(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 115168909 A
(43)申请公布日 2022.10.11

(21)申请号 202211086567.6

(22)申请日 2022.09.07

(71)申请人 翼方健数(北京)信息科技有限公司
地址 100000 北京市海淀区阜成路73号A座五层507,508,509,510,511,512号
申请人 翼健(上海)信息科技有限公司

(72)发明人 张李军 潘光明 张浩

(74)专利代理机构 北京沃杰永益知识产权代理事务所(普通合伙) 11905
专利代理师 杨杰

(51)Int.Cl.
G06F 21/62(2013.01)
G06F 21/60(2013.01)

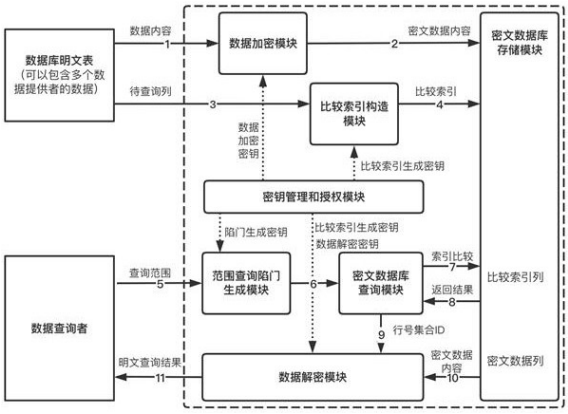
权利要求书4页 说明书15页 附图3页

(54)发明名称

一种基于比较索引的密文数据范围查询方法和系统

(57)摘要

本发明提供一种基于比较索引的密文数据范围查询方法和系统,从数据提供者提供的数据及待查询列数据开始处理,首先生成密文数据内容和待查询列的比较索引,存储在密文数据库存储模块中;然后数据查询者根据查询范围发起查询,范围查询陷门生成模块产生出查询陷门,密文数据库查询模块利用查询陷门进行密文范围查询获得数据库表行号集合ID;数据解密模块解密ID中对应行所需的密文数据列,最终将明文结果返回给数据查询者。本发明的查询算法高效并能并行化地执行,算法泄露的明文信息比特数少,适用于大规模数据量下的范围查询业务场景;同时基于按比特的索引大小比较算法能够保证范围查询结果的准确性。



CN 115168909 A

1.一种基于比较索引的密文数据范围查询方法,其特征在于,所述方法包括:

获取填有多条明文数据内容的数据库明文表以及数据加密密钥,采用数据加密密钥并利用第一算法对数据库明文表进行加密,得到对应的密文数据库表;

从数据库明文表中获取待查询列,利用第二算法对数据库明文表的待查询列生成比较索引,将生成的比较索引加入所述密文数据库表的比较索引列;

获取数据查询者提供的查询范围,基于查询范围并通过第三算法计算生成范围查询陷门;

基于范围查询陷门,并通过第四算法查询所述密文数据库表中的比较索引列,得到查询结果;

基于查询结果获取对应的数据解密密钥,并从密文数据库表中获取对应的密文数据内容,基于数据解密密钥并通过第五算法对密文数据内容进行解密,输出明文查询结果。

2.根据权利要求1所述的一种基于比较索引的密文数据范围查询方法,其特征在于,获取填有多条明文数据内容的数据库明文表以及数据加密密钥,采用数据加密密钥并利用第一算法对数据库明文表进行加密,得到对应的密文数据库表,具体包括:

预设数据库明文表包括由多条明文数据内容,每条明文数据内容包括多个明文列项数据;

基于不同的明文数据内容分别生成对应的数据加密密钥;

采用数据加密密钥并按照第一算法对相应明文数据内容中的明文列项数据进行加密得到对应的密文列项数据,将密文列项数据置于密文数据库表的密文列中,同时将数据加密密钥与明文数据内容在数据库明文表中的位置标识信息进行关联存储在密钥库中;

其中,所述第一算法的执行过程为:

预设数据库明文表中的明文列项数据为M,数据加密密钥为K;

