式中:$\boldsymbol{q}$为用户RSS数据,$\boldsymbol{p}_i$为数据库中第i个指纹数据。

$$l^*=\frac{1}{K}\sum_{i=1}^{K}l_i^* \tag{2}$$

式中:l_i^*为最近邻参考点坐标,l^*为估计位置坐标。

1.2 位置隐私保护技术

现有的实现位置隐私保护技术主要包括K-匿名技术、差分隐私技术、假名技术以及加密技术,如Paillier同态加密、ASPE加密等。

K-匿名技术将用户分为目标用户和$K-1$个其他用户,这K个用户所在的区域为匿名区域,K的大小即为匿名集的大小。通过$K-1$个其他用户的干扰,可以提高隐私保护性能。但用户分布稀疏时,匿名区域就会很大,导致查询结果不准确和开销较大。

差分隐私技术能够在保护用户隐私信息的前提下,确保查询结果对于任意特定用户的个人数据保持低敏感性,并显示数据集的正确统计信息。这种技术通过假设敌手掌握任意背景知识来为数据提供隐私保护,但假设的实际应用场景受到限制。

假名技术主要利用可信任的第三方服务器,为用户创建了一个虚假的身份标识,再向基于位置的服务(Location based service,LBS)服务器发送查询定位请求。这避免了用户真实信息的暴露,使得服务器不能通过用户的身份信息来将查询定位请求信息中的数据与具体的用户身份对应。基于这种技术,用户的敏感信息可以受到有效保护,同时不影响服务质量。但是,在假名技术中的混合区模型[13]下,用户一旦进入混合区,其身份就会与其他用户的身份混合在一起,导致用户的定位服务受到影响并且服务质量明显下降。

Paillier加密技术是加性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,具有加法同态和标量积同态特性,被广泛应用于电子现金交易、安全电子投票和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[25-26]。在定位服务中,通过Paillier加密技术,可以对加密的定位信息进行处理获得计算结果,进而保证定位服务中,用户个人

隐私信息和服务器的数据库受到保护。然而,这种加密方式存在不小的计算和通信开销,特别是当安全强度增加时,密钥长度高达2 048比特甚至更大时,定位服务的延迟会明显增加,导致影响服务质量。

相比Paillier加密算法,ASPE方案的计算量较少;同时克服了DPT算法中加密的DB易被攻击而泄露数据的问题,能够有效保护DB不被攻击者获得。但是,该算法中用户解密可得到指纹数据和最近邻参考点的位置。当用户成为攻击者时,则可轻易获取指纹库数据。

为了实现对服务器端的指纹数据更强的隐私保护,本文提出了一种改进的非对称标量积隐私定位算法。算法首先考虑把RP的物理位置信息随机嵌入到指纹后再进行加密运算,使得用户在完成解密定位后只获取自身的估计位置,而无法得知最近邻的K个参考点的位置和RSS信息,最大程度增强对服务器端指纹数据的保护;同时在隐私定位中引入了匿名粗定位,为不泄露服务器和用户定位时所用的AP信息,利用布隆滤波器(Bloom filter, BF)完成AP序列的匿名匹配来获得用户位置附近的候选RP集。

2 系统模型、威胁模型和设计目标

2.1 系统模型

本文的系统模型如图1所示,由3类实体组成:服务器端、客户端和第三方。

(1) 服务器端:完成系统的初始化,包括生成密钥、对指纹库进行加密处理等,并把加密指纹库发送给第三方;并将定位相关信息发送给用户。

(2) 用户端:通过密钥对自身扫描的定位RSS信息进行加密;发送加密查询请求到第三方;解密从第三方返回的加密信息,完成位置估计。

(3) 第三方:根据从服务器端接收的加密指纹库和从用户端接收的加密查询请求完成密文KNN检索,将查询结果发送到用户端。

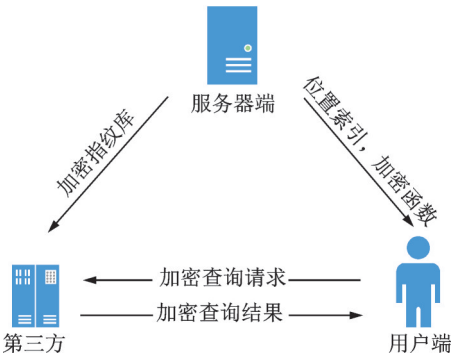

图1 系统模型

Fig.1 System model

2.2 威胁模型

本文的威胁模型定义如下:服务器和用户可信,第三方不可信且不与用户合谋。由此得出以下的攻击模型:

(1) 唯密文攻击(Ciphertext only attack, COA):第三方可以获得加密指纹库和加密查询请求,可以对密文进行穷举攻击。

(2) KPA:恶意用户可以通过查询结果,解密出部分指纹。通过多次恶意定位请求,构建相似指纹库。

2.3 设计目标

本文的目标是设计一种基于ASPE的具有位置隐私保护的定位算法。主要的设计要求包括:

(1) 隐私要求。

① 指纹数据库安全性:服务器端的指纹信息应该被保护,第三方和用户均不能获取指纹信息和参考点位置信息。

② 用户数据安全性:用户的请求指纹和定位结果应该被保护,第三方或恶意攻击者不能获取用户的数据信息。

③ 匹配安全性:粗定位的匿名匹配过程应该不泄露用户的位置信息。

(2) 定位性能。所提具有隐私保护的定位算法应该达到与明文域定位算法相当的定位精度。

(3) 定位效率。在不降低定位精度的前提下,所提具有隐私保护的定位算法应具有良好的实时性。

3 改进的非对称标量积算法

3.1 改进的ASPE算法框架

如图2所示,定位算法包含服务器端、用户端和第三方。服务器端包含有离线指纹库$\boldsymbol{D}$;第三方拥有加密的指纹库$\boldsymbol{P}'$;用户端采集数据并加密后发送定位请求到第三方。所提的具有隐私保护的定位算法包含5个阶段,分别是准备阶段、服务器端加密阶段、用户端生成定位请求、第三方定位处理和用户端解密定位。

