# 无线体域网中隐私保护安全*k*NN查询协议

张大方[1]，徐鸿玥[1]，李 睿[2]

(1. 湖南大学信息科学与工程学院 长沙 410082；2. 东莞理工学院计算机与网络安全学院 广东 东莞 523808)

**【摘要】**针对无线体域网中的数据隐私问题，提出了一种适用于无线体域网的安全*k*NN查询协议，能够保护数据隐私与访问权限控制。该协议主要分3个部分，首先采用非对称矩阵向量积保值加密机制(ASPE)对数据和查询条件分别进行加密，从而保护数据的隐私；其次基于R树的桶划分索引结构BRtree，将数据划分到桶节点后采用剪枝策略去除不必要的查询来提高查询效率；最后基于数据层面的访问权限授予与回收机制，从ASPE加密密钥中分解出权限密钥，通过可信第三方实现了访问权限控制和访问权限迁移。并在真实移动健康数据集上验证了该方案的有效性。

**关 键 词** 权限控制； 矩阵加密； 安全k邻近查询； 无线体域网

**中图分类号** TP393 **文献标志码** A doi:10.3969/j.issn.1001-0548.2017.05.014

# Privacy Preserving *k*NN Query Protocol for Wireless Body Sensor Networks

ZHANG Da-fang[1], XU Hong-yue[1], and LI Rui[2]

(1. College of Information Science and Engineering, Hunan University Changsha 410082;
2. School of Computer and Network Security, Dongguan University of Technology Dongguan Guangdong 523808)

**Abstract** For the data privacy in wireless body area network (WBAN), a secure privacy preserving *k*-nearest neighbor (*k*NN) query protocol for WBAN is proposed. This protocol can protect data privacy and access control by encrypting both data and queries with asymmetric scalar-product-preserving encryption (ASPE). To improving searching efficiency, we combine the technologies of R-tree and bucket partition and propose a data structure, named BRtree, for indexing data items. BRtree can significantly eliminate the unnecessary searching branches. In order to achieve access control, we separate an access key from the encryption key and introduce a trusted third authority to manage access rights and access rights transferring. The experimental results validate the efficiency of our scheme.

**Key words** access control; matrix encryption; secure k-nearest neighbor query; WBAN

随着云计算和无线传感器网络的迅速发展，无线体域网(WBAN)得到了广泛的应用[1]。在WBAN中包含3类节点：1) 医疗数据采集节点，该类节点部署在被检测对象的身体上，负责采集人体的生理数据，并通过个人基站(如PDA、智能手机等)将采集的数据上传到指定服务器；2) 服务器节点，该类节点往往位于医院指定的区域，负责接受授权病人发送来的医疗数据，并执行授权用户(病人的责任医生)发来的查询请求；3) 数据用户节点，授权用户通过该类节点向服务器发送数据查询请求，获得相应的查询结果。

然而，WBAN模型存在重大的安全隐患。由于用于采集数据的无线传感器网络与用于存储数据的服务器处在不同的安全领域，服务器可能泄漏其敏感医疗数据，导致病人隐私信息泄漏。近年来，医疗数据保护问题导致个人隐私信息泄漏的事件频发[2]，为无线体域网设计出相应的数据安全传输、查询、及分析协议至关重要。

本文研究隐私保护下的*k*邻近(*k*-nearest neighbor, *k*NN)查询问题。*k*NN在无线体域网中扮演重要的角色，医生可根据某疾病的特征值查询出患者医疗数据中与该病症特征值最邻近的*k*个数据来快速确定出现该病症的具体时间，进而掌握该患者的发病规律，制定出相应的治疗方案。

如何在无线体域网中实现隐私保护*k*NN查询是一个极具挑战性的问题：1) 为了实现对病人数据的

收稿日期：2016－01－18； 修回日期：2017－03－07
基金项目：国家自然科学基金(61370226, 61472130, 61672156)；国家973项目(2012CB315805)
作者简介：张大方(1959－)，男，博士，教授，主要从事网络安全、可信系统与网络、网络测试等方面的研究.

隐私保护，个人基站需对数据加密后再提交给服务器；同样，数据用户需对查询条件进行加密后发送给服务器进行查询处理。因此，需要解决服务器在既不知道数据的真实值，也不知道查询条件的情况下的*k*NN查询问题。2) 为了防止数据用户滥用病人医疗数据导致隐私信息的泄漏，需要从数据级上解决用户访问权限及访问权限的迁移问题。

目前对无线体域网中的数据安全研究集中在身份认证上[3-5]，鲜有文献讨论数据的隐私保护和访问控制问题。文献[6-8]研究了无线传感网中的安全认证协议，在两层传感器网络中，研究者就网络安全查询协议开展了大量的研究[9-13]，目前还没有工作探讨安全*k*NN查询协议。在云计算中，文献[14]提出了一种基于非对称矩阵向量积保值的加密机制ASPE，并证明该机制能抵抗已知部分明文攻击(KPA)，随后研究者基于ASPE机制提出了其他隐私保护协议[15-16]。