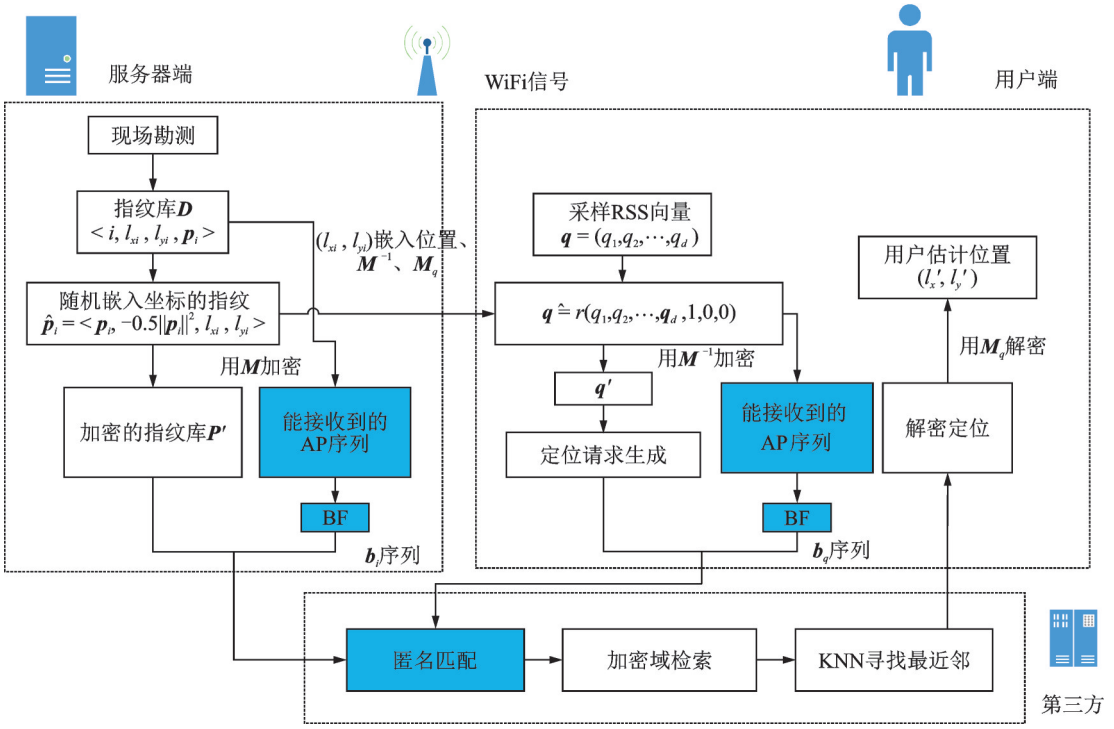

图2 基于密文KNN检索的定位方案系统框图

Fig.2 Diagram of the localization system based on ciphertext KNN retrieval

本文提出两种方案,分别是具有参考点位置隐私保护的基础方案(Basic scheme, BS)和基于BF的改进方案(Improved scheme based on Bloom filter, ISBBF)。每个方案均包含上面5个阶段。图2中加底色标注的部分为ISBBF基于基础方案的改进部分。

图2中,服务器端进行现场勘测获得指纹库$\boldsymbol{D}$,并进行位置嵌入和加密,获得加密指纹库$\boldsymbol{P}'$,再发送到第三方;而用户端进行RSS向量采样,将数据加密后发送到第三方;第三方再根据加密指纹库$\boldsymbol{P}'$和用户的数据$\boldsymbol{q}'$进行加密检索得到KNN最近邻,最后发送到用户端进行解密得到估计位置。图2中白底色框图即为本文提出的基础方案BS。

为了实现隐私保护的同时对用户实现在局部位置的粗定位,在BS的基础上进一步提出具有匿名RP匹配的基于BF的改进方案ISBBF。这个方案相比较BS增加了匿名匹配和BF滤波器映射,如图2中添加蓝底色框图所示。

3.2 基础方案BS

BS中采用位置嵌入方式,拓展位置信息到RSS指纹中,经过安全KNN计算后,第三方返回加密的用户估计位置;可避免用户在获得服务时解密最近邻参考点的位置。5个阶段包含的具体内容如下所示。

(1) 准备阶段:服务器将离线指纹库$\boldsymbol{D}$中d维的RSS指纹$\boldsymbol{p}_i$拓展为$d+3$维的$\hat{\boldsymbol{p}}_i$。其中,第$d+1$维添加$-0.5\|\boldsymbol{p}_i\|^2$,参考点坐标$(l_{xi},l_{yi})$随机嵌入到指纹$\boldsymbol{p}_i$中。假设嵌入的位置在末端,故第$d+2$维和$d+3$维分别添加$l_{xi},l_{yi}$。

$$\hat{\boldsymbol{p}}_i=(\boldsymbol{p}_i,-0.5\|\boldsymbol{p}_i\|^2,l_{xi},l_{yi})^{\mathrm{T}}=(p_{i1},p_{i2},\cdots,p_{id},-0.5\|\boldsymbol{p}_i\|^2,l_{xi},l_{yi})^{\mathrm{T}} \tag{3}$$

服务器生成由$(d+3)\times(d+3)$维的随机数构成的可逆矩阵密钥$\boldsymbol{M}$,并发送(l_{xi},l_{yi})嵌入位置、加密密钥$\boldsymbol{M}^{-1}$和解密密钥$\boldsymbol{M}_q$到用户端。其中,$\boldsymbol{M}_q$由位置提取矩阵$\boldsymbol{D}_K$和解密密钥$(\boldsymbol{M}^{\mathrm{T}})^{-1}$的矩阵乘积得到。位置提取矩阵$\boldsymbol{D}_K$是对角矩阵,根据服务器设定的$(l_{xi},l_{yi})$嵌入位置确定对角线上元素为1的位置。当采用式(3)中的嵌入位置时,提取矩阵$\boldsymbol{D}_K=\mathrm{diag}(0,0,0,\cdots,0,1,1)$,其对角线上只有$d+2$维和$d+3$维为1,其余均为0。

$$\boldsymbol{M}_q=\boldsymbol{D}_K(\boldsymbol{M}^{\mathrm{T}})^{-1} \tag{4}$$

(2) 服务器端加密阶段:服务器用密钥$\boldsymbol{M}$对数据库进行加密,发送加密的地图数据给第三方。

$$\boldsymbol{p}_i'=\boldsymbol{M}^{\mathrm{T}}\hat{\boldsymbol{p}}_i \tag{5}$$

(3) 用户端生成定位请求:用户扩展d维的RSS指纹$\boldsymbol{q}$为$d+3$维的$\hat{\boldsymbol{q}}$,用密钥$\boldsymbol{M}^{-1}$加密$\hat{\boldsymbol{q}}$得到$\boldsymbol{q}'$并发给第三方。其中,$\hat{\boldsymbol{q}}$的第$d+1$维置1。根据接收的(l_{xi},l_{yi})嵌入位置,第$d+2$和$d+3$维均置0。r为随机数,且$r>0$。引入随机数r不会影响最终的检索结果。

$$\hat{\boldsymbol{q}}=r(\boldsymbol{q},1,0,0)^{\mathrm{T}}=r(q_1,q_2,\cdots,q_d,1,0,0)^{\mathrm{T}} \tag{6}$$

$$\boldsymbol{q}'=\boldsymbol{M}^{-1}\hat{\boldsymbol{q}} \tag{7}$$

(4) 第三方定位处理:根据收到的用户加密RSS指纹$\boldsymbol{q}'$,第三方需在服务器发送的加密指纹库$\boldsymbol{P}'=[\boldsymbol{p}_1',\boldsymbol{p}_2',\cdots,\boldsymbol{p}_N']$中进行密文域的距离比较,获得距离用户指纹$\boldsymbol{q}'$最近的$K$个RP的加密指纹$[\boldsymbol{p}_1',\boldsymbol{p}_2',\cdots,\boldsymbol{p}_K']$,取平均值得到$\boldsymbol{p}^*$,并发送到用户端。

由密文域距离计算获得最近邻参考点的理论分析过程如下:可用欧式距离$d(\boldsymbol{p}_i,\boldsymbol{q})$表示数据库中任一参考点$\mathrm{RP}_i$的指纹$\boldsymbol{p}_i$和用户位置请求指纹$\boldsymbol{q}$的距离关系。$d$越小,可认为两者之间的距离越近。

$$d(\boldsymbol{p}_i,\boldsymbol{q})=\|\boldsymbol{p}_i\|^2-2\boldsymbol{p}\cdot\boldsymbol{q}+\|\boldsymbol{q}\|^2 \tag{8}$$

第三方在此过程中,处理的都是密文数据,计算所得的也是加密后包含了用户估计位置的平均指纹信息,未获得任何明文信息。

(5) 用户端解密定位:基于从第三方返回的加密定位数据$\boldsymbol{p}^*$、解密密钥$\boldsymbol{M}_q$,用户得到自己的预测坐标l_x'、l_y'。

$$\boldsymbol{M}_q\boldsymbol{p}^*=\boldsymbol{D}_K(\boldsymbol{M}^{\mathrm{T}})^{-1}\boldsymbol{p}^*=\begin{bmatrix}0&&&&\\&\cdots&&&\\&&0&&\\&&&1&\\&&&&1\end{bmatrix}(\boldsymbol{M}^{\mathrm{T}})^{-1}\boldsymbol{p}^*=(0,\cdots,0,l_x',l_y')=\boldsymbol{p}_{xy} \tag{9}$$

BS中,用户最终解密得到的数据仅包含估计的自身位置坐标,而不包含服务器RSS指纹信息和参考点位置信息,增强了对服务器端的数据库资源的保护。

3.3 基于BF的改进方案ISBBF

实际应用中,为提高定位精度或降低定位算法复杂度,有很多算法会选择将用户先粗定位在某一

个局部区域内,再在该局部区域内进行精确定位。目前有多种方法可实现用户的粗定位,最常用的是在离线阶段对RP进行分簇;在线定位时,首先通过簇匹配完成用户的粗定位,再在该簇成员RP内完成用户的精确定位[17,22,27];文献[15-16]提出了无需离线信号分析的粗定位方法,利用邻近的位置点通常经历相同的AP信号空间分布而共享一组AP的现象,把目标粗定位到部分RP构成的局部区域内。通常可利用在线时用户可接收的AP集合,与指纹库中各参考点能接收的AP集合,进行相似度比较来实现局部定位。为此,本文提出了ISBBF算法,在达成隐私保护的同时对用户进行粗定位,提高预测的定位精度并减少时间延迟。ISBBF的基本思想是,引入BF,以匿名的形式通过AP集的匹配将用户粗定位在若干RP内,得到候选RP集。KNN加密域检索阶段只需在候选RP集上进行,而不需要遍历整个数据库。

(1) 布隆滤波器BF

BF是一种以节省空间的方式表示一个元素集合的数据结构[28]。为特定集合生成的BF允许在不知道集合本身的情况下对集合进行元素查询,确定一个元素是否在集合中,也常用来打乱序列原有的顺序。

定义一个BF:$B(A)$,其中$A=\{\mathrm{AP}_1,\mathrm{AP}_2,\cdots,\mathrm{AP}_d\}$为待映射的AP序列,$H=\{h_1,h_2,\cdots,h_k\}$是用于映射的$k$个哈希函数,则可用式(10)来描述$B(A)$

$$B(A)=\bigcup_{AP_i\in A}B(\mathrm{AP}_i) \tag{10}$$

式中$\boldsymbol{B}(\mathrm{AP}_i)=\{h_1(\mathrm{AP}_i),h_2(\mathrm{AP}_i),\cdots,h_k(\mathrm{AP}_i)\}$。

$$h_i\in H:\{0,1\}\overset{*}{\rightarrow}\{1,2,\cdots,m\} \tag{11}$$

其中哈希函数h_i将AP的序号作为输入,然后输出一个在$\{1,2,\cdots,m\}$中随机选择、均匀分布的数字。这样,布隆滤波器$B(A)$可以表示为一个m位长度的二进制向量$\boldsymbol{b}$,$\boldsymbol{b}$向量的第t位可表示为

$$b[t]=\begin{cases}1 & t\in B(A)\\ 0 & t\notin B(A)\end{cases} \tag{12}$$

本文采用BF对参考点或用户可接收的AP序列进行映射,映射后的每一个$\boldsymbol{b}$向量,都代表了一组特定的AP序列。本算法中哈希函数的数量k设置可为1,序列长度m可根据定位区域内AP的数量而设置。

(2) AP序列匿名粗定位

根据指纹库中每个RP的RSSI向量,可确定相应的能接收到的AP集。第n个RP能接收到的AP集为$\{\mathrm{AP}_n{}^j\}_{1\leqslant n\leqslant N}^{1\leqslant j\leqslant d}$,$d$表示整个定位区域内能接受到的AP总数,$N$表示RP总数。已有的研究表明,可接收AP序列相似的RP通常都分布在一个局部范围内,所以可用可接收AP序列来寻找匹配的RP完成粗定位[15-16]。考虑到直接基于AP集进行粗定位,则每个参考点能接收的AP集和用户当前接收的AP集都会暴露给第三方。因此引入了BF对AP集进行映射,再进行基于$\boldsymbol{b}$向量的匿名匹配。

图3给出了进行BF滤波和匿名匹配的一个实例。假设服务器端的RP_n接收到3个AP$\{\mathrm{AP}_1,\mathrm{AP}_2,\mathrm{AP}_6\}$,用户端接收到2个AP$\{\mathrm{AP}_1,\mathrm{AP}_6\}$。服务器端AP序列的映射过程为:$\mathrm{AP}_1\rightarrow b_n[3]$,$\mathrm{AP}_2\rightarrow b_n[8]$,$\mathrm{AP}_6\rightarrow b_n[5]$;用户端的AP序列的映射过程为:$\mathrm{AP}_1\rightarrow b_n[3]$,$\mathrm{AP}_6\rightarrow b_n[5]$。匿名匹配采用$\boldsymbol{b}_n$和$\boldsymbol{b}_q$的汉明距离来完成AP集相似度的计算。图3中$\boldsymbol{b}_n$和$\boldsymbol{b}_q$的汉明距离为1,与在AP集明文对应的距离计算中一致。通过计算每个RP与用户之间$\boldsymbol{b}$向量的距离大小,可得到距离最近的若干个RP,由此把用户粗定位至数据库中这些RP所在的区域。当$m\geqslant d$时,不会出现哈希碰撞,映射后的数值一一对应,即匿名匹配过程不会影响粗定位的结果。