# 1　非对称矩阵向量积保值加密机制

文献[14]提出了一种基于ASPE的隐私保护安全*k*NN查询机制。假设数据采集单元采集的数据为集合$\boldsymbol{D}$，$\boldsymbol{D}$中数据有$u$个属性，表示为$A^1,A^2,\cdots,A^u$，$\boldsymbol{D}$中某条数据表示为$(d^1,d^2,\cdots,d^u)$。查询条件是$\{\boldsymbol{q}, k\}$，$\boldsymbol{q}$与数据$\boldsymbol{D}$在同一维度空间，即$\boldsymbol{q}=(q^1,q^2,\cdots,q^u)$。*k*NN的查询结果是返回数据集合$\boldsymbol{D}$中与查询条件$\boldsymbol{q}$距离最近的$k$个数据。$\boldsymbol{d}_i$和$\boldsymbol{d}_j$ $(1\leqslant i,j\leqslant n, i\neq j)$为采集到的两条数据，需判断$\boldsymbol{d}_i$和$\boldsymbol{d}_j$哪一个距离查询条件$\boldsymbol{q}$更近。

为了对数据进行隐私保护，$\boldsymbol{D}$中所有病人的数据在上传给服务器前都要进行加密，用Enc($\boldsymbol{D}$)表示。为了保证*k*NN查询的安全，通常采用可恢复距离加密方法(distance-recoverable encryption, DRE)，该方法能通过密文Enc($\boldsymbol{D}$)计算$\boldsymbol{d}_i$和$\boldsymbol{d}_j$间的距离dist($\boldsymbol{d}_i$, $\boldsymbol{d}_j$)。例如，使用密钥$\boldsymbol{K}$分别对数据$\boldsymbol{d}_i$和$\boldsymbol{d}_j$进行加密，得到Enc($\boldsymbol{d}_i$, $\boldsymbol{K}$)、Enc($\boldsymbol{d}_j$, $\boldsymbol{K}$)，DRE可利用Enc($\boldsymbol{d}_i$, $\boldsymbol{K}$)、Enc($\boldsymbol{d}_j$, $\boldsymbol{K}$)计算得到dist($\boldsymbol{d}_i$, $\boldsymbol{d}_j$)。

ASPE的基本思路是通过密文计算得到数据记录与查询条件之间的距离关系。设Enc($\boldsymbol{d}$, $\boldsymbol{K}_d$)表示用密钥$\boldsymbol{K}_d$加密后的采集数据$\boldsymbol{d}$；Enc($\boldsymbol{q}$, $\boldsymbol{K}_q$)表示用密钥$\boldsymbol{K}_q$加密后的查询条件$\boldsymbol{q}$；ASPE须满足条件：

$$\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{q}=\mathrm{Enc}(\boldsymbol{d}_i,\boldsymbol{K}_d)^{\mathrm{T}}\mathrm{Enc}(\boldsymbol{q},\boldsymbol{K}_q)$$

$$\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{d}_j\neq\mathrm{Enc}(\boldsymbol{d}_i,\boldsymbol{K}_d)^{\mathrm{T}}\mathrm{Enc}(\boldsymbol{d}_j,\boldsymbol{K}_d)$$

可见，$\boldsymbol{d}^{\mathrm{T}}\boldsymbol{I}\boldsymbol{q}=(\boldsymbol{d}^{\mathrm{T}}\boldsymbol{M})(\boldsymbol{M}^{-1}\boldsymbol{q})$。为了满足ASPE，加密过程为：$\mathrm{Enc}(\boldsymbol{d}_i,\boldsymbol{K}_d)=\boldsymbol{M}^{\mathrm{T}}\boldsymbol{d}_i$，$\mathrm{Enc}(\boldsymbol{q},\boldsymbol{K}_q)=\boldsymbol{M}^{-1}\boldsymbol{q}$。

## 1.1　ASPE加密原理

假设$\boldsymbol{d}_i$距离$\boldsymbol{q}$比$\boldsymbol{d}_j$更近，则不等式成立：

$$\mathrm{dist}(\boldsymbol{d}_i,\boldsymbol{q})\leqslant\mathrm{dist}(\boldsymbol{d}_j,\boldsymbol{q})$$

$$\sqrt{\|\boldsymbol{d}_i\|^2+2\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{q}+\|\boldsymbol{q}\|^2}\leqslant\sqrt{\|\boldsymbol{d}_j\|^2+2\boldsymbol{d}_j^{\mathrm{T}}\boldsymbol{q}+\|\boldsymbol{q}\|^2}$$

$$\|\boldsymbol{d}_i\|^2-\|\boldsymbol{d}_j\|^2+2(\boldsymbol{d}_j^{\mathrm{T}}-\boldsymbol{d}_i^{\mathrm{T}})\boldsymbol{q}\leqslant 0 \tag{1}$$

可见，ASPE能对$\boldsymbol{d}^{\mathrm{T}}\boldsymbol{q}$的值进行隐私保护。每一个数据$\boldsymbol{d}$对应的$\|\boldsymbol{d}\|$是一个定值，为了方便后续计算，将$\|\boldsymbol{d}\|$与数据$\boldsymbol{d}$一同进行预存储。因此，加密前将数据$\boldsymbol{d}$增加到$u$+1维，用向量$\hat{\boldsymbol{d}}$表示，并将第$u$+1维设为$-0.5\|\boldsymbol{d}\|^2$，即$\hat{\boldsymbol{d}}^{\mathrm{T}}=(\boldsymbol{d}^{\mathrm{T}},-0.5\|\boldsymbol{d}\|^2)$；为了使$\boldsymbol{q}$与$\boldsymbol{d}$在同一维度空间，将$\boldsymbol{q}$也增加到$u$+1维，并将第$u$+1维设为1，即$\hat{\boldsymbol{q}}=(\boldsymbol{q}^{\mathrm{T}},1)$。为了使每个$\boldsymbol{q}$都不可估计，对其乘以随机系数$r$，表示为向量$\hat{\boldsymbol{q}}=r(\boldsymbol{q}^{\mathrm{T}},1)$。

$$\|\boldsymbol{d}\|=\sqrt{(d^1)^2+(d^2)^2+\cdots+(d^u)^2} \tag{2}$$

数据$\boldsymbol{d}_i$和$\boldsymbol{d}_j$增加一维后得到向量$\hat{\boldsymbol{d}}_i,\hat{\boldsymbol{d}}_j$，使用ASPE对向量$\hat{\boldsymbol{d}}$和$\hat{\boldsymbol{q}}$加密，对于$\boldsymbol{d}_i$, $\boldsymbol{d}_j$通过计算$(\mathrm{Enc}(\boldsymbol{d}_i,\boldsymbol{K}_d)^{\mathrm{T}}-\mathrm{Enc}(\boldsymbol{d}_j,\boldsymbol{K}_d)^{\mathrm{T}})\mathrm{Enc}(\boldsymbol{q},\boldsymbol{K}_q)$的值来判断$\boldsymbol{d}_i$和$\boldsymbol{d}_j$哪一个距离查询条件$\boldsymbol{q}$更近。

使用矩阵$\boldsymbol{M}$加密$\boldsymbol{d}_i$和$\boldsymbol{d}_j$，并使用$\boldsymbol{M}^{-1}$加密$\boldsymbol{q}$，$\boldsymbol{M}$是($u$+1)*($u$+1)维的矩阵。可见加密后的的矩阵向量积与加密前相等：

$$(\mathrm{Enc}(\hat{\boldsymbol{d}}_i,\boldsymbol{K}_d)^{\mathrm{T}}-\mathrm{Enc}(\hat{\boldsymbol{d}}_j,\boldsymbol{K}_d)^{\mathrm{T}})\mathrm{Enc}(\hat{\boldsymbol{q}},\boldsymbol{K}_q)=$$

$$(\boldsymbol{M}^{\mathrm{T}}\hat{\boldsymbol{d}}_i-\boldsymbol{M}^{\mathrm{T}}\hat{\boldsymbol{d}}_j)^{\mathrm{T}}\boldsymbol{M}^{-1}\hat{\boldsymbol{q}}=(\hat{\boldsymbol{d}}_i-\hat{\boldsymbol{d}}_j)^{\mathrm{T}}\hat{\boldsymbol{q}} \tag{3}$$

$\hat{\boldsymbol{d}}_i,\hat{\boldsymbol{d}}_j$和$\hat{\boldsymbol{q}}$的相对距离：

$$(\hat{\boldsymbol{d}}_i-\hat{\boldsymbol{d}}_j)^{\mathrm{T}}\hat{\boldsymbol{q}}=$$

$$(\boldsymbol{d}_i-\boldsymbol{d}_j)^{\mathrm{T}}r\boldsymbol{q}+(-0.5\|\boldsymbol{d}_i\|^2+0.5\|\boldsymbol{d}_j\|^2)=$$

$$0.5r((\|\boldsymbol{d}_j\|^2+2\boldsymbol{d}_j\boldsymbol{q}+\boldsymbol{q}^2)-(\|\boldsymbol{d}_i\|^2+2\boldsymbol{d}_i\boldsymbol{q}+\boldsymbol{q}^2))=$$

$$0.5r(\mathrm{dist}(\boldsymbol{d}_j,\boldsymbol{q})^2-\mathrm{dist}(\boldsymbol{d}_i,\boldsymbol{q})^2) \tag{4}$$

可得出：

$$(\hat{\boldsymbol{d}}_i-\hat{\boldsymbol{d}}_j)^{\mathrm{T}}\hat{\boldsymbol{q}}>0\Leftrightarrow\mathrm{dist}(\boldsymbol{d}_i,\boldsymbol{q})<\mathrm{dist}(\boldsymbol{d}_j,\boldsymbol{q}) \tag{5}$$

## 1.2　ASPE加密过程

为了提高ASPE加密过程中数据的随机性，在上述ASPE机制中引入随机划分向量$\boldsymbol{S}$，并根据$\boldsymbol{S}$对$\boldsymbol{d}$和$\boldsymbol{q}$进行划分。$\boldsymbol{S}$是由0和1组成的向量，由数据采集单元与数据用户共享，用于将$\boldsymbol{d}$划分为$\boldsymbol{d}'$和$\boldsymbol{d}''$、将$\boldsymbol{q}$划分为$\boldsymbol{q}'$和$\boldsymbol{q}''$。当$S[i]$=1时，$\boldsymbol{d}[i]=\boldsymbol{d}'[i]+\boldsymbol{d}''[i]$，$\boldsymbol{q}[i]=\boldsymbol{q}'[i]=\boldsymbol{q}''[i]$，其中$\boldsymbol{d}'[i]$随机。当$S[i]$=0时，$\boldsymbol{d}[i]=\boldsymbol{d}'[i]=\boldsymbol{d}''[i]$，$\boldsymbol{q}[i]=\boldsymbol{q}'[i]+\boldsymbol{q}''[i]$，其中$\boldsymbol{q}'[i]$随机。

对划分后的$\boldsymbol{d}$和$\boldsymbol{q}$，使用$\boldsymbol{M}_1$和$\boldsymbol{M}_2$加密。加密过程表示为：$E(\boldsymbol{d}')=\boldsymbol{M}_1^{\mathrm{T}}\boldsymbol{d}'$，$E(\boldsymbol{d}'')=\boldsymbol{M}_2^{\mathrm{T}}\boldsymbol{d}''$，$E(\boldsymbol{q}')=$