利用匿名匹配选取的距离最近的N'个RP,作为后续精确定位的候选RP集,然后在此集合内利用

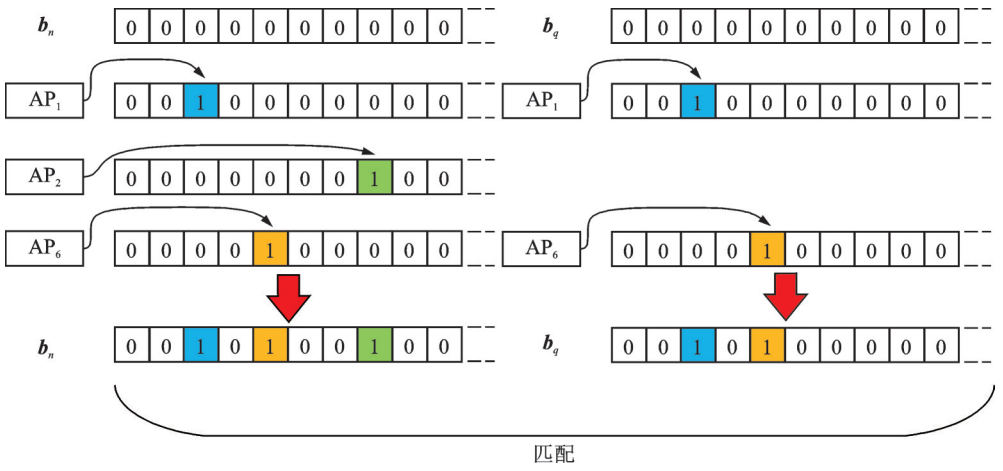

图3 AP集的BF滤波及匿名匹配过程

Fig.3 BF filtering and anonymous matching process for AP sets

BS完成加密定位。ISBBF算法流程中服务器端、用户端和第三方的伪代码如下所示:

输入:指纹库$\boldsymbol{D}$、请求定位服务用户的RSSI向量$\boldsymbol{q}$、加密矩阵$\boldsymbol{M}$、哈希函数$\boldsymbol{H}$、算法各参数:候选RP数量N'、KNN近邻数K、BF长度m

输出:目标用户的估计位置(l_x', l_y')

服务器端:

利用式(3,5)完成指纹拓展和加密

for $n=1,2,\cdots,N$ do

确定RP_n指纹库候选AP集:$A_d=\{\mathrm{AP}_1,\mathrm{AP}_2,\cdots,\mathrm{AP}_{d'}\}$

将$\boldsymbol{b}_n$所有位清零

for $d_i=1,2,\cdots,d'$ do

利用式(12)将AP编号AP_{di}映射得到t_{di},设$\boldsymbol{b}_n[t_{di}]=1$

end for

end for

将$\boldsymbol{b}_n$, $n\in\{1,2,\cdots,N\}$发送到第三方

用户端:

利用式(6,7),完成RSSI向量加密,并发送给第三方

用户确定自身候选AP集:$A_{dq}=\{\mathrm{AP}_1,\mathrm{AP}_2,\cdots,\mathrm{AP}_{d''}\}$

将$\boldsymbol{b}_q$所有位清零

for $d_{qi}=1,2,\cdots,d''$ do

利用式(12)得到t_{dqi},设$b_q[t_{dqi}]=1$

end for

将$\boldsymbol{b}_q$发送到第三方

基于加密数据p_K^*和解密密钥$\boldsymbol{M}_q$,得到估计位置(l_x', l_y')

第三方:

for $n=1,2,\cdots,N$ do

计算RP_n与用户$\boldsymbol{b}$向量的汉明距离：$d_{\mathrm{HD}}[n]=\sum_{t=1}^{m}b_n[t]\oplus b_q[t]$

end for

获取N'个最小汉明距离的$\mathrm{RP}_{n'}$，确定其候选RP集：$R_{N'}=\{\mathrm{RP}_1,\mathrm{RP}_2,\cdots,\mathrm{RP}_{n'},\cdots,\mathrm{RP}_{N'}\}$

在$R_{N'}$上进行加密检索，得到参考RP集$\{\mathrm{RP}_1,\mathrm{RP}_2,\cdots,\mathrm{RP}_k,\cdots,\mathrm{RP}_K\}$所对应的$K$个指纹$\boldsymbol{p}_K=\{\boldsymbol{p}_1',\boldsymbol{p}_2',\cdots,\boldsymbol{p}_k',\cdots,\boldsymbol{p}_K'\}$

取平均值后得到$\boldsymbol{p}_K^*$，发送到用户端

3.4 方案正确性分析

本文提出的隐私保护定位算法能够根据加密指纹库和加密的查询请求的内积，计算并返回前K个近邻位置。下面给出方案的正确性分析。

定理1 正确性。在隐私定位算法的密文域检索中，对于两个检索向量$\boldsymbol{p}_1'$和$\boldsymbol{p}_2'$，第三方可以通过对应的加密指纹库向量和加密查询用户请求向量$\boldsymbol{q}'$之间的内积大小关系，确定两个检索向量与查询向量之间的欧式距离大小关系。即当$\boldsymbol{p}_i'\cdot\boldsymbol{q}'\geqslant\boldsymbol{p}_j'\cdot\boldsymbol{q}'$时，$d(\boldsymbol{p}_i',\boldsymbol{q}')\leqslant d(\boldsymbol{p}_j',\boldsymbol{q}')$，则用户更靠近$\boldsymbol{p}_i$参考点；反之，用户则更靠近$\boldsymbol{p}_j$参考点。

以指纹库中任意2个RP的加密指纹$\boldsymbol{p}_1'$和$\boldsymbol{p}_2'$与用户加密指纹$\boldsymbol{q}'$的距离计算为例，密文域距离计算获得最近邻参考点的证明过程如下：

证明：

$$(\boldsymbol{p}_1'-\boldsymbol{p}_2')\cdot\boldsymbol{q}'=(\boldsymbol{p}_1'-\boldsymbol{p}_2')^{\mathrm{T}}\boldsymbol{q}'=(\boldsymbol{M}^{\mathrm{T}}\hat{\boldsymbol{p}}_1-\boldsymbol{M}^{\mathrm{T}}\hat{\boldsymbol{p}}_2)^{\mathrm{T}}\boldsymbol{M}^{-1}\hat{\boldsymbol{q}}=(\hat{\boldsymbol{p}}_1-\hat{\boldsymbol{p}}_2)^{\mathrm{T}}\hat{\boldsymbol{q}}=(\boldsymbol{p}_1-\boldsymbol{p}_2,-0.5\|\boldsymbol{p}_1\|^2+0.5\|\boldsymbol{p}_2\|^2,l_{x1}-l_{x2},l_{y1}-l_{y2})\cdot r(\boldsymbol{q},1,0,0)^{\mathrm{T}}=(\boldsymbol{p}_1-\boldsymbol{p}_2)^{\mathrm{T}}(r\boldsymbol{q})+(-0.5\|\boldsymbol{p}_1\|^2+0.5\|\boldsymbol{p}_2\|^2)r=r(\boldsymbol{p}_1-\boldsymbol{p}_2)^{\mathrm{T}}\boldsymbol{q}-0.5r\|\boldsymbol{p}_1\|^2+0.5r\|\boldsymbol{p}_2\|^2=0.5r\left(\left(\|\boldsymbol{p}_2\|^2-2\boldsymbol{p}_2\cdot\boldsymbol{q}+\|\boldsymbol{q}\|^2\right)-\left(\|\boldsymbol{p}_1\|^2-2\boldsymbol{p}_1\cdot\boldsymbol{q}+\|\boldsymbol{q}\|^2\right)\right)=0.5r(d(\boldsymbol{p}_2,\boldsymbol{q})-d(\boldsymbol{p}_1,\boldsymbol{q})) \tag{13}$$