$\boldsymbol{M}_1^{-1}\boldsymbol{q}'$，$E(\boldsymbol{q}'')=\boldsymbol{M}_2^{-1}\boldsymbol{q}''$，如式(6)所述，加密后$\boldsymbol{d}^{\mathrm{T}}\boldsymbol{q}$的值不被泄露。

$$E(\boldsymbol{d}')^{\mathrm{T}}E(\boldsymbol{q}')+E(\boldsymbol{d}'')^{\mathrm{T}}E(\boldsymbol{q}'')=$$
$$(\boldsymbol{M}_1^{\mathrm{T}}\boldsymbol{d}')^{\mathrm{T}}\boldsymbol{M}_1^{-1}\boldsymbol{q}'+(\boldsymbol{M}_2^{\mathrm{T}}\boldsymbol{d}'')^{\mathrm{T}}\boldsymbol{M}_2^{-1}\boldsymbol{q}''=$$
$$\boldsymbol{d}'^{\mathrm{T}}\boldsymbol{q}'+\boldsymbol{d}''^{\mathrm{T}}\boldsymbol{q}''=\boldsymbol{d}^{\mathrm{T}}\boldsymbol{q} \tag{6}$$

## 2 基于R树的分桶索引BRtree

显然，随着数据记录的增多ASPE的查询过程非常耗时。为此，本文提出一种基于Rtree构建的桶划分索引——BRtree，该索引能有效提高查询效率。

### 2.1 BRtree的定义

图1为一个Brtree结构，它是一棵二维满二叉树，每个节点对应一个分桶。根节点所在层记为第一层，树的第一层分辨器对应第一个数据维度，即$j$=0 mod 2；树的第二层对应第二个数据维度，即$j$=1 mod 2。非叶子节点对应的桶$B_1$、$B_2$、$B_3$中保存了节点的最小边界值$V_{n_l}$和最大边界值$V_{n_h}$($n$=1, 2, 3)，叶子节点对应的桶$B_4$、$B_5$、$B_6$、$B_7$中不仅保存了节点的最小边界值$V_{n_l}$和最大边界值$V_{n_h}$($n$=4, 5, 6)，还保存了节点的数据集合$D_4, D_5, D_6, D_7$。子节点个数满足$2^{h-1}$，$h$=3。

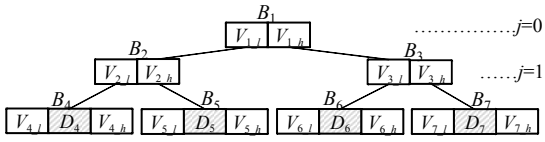


图1 BRtree定义

### 2.2 BRtree创建过程

假设数据采集单元$C$={$C_1$, $C_2$, …, $C_n$}采集的数据集合是$D$={$D_1, D_2$, …, $D_n$}, $D \subseteq D_R$。假设叶子节点对应的分桶大小为$m$，即叶子节点对应分桶中数据集合$D$所包含的数据个数最大为$m$。下面介绍BRtree的创建过程。

1) 根据 $D_R$ 创建 BRtree 索引

创建过程如图2所示，$B_1$表示初始分桶根节点，$B_1$上下界{$V_1$, $V_2$}分别是数据集合$D_R$中的最小边界值和最大边界值。树的第一层分辨器对应第一维数据，记作 $j$=0 mod 2。对分桶$B_1$在第一维数据上进行折半划分，得到分桶$B_2$={$V_1$, $V_3$}和$B_3$={$V_4$, $V_2$}。再将$B_2$和$B_3$分别按照分辨器$j$=1 mod 2进行折半划分，得到分桶$B_4, B_5, B_6, B_7$。每划分一个数据维度，树的高度增加一层。按照上述方法递增分辨器循环每一个数据维度对节点进行折半划分，直到分桶中的数据个数小于或等于$m$。对每一个分桶使用ASPE机制进行加密得到[$B_i$]={$E(V_l)$, $E(V_h)$}，并将加密后的BRtree分桶索引上传到服务器。

2) 将采集数据插入 BRtree 分桶中

假设传感器单元$C$采集的数据集合$D$={$d_1, d_2$, …, $d_n$}。首先，将$D$中的数据逐个放置到满足边界值范围的相应叶子节点分桶中。分桶$B_i$中的数据是{$d_1, d_2, \cdots, d_{m'}$}，$0 \leqslant m' \leqslant m$。为采集数据找到相应的分桶后，对分桶中的数据使用ASPE进行加密，记作$E(B_i)$={$E(d_1), E(d_2)$, …, $E(d_m)$}。将$E(B_i)$发送至服务器并存储在相应的分桶中。因此，叶子节点的分桶$B_i$中除了包含上下边界值[$B_i$]={$E(V_l)$, $E(V_h)$}，还包含了当前分桶中的加密数据$E(B_i)$={$E(d_1)$, $E(d_2), \cdots, E(d_m)$}。所有的叶子节点构成的桶集合表示为：$B_{\mathrm{leaf}}$={$B_{l\_1}, B_{l\_2}, \cdots, B_{l\_t}$}，叶子节点$B_{l_i}$($1 \leqslant i \leqslant t$)中的数据为$B_{l_i}$={$d_1, d_2$, …, $d_m$}；所有的非叶子节点构成的桶表示为：$B_{\mathrm{node}}$={$B_{n\_1}, B_{n\_2}, \cdots, B_{n\_z}$}；创建的BRtree共$h$层，第$c$ ($1 \leqslant c \leqslant h-1$) 层包含$y$个桶节点，$c$层构成的桶集合为$B_{l=c}$={$B_1$, $B_2, \cdots, B_y$}。然后，数据采集单元将创建好的BRtree索引$B$={[$B_{n\_1}$], [$B_{n\_2}$], …, [$B_{n\_z}$], [$B_{l\_1}$], [$B_{l\_2}$], …, [$B_{l\_t}$]}和$E(B)$={$E(B_{l\_1})$, $E(B_{l\_2})$, …, $E(B_{l\_t})$}发送至服务器。