因为r为随机正数，当$(\boldsymbol{p}_1'-\boldsymbol{p}_2')\cdot\boldsymbol{q}'>0$，即$\boldsymbol{p}_1'\cdot\boldsymbol{q}'>\boldsymbol{p}_2'\cdot\boldsymbol{q}'$时，$d(\boldsymbol{p}_2,\boldsymbol{q})>d(\boldsymbol{p}_1,\boldsymbol{q})$，说明在$\boldsymbol{p}_1$和$\boldsymbol{p}_2$对应的两个RP中，$\boldsymbol{p}_1$对应的$\mathrm{RP}_1$是相对近邻参考点。

证毕。

4 安全性分析

本文假设服务器和用户可信，第三方不可信且不与用户合谋，主要分析第三方对服务器和用户发起的攻击。基于以上假设，对指纹数据库安全性、用户数据安全性和匹配安全性进行分析。

4.1 指纹数据安全性

在ASPE方案[19]中，用户端执行1次定位即可从第三方获取服务器的K个指纹信息。在本文所提的BS中，服务器将各RP的物理位置l_x和l_y嵌入到RSS中，经加密运算后保护了参考点的位置信息；相比ASPE方案，BS增强了对服务器指纹库中RP物理位置和RSS的保护：(1) 对于第三方，从服务器端接收到的是加密的数据$\boldsymbol{P}'=\{\boldsymbol{p}_i'\}$，即$\boldsymbol{p}_i'=\boldsymbol{M}^{\mathrm{T}}\hat{\boldsymbol{p}}_i$，无法获得参考点的位置信息，只能进行唯密文攻击。攻击者若想得到$\boldsymbol{p}_i$，需要获得$\boldsymbol{M}$并由此解密得到$\hat{\boldsymbol{p}}_i$，然后再推断出$\boldsymbol{p}_i$。显然，攻击者无法推出$(d+3)\times(d+3)$维的随机矩阵$\boldsymbol{M}$。因此，在无密钥的情况下，攻击者无法获得明文指纹数据库。(2) 对于已知明文攻击，敌手获取了指纹库的部分明文信息，假设有y个。根据方程组求解规律，敌手需要利用多项式求解方式来推出获取的部分明文和密文的对应关系，但其明密文组合至少有$O(n^y)$种可能，代价是指数级的。(3)对于用户端，第三方返回给用户的是加密的预测位置，用户解密后只能获得自身的预测位置而无法获知参考点的指纹数据。进一步假设用户不可信。恶意用户可以发送多次定位请求到第三方，

但容易因定位请求过于频繁,被第三方停止定位服务。恶意用户无法建立与服务器相似的指纹库,因此方案中指纹数据库是安全的。

4.2 用户数据安全性

用户发送到第三方的数据经过加密,第三方无法获得明文的用户指纹数据,攻击者在无密钥的情况下只能进行唯密文攻击。攻击者若想得到$\boldsymbol{q}$,需要获得$\boldsymbol{M}$并由此解密得到$\hat{\boldsymbol{q}}$,显然,攻击者无法推出$(d+3)\times(d+3)$维的随机矩阵$\boldsymbol{M}$。进一步假设攻击者通过某种方式获取了密钥$\boldsymbol{M}$,需要从$\hat{\boldsymbol{q}}$推断出$\boldsymbol{q}$。由于用户在加密时引入了一次一密的随机数r,攻击者仍无法获取用户的请求指纹。

同样,对于用户的估计位置信息,攻击者在无密钥的情况下不能解密包含位置信息的定位数据$\boldsymbol{p}^*$,进一步假设攻击者通过某种方式获取了密钥$\boldsymbol{M}$,还需知道指纹数据中随机嵌入的位置信息才能获取用户的估计位置,因此方案中用户数据是安全的。

4.3 匹配安全性

在ISBBF中,第三方接收服务器和用户发送的$\boldsymbol{b}$序列进行匹配,来实现用户的粗定位。第三方不知道生成$\boldsymbol{b}$序列的输入信号:各RP可接收的AP序列,也不知道相应的哈希函数集$H=\{h_1,h_2,\cdots,h_k\}$。所以第三方利用$\boldsymbol{b}$序列不能获取服务器和用户的有用信息,只能获得用户能接收到的大致AP个数;进一步,匿名匹配后,第三方获知的是用户的最近邻RP的序号及对应的加密指纹。在不知道密钥$\boldsymbol{M}$的情况下,第三方不能由此推测或获知有关服务器和用户的信息,因此所提的匿名匹配过程是安全的。

5 实验结果与分析

5.1 实验配置

本文采用两个在实际环境中采集的WiFi RSSI数据集来衡量所提算法的性能。所用的公共数据集[29-30]的采集环境是一个覆盖面积为50 m×30 m的办公楼层,包括了房间、走廊和大厅等场所。其中,参考点的采样间隔为1 m,整个环境内检测的AP个数为32。在公共数据集中的711个采样点中,随机选择200个作为测试用户,即用户端定位测试点。

另一个实验数据集是在本校光电大楼的9楼自主采集获取。这个采集环境包括了走廊和大厅等要素,并以1.8 m为采样间隔,设定了135个参考点,构建了实验室离线指纹库。同时,又选取了43个位置点用于在线用户端测试。在整个楼层中,采集到的AP数量为349个。表1给出两个数据集的详细参数。

文献[29]给出了公共数据的详细室内布局图,实验室数据集的室内布局图如图4所示,其中黑色圆形表示参考点,红色三角形表示测试点。

BS、ISBBF和ASPE均基于Matlab2021a平台进行测试,所运行的PC环境如下:MacBookPro17,1@Apple M1,16GB RAM,64bit操作系统。

表1 两个数据集的详细参数

Table 1 Detailed parameters for two datasets

数据集	参考点数量	测试点数量	AP数量
公共数据集	511	200	32
实验室数据集	135	43	349

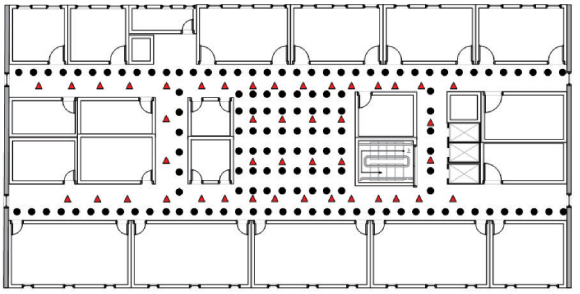

图4 实验室数据集的室内布局

Fig.4 Indoor layout of the lab dataset

5.2 时间效率

在公共数据集测试中，选择的测试点为200个。实验中的各个参数设定如下：参考点个数$N=511$，AP个数$d=32$，序列长度$m=50$，KNN最近邻数$K=3$，候选RP集个数$N'=10$。