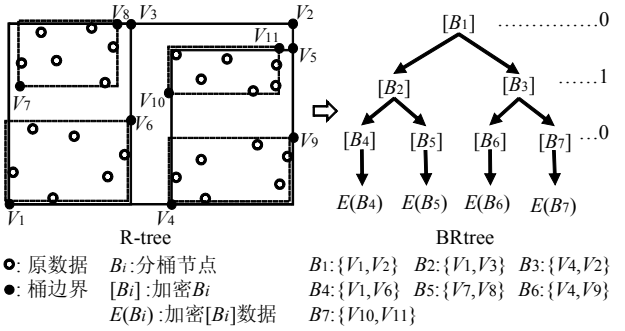


图2 BRtree索引创建过程

### 2.3 基于BRtree索引的剪枝查询

执行一次查询$\boldsymbol{q}$时，首先，从根节点开始递归BRtree的每一层，即$c$=0, 1, …, $h$–1。计算查询$q$到$B_{l=c}$={$B_1$, $B_2, \cdots, B_y$}中每一个非叶子节点$B_x$($B_x \in B_{l=c}$)的边界距离参数：最小边界距离$\mathrm{dist}_{\min}(B_x, \boldsymbol{q})$、最大边界距离$\mathrm{dist}_{\max}(B_x, \boldsymbol{q})$和最小最大边界距离$\mathrm{dist}_{\min_\max}(B_{l=c}, \boldsymbol{q})$。其中，最小边界距离$\mathrm{dist}_{\min}(B_x, \boldsymbol{q})$是查询条件$q$距离当前节点上界和下界的最小值，即$\mathrm{dist}_{\min}(B_i, \boldsymbol{q})=\min\{\mathrm{dist}(B_{i_\mathrm{low}}, \boldsymbol{q}), \mathrm{dist}(B_{i_\mathrm{high}}, \boldsymbol{q})\}$；同理，最大边界距离$\mathrm{dist}_{\max}(B_i, \boldsymbol{q})=\max\{\mathrm{dist}(B_{i_\mathrm{low}}, \boldsymbol{q}), \mathrm{dist}(B_{i_\mathrm{high}}, \boldsymbol{q})\}$；此外，还需计算查询条件距离当前层所有节点的最小最大边界距离$\mathrm{dist}_{\min_\max}(B_{l=c}, \boldsymbol{q})$，表示该层节点中距离查询条件$q$的最大边界距离中的最小值，即$\mathrm{dist}_{\min_\max}(B_{l=c}, \boldsymbol{q}) = \min\{\mathrm{dist}_{\max}(B_1, \boldsymbol{q}),$

$\text{dist}_{\max}(B_2, \boldsymbol{q}),\cdots, \text{dist}_{\max}(B_y, \boldsymbol{q})\}$。

然后，根据计算得到的边界距离参数进行剪枝。当且仅当$\text{dist}_{\min}(B_x, \boldsymbol{q})>\text{dist}_{\min_\max}(B_{l=c}, \boldsymbol{q})$时，$B_x$节点所在的分枝将被剪去。图3为一个BRtree分枝节点，该节点包含的两个非叶子的桶节点分别是$B_a$和$B_{a+1}$，它们所在的层是$c$。当执行查询$\boldsymbol{q}$时，计算得$\boldsymbol{q}$到$B_a$的最大距离是该层的最小最大距离，即$\text{dist}_{\min_\max}(B_{l=c}, \boldsymbol{q})=\text{dist}_{\max}(B_a, \boldsymbol{q})$。被查询的当前桶是$B_{a+1}$，计算$\boldsymbol{q}$到$B_{a+1}$的最小距离大于该层节点的最小最大距离：$\text{dist}_{\min}(B_{a+1}, \boldsymbol{q})>\text{dist}_{\min_\max}(B_{l=c}, \boldsymbol{q})$，根据上述剪枝条件，将删除桶节点$B_{a+1}$及其孩子节点。

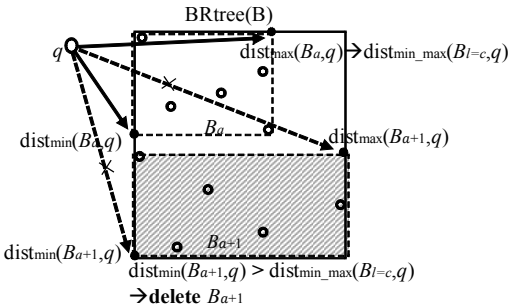


图3　剪枝过程

# 3　访问权限控制

为了防止数据用户滥用病人医疗数据导致隐私信息的泄漏，本文引入可信第三方实体，实现对病人信息的访问权限控制与迁移。

## 3.1　访问权限迁移

可信第三方实体根据病人密钥和其责任医生密钥，生成该病人数据的访问权限密钥，并将密钥分配给服务器。服务器只有从可信第三方获得正确权限密钥，才能访问病人的医疗数据，并为授权用户提供正确的查询结果。一旦用户的责任医生发生更换，原来的权限密钥将会失效，服务器须从可信第三方获取新的权限密钥，同时新责任医生被授权。