表2给出了BS、ISBBF和ASPE各算法执行单次定位的时间开销，其中，“服务器端”在整个定位服务中只需执行1次数据库加密，ISBBF中$\boldsymbol{b}$序列长度为50。从表2可观察到，BS和ASPE在200个测试点的定位总体开销基本一致，因为与ASPE相比较，BS只增加了嵌入指纹位置的处理过程；而ISBBF所需的时间开销则明显减低，且其主要开销在服务器端，这是由于服务器端对每个RP进行了$\boldsymbol{b}$向量的生成过程；另一个明显降低的时间开销体现在第三方。相比较ASPE和BS，ISBBF引入了在线RP的匿名匹配过程。在ASPE和BS中，KNN加密域检索需要在整个数据库511个RP中进行；而在ISBBF中，只需要在候选RP集$\{\mathrm{RP}\}_{N'}$中的元素中进行检索比较。实验中N'设置为10，减少了大量的时间消耗。所以，相比较其他两种算法，ISBBF的时延较小，在实际定位中有更好的实时性。

表2 公共数据集上的算法性能

Table 2 Algorithm performance on public datasets

算法	时间开销/ms				平均定位误差/m
	定位总体开销	服务器端	用户端	第三方	
ASPE	155.32	0.85	0.18	154.29	4.11
BS	147.23	0.87	0.33	146.03	4.11
ISBBF	55.18	54.76	0.21	0.21	4.12

图5和图6分别给出了KNN不同最近邻个数K对第三方和用户端进行单次定位的时间开销的影响。从图5，6可以看出，在增大K的情况下，3个方案中第三方和用户端的耗时基本保持稳定，表明最近邻个数K对时间开销的影响不大。在相同的最近邻个数下，相比ASPE和BS，ISBBF中第三方的时间开销明显较低；而3种方案中用户端定位的时间开销均很接近。这表明，将用户粗定位在局部区域内的方式能有效降低第三方的时间开销。

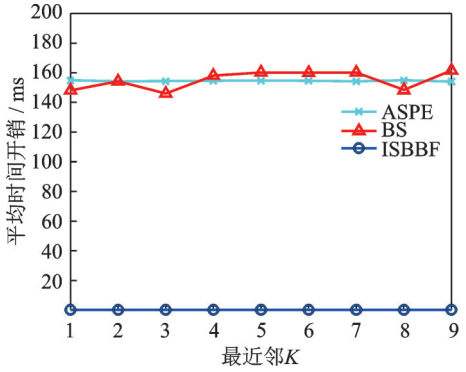

图5 第三方定位的时间开销

Fig.5 Average time overhead for third-party targeting

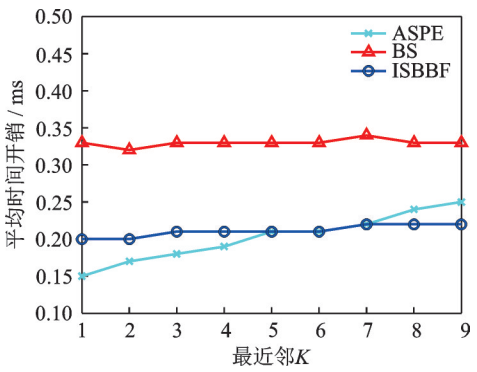

图6 用户端定位的时间开销

Fig.6 Average time overhead for user-side targeting

5.3 定位性能

图7给出了在公共数据集中第32个测试点的定位结果。从图7中可观察到，ISBBF采用BF后的匹配RP与不采用BF后匹配的RP一致，匿名匹配最终找到的最近邻RP也与明文时找到的最近邻RP一样。

图8给出了ASPE、BS和ISBBF的候选RP的位置，可以看到ASPE和BS的定位结果完全一致，这也符合算法的设计，均是在密文域找到K近邻进行定位，BS中指纹位置的嵌入并不影响计算结果；而

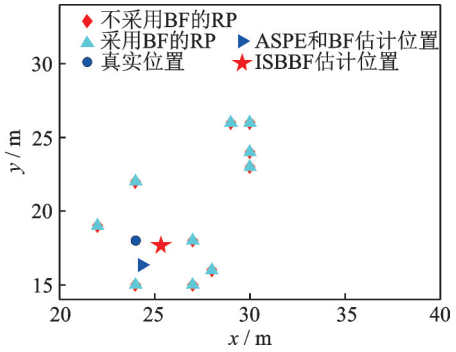

图7 第32个测试点的定位结果

Fig.7 Positioning results for the 32nd test point

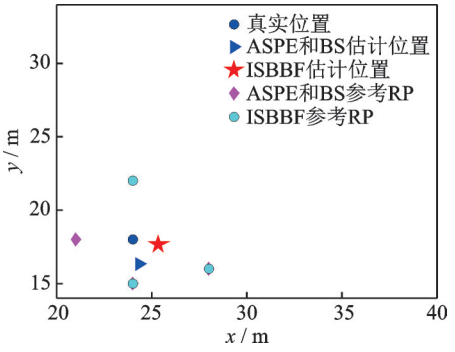

图8 第32个测试点的RP

Fig.8 RP for the 32nd test point

ISBBF与它们略有差别,具体分析可知有2个RP位置重合,1个RP位置不重合。

图9给出3种方案在公共数据集上每个测试点的定位误差累积分布函数(Cumulative distribution function, CDF)。从图9可以看出3种方案的CDF曲线几乎重合,这表明ISBBF中基于BF匿名的RP选择不会影响定位精度,但能保护服务器指纹数据库中AP列表和用户RSS信号的AP序列,并降低了计算开销,性能表现良好。在两种方案中,90%测试点的误差均在8 m以内。

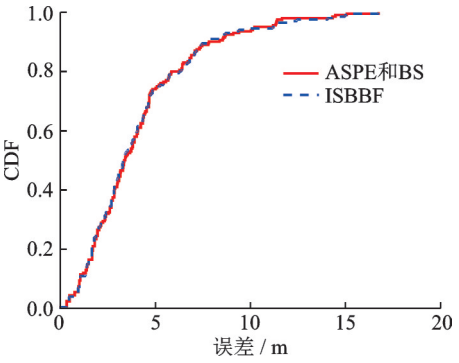

图9 公共数据集上的定位误差累积分布函数

Fig.9 Positioning error cumulative distribution function on a public dataset

上述实验结果表明,在线匿名匹配能有效降低第三方的时间开销,但不会增大定位精度和泄露隐私。ISBBF具有隐私保护性能良好、低延时、定位精度较高且算法相对简单等优势。

5.4 方案在实验室数据集中的具体性能

为了进一步分析AP和RP数量对算法的影响,在分布有大量AP和较少RP的实验室数据集中对算法性能进行测试。

5.4.1 时间效率

实验室数据集上的算法性能如表3所示。结合表2,3可以发现,与公共数据集中算法的性能比较,在实验室数据集中ASPE和BS在服务器端及用户端开销较大。这是因为实验室数据集中AP数量多,相应需要较大阶数的密钥进行加密和解密计算。而第三方则开销较小。这是因为两种算法在每次定

表3 实验室数据集上的算法性能

Table 3 Algorithm performance on lab datasets

算法	时间开销/ms				平均定位误差/m
	定位总体开销	服务器端	用户端	第三方	
ASPE	47.47	7.04	27.77	12.66	2.06
BS	32.20	7.04	12.97	12.19	2.06
ISBBF (m=50)	77.77	64.95	12.64	0.18	2.54
ISBBF (m=500)	117.22	103.71	13.06	0.45	2.18

位中都需在整个加密的离线指纹库内进行KNN检索,以得到距离最小的K个RP。实验室数据集中RP的数量较公共数据集少,降低了KNN检索的时间。对于ASPE和BS算法,数据集中AP的数量会影响服务器端和用户端的时间开销,而RP数则影响第三方的时间开销。

ISBBF采用在线匿名匹配和BF得到候选RP集$\{\mathrm{RP}\}_{N'}$,并在这个集合上进行KNN加密域检索,数据库中的RP个数的增加不会增加ISBBF第三方的计算开销。从表3可以发现,ISBBF算法的总定位开销较大,其主要是由服务器端开销增大引起。当数据集中有大量AP时,服务器端需要进行哈希运算完成映射的AP数量较多,导致计算开销较大。值得注意的是服务器端对指纹的加密、映射等处理在离线阶段已完成。在线定位时的计算开销主要取决于用户端和第三方的时间开销。在具有大量AP的数据集中,ISBBF仍可以有效提高定位实时性。

5.4.2 定位精度

图10给出了3种方案在实验室数据集上每个测试点的误差累积分布函数图。3种方案中,均有90%的测试点误差都在4.5 m以内。m较小时,ISBBF(m=50)定位误差较大;m较大时,ISBBF(m=500)定位误差很小。BF中m的取值会影响定位性能。

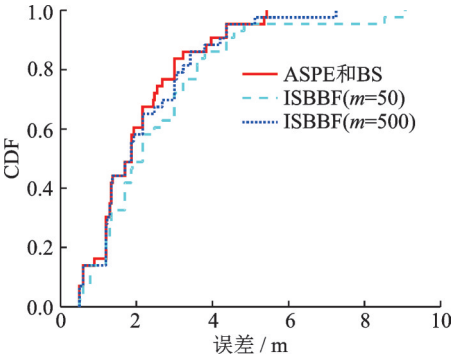

图10 两种方案在实验室数据集上的CDF图

Fig.10 CDF figures of two protocols on laboratory datasets

图3表明BF映射中的哈希碰撞会影响定位精度。公共数据集中AP个数小于50,所以映射的结果不会出现覆盖;然而实验室数据中的AP个数达到了349个,使得不同的AP映射到$\boldsymbol{b}$序列的位置出现碰撞,AP序列与$\boldsymbol{b}$向量不再唯一对应。碰撞结果直接影响在线匿名匹配所确定的候选RP集,KNN定位后的位置估计存在偏差。所以在图10中,ISBBF(m=50)的定位精度较ASPE和BS下降,且出现了大的定位误差。当增大$\boldsymbol{b}$序列的长度m时,ISBBF(m=500)的定位精度提高,与ASPE和BS很接近。在不同的环境中应用ISBBF算法时,为保证定位性能,一般选取$\boldsymbol{b}$序列的长度要大于定位所用的AP个数,保证最后预测定位的K近邻参考RP选择的正确性。实验结果中,ASPE和BS的参考RP集一致。表4给出了在两个数据集中,ASPE和BS的参考RP集与ISBBF参考RP集合的匹配正确率。实验中,定义严格匹配条件为:每个测试点的参考RP集一一对应为成功匹配,对应的相同参考点个数和为匹配点个数;宽松匹配条件为:每个测试点的参考RP集内存在相同元素即为成功匹配,对应的相同元素个数为匹配点个数;匹配正确率定义为:匹配点总个数与参考点总个数的比值。

表4 参考点匹配正确率

Table 4 Reference point matching accuracy

匹配条件	公共数据集匹配正确率/%	实验室数据集匹配正确率/%	
	m = 50	m = 50	m = 500
严格匹配	93.17	62.79	88.37
宽松匹配	95.67	73.64	92.25

结合表3,4,可以发现匹配率越高,平均定位误差越低。在实验室数据集中,ISBBF(m=50)匹配率较低,定位精度较ASPE和BS下降。当m大于AP个数时,ISBBF(m=500)的匹配率较ISBBF(m = 50)明显上升,定位精度提高。

综上所述,ASPE、BS和ISBBF的计算开销都在毫秒级,能保证用户定位的实时性。采用了BF的ISBBF通过在线匿名匹配过程在不向第三方泄露数据库信息的同时进一步降低了在线定位时的计算开销。为了提高匹配率和定位精度,需要保证m的取值要大于AP个数。

6 结束语

针对目前各种基于加密的隐私保护算法会泄漏部分数据库指纹信息的问题,本文提出了BS和ISBBF两种改进的KNN加密域检索的室内定位隐私保护算法,能够进一步降低计算开销并提高服务器端和用户端隐私数据的安全性。其中ISBBF算法利用BF滤波器可通过AP的匿名匹配,实现用户的粗定位,这不仅降低了本隐私保护算法中第三方大量的计算开销,同时可与各类室内定位算法相结合,实现隐私保护下对用户的局部区域定位。理论分析和实验结果表明,所提的定位算法在保护服务器指纹库数据、用户端定位隐私的同时,能够在保证定位精度的前提下减少时间开销。与同类隐私保护算法相比较,在服务器和用户端的隐私保护上具有显著优势。

参考文献:

[1] YANG X, WU Z, ZHANG Q. Bluetooth indoor localization with Gaussian-Bernoulli restricted Boltzmann machine plus liquid state machine[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71(1): 1-8.

[2] IBNATTA Y, KHALDOUN M, SADIK M. Indoor localization system based on mobile access point model MAPM using RSS with UWB-OFDM[J]. IEEE Access, 2022, 10(1): 46043-46056.

[3] LEE S, KIM J, MOON N. Random forest and WiFi fingerprint-based indoor location recognition system using smart watch[J]. Human-Centric Computing and Information Sciences, 2019, 9(6): 1-14.

[4] HUANG G, HU Z, WU J, et al. WiFi and vision-integrated fingerprint for smartphone-based self-localization in public indoor scenes[J]. IEEE Internet of Things Journal, 2020, 7(8): 6748-6761.

[5] LI S, HEDLE M, BENGSTON K, et al. Passive localization of standard WiFi devices[J]. IEEE Systems Journal, 2019, 13(4): 3929-3932.