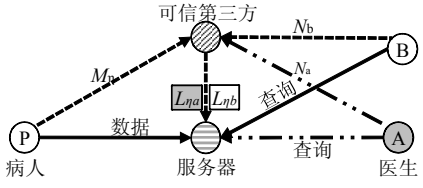


图4　访问权限控制

访问权限控制如图4所示，病人$P$持有密钥$M_\eta$，其责任医生$A$所持密钥$N_a$，此时可信第三方根据病人$P$的密钥和其责任医生$A$的密钥生成权限密钥$L_{\eta a}$。当病人$P$更换为持有密钥$N_b$的责任医生$B$后，可信第三方生成病人$P$和新责任医生$B$的权限密钥$L_{\eta b}$，并将其发送至服务器。此时，医生$B$被授予访问病人$P$的医疗数据权限，同时医生$B$能够通过服务器查询到正确的结果，医生$A$的访问权限$L_{\eta a}$失效，访问权限被撤销，无法查询到正确的结果。

## 3.2　基于ASPE的权限控制数据交换过程

本文在ASPE机制中引入访问权限控制过程，其数据的流动过程如图5所示。数据采集单元$C_\eta$的密钥是$\{\boldsymbol{M}_{\eta1},\boldsymbol{M}_{\eta2}\}$，$C_\eta$采集的数据向量是$\boldsymbol{d}_\eta$，对划分后的数据向量$\boldsymbol{d}_\eta',\boldsymbol{d}_\eta''$通过ASPE进行加密，得到$E(\boldsymbol{d}_\eta)=\{E(\boldsymbol{d}_\eta'),E(\boldsymbol{d}_\eta'')\}$，其中，$E(\boldsymbol{d}_\eta')=\boldsymbol{M}_{\eta1}^{\mathrm{T}}\boldsymbol{d}_\eta'$，$E(\boldsymbol{d}_\eta'')=\boldsymbol{M}_{\eta2}^{\mathrm{T}}\boldsymbol{d}_\eta''$将其发送至服务器。

服务器接收到数据采集单元$C_\eta$发送的加密数据$\mathrm{E}(\boldsymbol{d}_\eta)$后，使用密钥$\boldsymbol{L}_{\eta1}$和$\boldsymbol{L}_{\eta2}$分别对$E(\boldsymbol{d}_\eta')$和$E(\boldsymbol{d}_\eta'')$进行加密，其中，$\boldsymbol{L}_{\eta1}=\boldsymbol{M}_{\eta1}^{-1}\boldsymbol{N}_1$，$\boldsymbol{L}_{\eta2}=\boldsymbol{M}_{\eta2}^{-1}\boldsymbol{N}_2$，加密后表示为$\{\delta(E(\boldsymbol{d}_\eta')),\delta(E(\boldsymbol{d}_\eta''))\}$，其中$\delta(E(\boldsymbol{d}_\eta'))=E(\boldsymbol{d}_\eta')\boldsymbol{L}_{\eta1}$，$\delta(E(\boldsymbol{d}_\eta''))=E(\boldsymbol{d}_\eta'')\boldsymbol{L}_{\eta2}$。

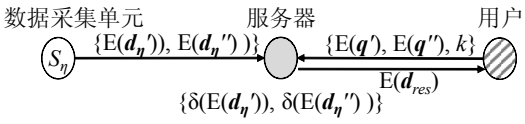


图5　基于ASPE的数据交换过程

授权用户执行一次查询$\boldsymbol{q}$时，用户密钥为$\{\boldsymbol{N}_1,\boldsymbol{N}_2\}$。用户使用密钥$\boldsymbol{N}_1$和$\boldsymbol{N}_2$对划分后的查询向量$\boldsymbol{q}',\boldsymbol{q}''$分别进行加密，得到$E(\boldsymbol{q}')=\boldsymbol{N}_1^{-1}\boldsymbol{q}'$，$E(\boldsymbol{q}'')=\boldsymbol{N}_2^{-1}\boldsymbol{q}''$。用户将加密后的查询条件$\{E(\boldsymbol{q}'),E(\boldsymbol{q}''),k\}$一并发送至服务器。

服务器根据已有的加密数据，对授权用户发送的查询条件进行计算，得到满足条件的$k$个查询结果$E(\boldsymbol{d}_{\text{res}})=\{E(\boldsymbol{d}_i), E(\boldsymbol{d}_{i+1}),\cdots, E(\boldsymbol{d}_{i+k})\}$, $E(\boldsymbol{d}_{\text{res}})\in E(\boldsymbol{d}_\eta)$，将结果$E(\boldsymbol{d}_{\text{res}})$发送给授权用户，用户使用密钥$\{\boldsymbol{N}_1, \boldsymbol{N}_2\}$解密得到明文信息。

# 4　安全性分析

## 4.1　复杂度分析

假定某传感单元采集了$n$个$u$维数据， BRtree剪枝后剩余计数$\alpha$。表1给出了本文协议在最坏情况下的计算复杂度、通信开销，及空间复杂度。

**表1　本文协议的算法复杂度分析**

| | 计算复杂度 | 通信开销 | 空间复杂度 |
|---|---|---|---|
| 数据采集单元 | $O(n\times u^2)$ 查询 | $O(n\times u)$ | |
| 存储节点 | $O(\alpha\times u^2)$BRtree查询 | $O(n\times u)$ | $O(nu)$ |
| 数据用户 | $O(u^2)$ 查询条件处理 | $O(u)$ | |