[6] 金施嘉珞, 乐燕芬, 许远航,等. 跨异构设备的室内Wi-Fi指纹定位方法[J]. 数据采集与处理, 2022, 37(3): 703-714.
JIN Shijialuo, LE Yanfen, XU Yuanhang, et al. Indoor Wi-Fi fingerprint location method across heterogeneous devices[J]. Journal of Data Acquisition and Processing, 2022, 37(3): 703-714.

[7] 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.
WANG Huiqiang, GAO Kaixuan, LV Hongwu. Survey of high-precision localization and the prospect of future evolution[J]. Journal on Communications, 2021, 42(7): 198-210.

[8] LAN T, WANG X, CHEN Z, et al. Fingerprint augment based on super-resolution for WiFi fingerprint based indoor localization[J]. IEEE Sensors Journal, 2022, 22(12): 12152-12162.

[9] 张笑宇, 沈超, 蔺琛皓,等. 面向机器学习模型安全的测试与修复[J]. 电子学报, 2022, 50(12): 2884-2918.
ZHANG Xiaoyu, SHEN Chao, LIN Chenhao, et al. The testing and repairing methods for machine learning model security[J]. Acta Electronica Sinica, 2022, 50(12): 2884-2918.

[10] GRUTESER M, GRUNWALD D. Anonymous usage of location-based services through spatial and temporal cloaking[C]//Proceedings of the 1st International Conference on Mobile Systems, Applications and Services. San Francisco, USA: USENIX Association, 2003: 31-42.

[11] YIU M L, JENSEN C S, HUANG X, et al. Spacetwist: Managing the trade-offs among location privacy, query performance, and query accuracy in mobile services[C]//Proceedings of the 24th International Conference on Data Engineering. Cancun, Mexico: IEEE, 2008: 366-375.

[12] BERESFORD A R, STAJANO F. Mix zones: User privacy in location-aware services[C]//Proceedings of IEEE Annual Conference on Pervasive Computing and Communications Workshops. Orlando, USA: IEEE, 2004: 127-131.

[13] 左开中, 刘蕊, 赵俊, 等. 融合语义信息的时空关联位置隐私保护方法[J]. 西安电子科技大学学报, 2022, 49(1): 67-77.
ZUO Kaizhong, LIU Rui, ZHAO Jun, et al. Method for the protection of spatiotemporal correlation location privacy with semantic information[J]. Journal of Xidian University, 2022, 49(1): 67-77.

[14] HIGUCHI T, MARTIN P, CHAKRABORTY S, et al. AnonyCast: Privacy-preserving location distribution for anonymous crowd tracking systems[C]//Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous

Computing. Osaka, Japan: ACM, 2015: 1119-1130.

[15] ZHANG T, CHOW S S M, ZHOU Z, et al. Privacy-preserving Wi-Fi fingerprinting indoor localization[C]//Proceedings of Advances in Information and Computer Security: 11th International Workshop on Security. Tokyo, Japan: Springer, 2016: 215-233.

[16] LI H, SUN L, ZHU H, et al. Achieving privacy preservation in WiFi fingerprint-based localization[C]//Proceedings of IEEE INFOCOM 2014—IEEE Conference on Computer Communications. Toronto, Canada: IEEE, 2014: 2337-2345.

[17] OLIVEIRA S R M, ZAIANE O R. Privacy preserving clustering by data transformation[J]. Journal of Information and Data Management, 2010, 1(1): 37-37.

[18] LIU K, GIANNELLA C, KARGUPTA H. An attacker's view of distance preserving maps for privacy preserving data mining [C]//Proceedings of the 10th European Conference on Principles and Practice of Knowledge Discovery in Databases. Berlin, Germany: [s.n.], 2006: 297-308.

[19] WONG W K, CHEUNG D W, KAO B, et al. Secure KNN computation on encrypted databases[C]//Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data. Rhode Island, USA: ACM, 2009: 139-152.

[20] 乐燕芬, 金施嘉珞, 朱一鸣, 等. 基于联合分簇和LASSO的室内指纹定位算法[J]. 数据采集与处理, 2020, 35(7): 1097-1105.
LE Yanfen, JIN Shijialuo, ZHU Yiming, et al. Fingerprinting indoor localization using hybrid clustering and LASSO[J]. Journal of Data Acquisition and Processing, 2020, 35(7): 1097-1105.

[21] 乐燕芬, 许远航, 施伟斌. 基于DPC指纹子空间匹配的室内WiFi定位方法[J]. 仪器仪表学报, 2021, 42(11): 106-114.
LE Yanfen, XU Yuanhang, SHI Weibin. WiFi fingerprint based indoor positioning with subspace matching and DPC[J]. Chinese Journal of Scientific Instrument, 2021, 42(11): 106-114.

[22] LE Yanfen, ZHANG Hena, SHI Weibin, et al. Received signal strength based indoor positioning algorithm using advanced clustering and kernel ridge regression[J]. Frontiers of Information Technology & Electronic Engineering, 2021, 22(6): 827-838.

[23] LIU Guiqi, QIAN Zhihong, LI Hualiang, et al. Indoor positioning algorithm based on effective AP selection and multi-classification LDA[J]. Journal on Communications, 2021, 42(11): 109-120.

[24] ZHANG Wei, LIU Kan, ZHANG Weidong, et al. Deep neural networks for wireless localization in indoor and outdoor environments[J]. Neurocomputing, 2016, 194: 279-287.

[25] HOFHEINZ D, JAGER T. Tightly secure signatures and public-key encryption[J]. Designs, Codes and Cryptography, 2016, 80: 29-61.

[26] MAO X, LAI J, MEI Q, et al. Generic and efficient constructions of attribute-based encryption with verifiable outsourced decryption[J]. IEEE Transactions on Dependable and Secure Computing, 2015, 13(5): 533-546.

[27] KHALAJMEHRABADI A, GATSIS N, PACK D J, et al. A joint indoor WLAN localization and outlier detection scheme using LASSO and elastic-net optimization techniques[J]. IEEE Transactions on Mobile Computing, 2016, 16(8): 2079-2092.

[28] CALDERONI L, PALMIERI P, MAIO D. Location privacy without mutual trust: The spatial Bloom filter[J]. Computer Communications, 2015, 68: 4-16.

[29] TÓTH Z, TAMÁS J. Miskolc IIS hybrid IPS: Dataset for hybrid indoor positioning[C]//Proceedings of 2016 26th International Conference Radioelektronika. Kosice, Slovakia: IEEE, 2016: 408-412.

[30] TÓTH Z, TAMÁS J. Miskolc IIS hybrid IPS[EB/OL]. [2016-07-03]. https://archive.ics.uci.edu/dataset/375/miskolc+iis+hybrid+ips.

作者简介:

欧锦添(1997-),男,硕士研究生,研究方向:室内定位和隐私保护,E-mail:212230405@st.usst.edu.cn。

乐燕芬(1978-),**通信作者**,女,博士,副教授,研究方向:室内定位、无线网络和多媒体,E-mail: leyanfen@usst.edu.cn。

施伟斌(1967-),男,博士,副教授,研究方向:无线传感器网络与物联网,E-mail: shiweibin@usst.edu.cn。

(编辑:陈珺)