### 4.2 隐私性分析

本文提出的协议能够有效保护数据隐私。1) 假设存储节点被妥协，攻击者没有加密密钥，仅通过存储节点存储的密文数据来计算明文数据是十分困难的。2) 无线体域网络中数据维数大，未知量计数远超已知密文信息的维度。因此，当暴露部分明文数据记录、密文索引信息时，攻击者仍无法计算出密钥，也无法从密文中获取有用的明文信息。

## 5 实 验

实验数据来自移动健康数据集，采集了10名志愿者在进行不同体育活动时的生命体征数据记录。数据集共包含10个传感器单元采集的23维属性数据文件，每个文件数据量在10～15万，将数据划分为不重叠的周期$T$，每10分钟为一个周期。

图6为创建索引时间开销随数据维度$u$的变化曲线。其中$m$=200，$k$=50。随$u$由13～23维递增，ASPE索引创建时间从0.41～0.88 s递增，BRtree索引创建时间从0.49～0.97 s递增，BRtree的平均创建索引时间是ASPE的1.15倍。

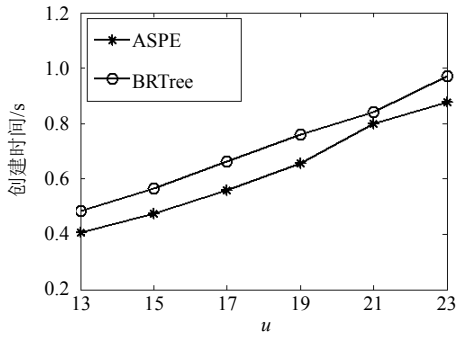


图6 $u$-创建时间开销

图7为创建索引时间开销随时间周期$T$的变化曲线。其中$u$=23, $m$=200, $k$=50。BRtree的创建索引时间从0.21～2.18 s递增，ASPE的创建索引时间从0.17～1.73 s递增，BRtree的平均创建索引时间是ASPE的1.14倍。

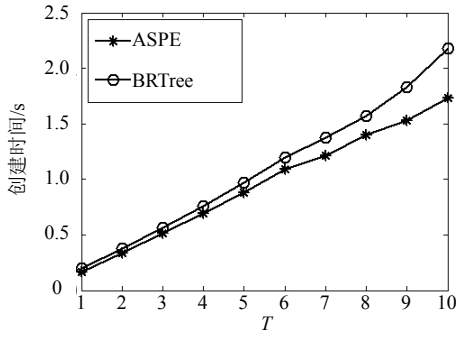


图7 $T$-创建时间开销

图8为查询时间开销随数据维度$u$的变化曲线。其中$m$＝200，$k$＝50。ASPE查询过程各周期查询时间从7.60～10.77 s递增，BRtree查询过程各周期查询时间从5.30～6.20 s变化。BRtree查询时间比ASPE查询时间加快了0.47倍。

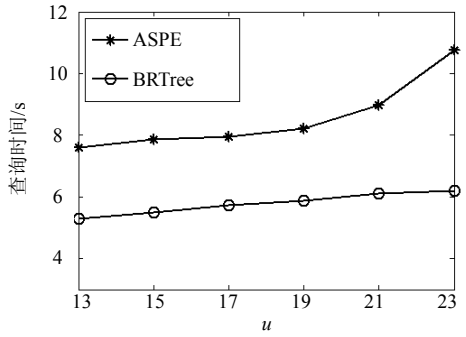


图8 $u$-查询时间开销

图9为查询时间开销随时间周期的变化曲线。其中$u$=23, $m$=200, $k$=50。ASPE各周期查询时间从2.01～22.62 s递增，BRtree查询时间从0.98～11.91 s递增。BRtree查询查询时间比ASPE加快了0.48倍。

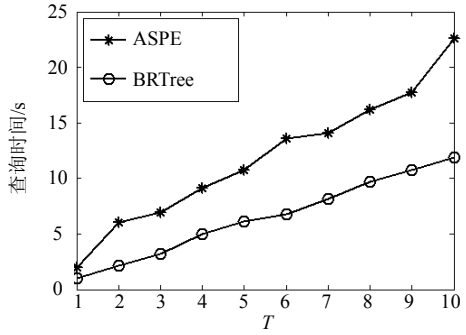


图9 $T$-查询时间开销

图10为查询时间开销随时间周期的变化曲线，其中$u$=23, $k$=30。图中4条曲线分别是ASPE和BRtree索引分桶大小为$m$=100, 200, 300时的查询时间曲线。$m$=200比$m$=100查询时间加快了0.21倍，$m$=300比$m$=100查询时间加快了0.48倍。

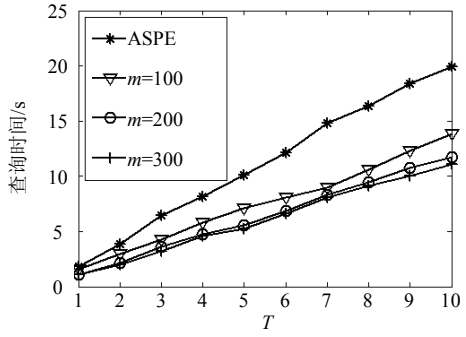


图10 $T$-分桶查询时间开销

图11为查询时间随$k$的变化曲线，其中$u$=23,

*m*=200。图中3条曲线分别表示TimeSlot=1, 5, 10时查询时间的变化曲线。可见，*k*增加对查询时间的影响非常小。TimeSlot=1, 5, 10的平均查询时间分别是1.03, 5.71, 11.87 s。

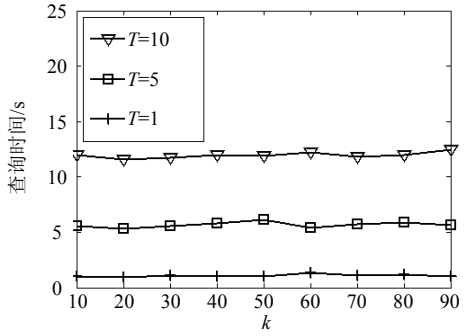


图11　*k*-查询时间开销

## 6　结 束 语

在无线体域网中实现了隐私保护安全的*k*NN查询协议。通过ASPE机制，在服务器既不知道数据真实值，也不知道查询条件的情况下实现了安全的*k*NN查询机制，该机制能够抵抗已知部分明文攻击。此外，通过基于R树的桶划分索引BRtree，采用剪枝策略有效地提高了*k*NN的查询效率。最后，根据ASPE机制特征，从ASPE加密矩阵中分解出权限矩阵，通过引入可信第三方，解决了无线体域网中数据级访问权限控制以及访问权限的迁移问题。

## 参 考 文 献

[1] MOVASSAGHI S, ABOLHASAN M, LIPMAN J, et al. Wireless body area networks: a survey[J]. IEEE Communications Survey&Tutorials, 2014, 16(3): 1658-1686.

[2] TEC. Behind the medical data leak who moved the patient's cheese[EB/OL]. [2016-10-25]. http://www.ip-guard.net/blog/? p=1664.

[3] SHI L, LI M, YU S, et al. Bana: Body area network authentication exploiting channel characteristics[J]. IEEE Journal on Selected Areas in Communications, 2013, 31(9): 1803-1816.

[4] LIU J, ZHANG Z, CHEN X, et al. Certificateless remote Anonymous authentication schemes for wireless body area networks[J]. IEEE Transactions on Parallel and Distributed Systems, 2014, 25(2): 332-342.

[5] ZHAO Z. An efficient anonymous authentication scheme for wireless body area networks using elliptic curve cryptosystem[J]. Journal of Medical Systems, 2014, 38(2): 1-7.

[6] 韩坚华, 吴柳飞. 无线传感器网络EMSR协议的安全性分析[J]. 电子科技大学学报, 2009, 38(3): 401-405.
HAN Jian-hua, WU Liu-fei. Analysis on security of EMSR protocol in wireless sensor network[J]. Journal of University of Electronic Science and Technology of China, 2009, 38(3): 401-405.

[7] 贾晨军, 廖永建, 陈抗生. 无线传感器网络中的高效签名算法[J]. 电子科技大学学报, 2009, 38(4): 537-541.
JIA Chen-jun, LIAO Yong-jian, CHEN Kang-sheng. Efficient signature algorithm in wireless sensor network[J]. Journal of University of Electronic Science and Technology of China, 2009, 38(4): 537-541.

[8] 汪小芬, 李胜强, 肖国镇. 认证群密钥协商协议的安全性分析与改进[J]. 电子科技大学学报, 2009, 38(1): 51-54.
WANG Xiao-fen, LI Sheng-qiang, XIAO Guo-zhen. Analysis and improvement of an authenticated group key agreement protocol[J]. Journal of University of Electronic Science and Technology of China, 2009, 38(1): 51-54.

[9] YI Ye-qing, LI Rui, CHEN Fei, et al. A digital watermarking approach to secureand precise range query processing in sensor networks[C]//IEEE Conference on Computer Communications(INFOCOM 2013). Turin, Italy: IEEE, 2013.

[10] ZHANG Rui, SHI Jing, ZHANG Yan-chao, et al. Secure top-k query processing in unattended tiered sensor networks[J]. IEEE Transactions on Vehicular Technology (TVT), 2014, 9(63): 4681-4693.

[11] 范永健, 陈红, 张晓莹, 等. 两层传感器网络中可验证隐私保护的 top-k 查询协议[J]. 计算机学报, 2014, 37(4): 915-926.
FAN Yong-jian, CHEN Hong, ZHANG Xiao-ying, et al. Verifiable privacy-preserving top-k query protocol in two-tiered sensor networks[J]. Chinese Journal of Computers, 2014, 37(4): 915-926.

[12] LIAO X, LI J. Privacy-preserving and secure top-k query in two-tiered wireless sensor[C]//Proceedings of IEEE Global Communications Conference(GLOBECOM). [S.l.]: IEEE, 2012: 335-341.

[13] 李睿, 林亚平, 易叶青, 等. 两层传感器网络中安全Top-k查询协议[J]. 计算机研究与发展, 2012, 49(9): 1947-1958.
LI Rui, LIN Ya-ping, YI Ye-qing, et al. A secure top-k query protocol in two-tiered sensor networks[J]. Computer Research and Development, 2012, 49(9): 1947-1958.

[14] WONG W K, CHEUNG D W, KAO B, et al. Secure *k*NN computation on encrypted database[C]//the 2009 ACM SIGMOD International Conference on Management of Data (SIGMOD2009). New York: ACM, 2009: 139-152.

[15] YUAN Jia-wei, YU Shu-cheng. Efficient privacy-preserving biometric identification in cloud computing[C]// IEEE Conference on Computer Communications. Turin, Italy: IEEE, 2013: 2652-2660.

[16] CAO Ning, WANG Cong, LI Ming, et al. Privacy-preserving multi-keyword ranked search over encrypted cloud data[C]//2011 IEEE Conference on Computer Communications (INFOCOM2011). [S.l.]: IEEE, 2011: 829-837.

编　辑　叶　芳