根据算法C=Enc(M,K),计算出对应的密文列项数据C,其中Enc为对称加密算法,优选为AES或国密SM4算法。

3.根据权利要求1所述的一种基于比较索引的密文数据范围查询方法,其特征在于,从数据库明文表中获取待查询列,利用第二算法对数据库明文表的待查询列生成比较索引,具体包括:

预设数据库明文表中待查询列的明文列项数据为{M_1,M_2,…, M_n};

预设安全参数为λ_1,随机选择λ_1比特长度的索引生成密钥k_1;

设伪随机函数$F:\{0,1\}^{\lambda_1}\times\{0,1\}^{m-1}\to\mathbb{Z}_3$,其中$\{0,1\}^{\lambda_1}$和$\{0,1\}^{m-1}$分别表示 λ_1比特和m-1比特长度的二进制字符串, $\mathbb{Z}_3$表示模3的整数剩余类环;

对每个明文列项数据M,M取自于集合{M_1,M_2,…, M_n},设M的二进制表示为$b_1b_2\cdots b_m$,令$[m]$表示整数集合{1, 2, …,m},对 $i\in[m]$,计算:

$u_i=F(k_1,b_1b_2\cdots b_{i-1}\,||\,0^{m-i})+b_i(\mathrm{mod}\ 3)$,符号 || 表示字符串级联, mod表示取模运算符;

计算出明文列项数据M的比较索引为$u=(u_1,u_2,\cdots,u_m)$;

输出索引生成密钥k和待查询列的所有明文列项数据的比较索引,并基于所有明文列项数据的比较索引形成比较索引列。

4.根据权利要求1所述的一种基于比较索引的密文数据范围查询方法,其特征在于,获

取数据查询者提供的查询范围，基于查询范围并通过第三算法计算生成范围查询陷门，具体包括：

预设安全参数为λ_2，随机选择λ_2比特长度的陷门生成密钥k_2；

获取查询请求信息$(m_1,m_2,token)$，所述查询请求信息包括查询范围和token信息，m_1,m_2表示查询范围的上下区间值，$token$本次数据查询者的请求令牌；

选择伪随机函数$R:\{0,1\}^{\lambda_2}\times T\to\mathbb{Z}_3^m$，其中$T$表示token的取值空间，$\mathbb{Z}_3^m$表示m维的$\mathbb{Z}_3$向量，$\mathbb{Z}_3$表示模3的整数剩余类环；

根据查询请求信息$(m_1,m_2,token)$，利用第三算法计算：

$c_1=TrapGen(m_1)$，$c_2=TrapGen(m_2)$；$TrapGen$表示查询陷门生成算法；

$h=R(k_2,token)$；

$c_{1t}=c_1\boxplus h$，$c_{2t}=c_2\boxplus h$，其中符号$\boxplus$表示m维向量按分量模3相加；

输出范围查询陷门$T=(c_{1t},c_{2t},h)$。

5.根据权利要求4所述的一种基于比较索引的密文数据范围查询方法，其特征在于，基于范围查询陷门，并通过第四算法查询所述密文数据库表中的比较索引列，得到查询结果，具体包括：

对两个索引值c和c'，定义索引比较函数$Compare(c,c')$：

设$c=(u_1,u_2,\cdots,u_m)$，$c'=(u'_1,u'_2,\cdots,u'_m)$，如果$c=c'$，输出0；

否则令i是$u_i\neq u'_i$的最小正整数，若满足$u'_i=u_i+1\ (mod\ 3)$，则输出-1

若满足$u_i=u'_i+1\ (mod\ 3)$，则输出1；

比较函数输出1表示c对应的明文数据值大于c'对应的明文数据值，-1表示c对应的明文数据值小于c'对应的明文数据值，0则表示两者相等；

对查询陷门$T=(c_{1t},c_{2t},h)$，初始化集合ID为空，进行下面的计算：

还原范围查询的端点值$c_1=c_{1t}\boxminus h$和$c_2=c_{2t}\boxminus h$，符号$\boxminus$表示m维向量按分量模3相减；

对密文数据库表中查询列的每个索引值u，利用比较函数计算函数值：

$r_1=Compare(u,c_1)$，$r_2=Compare(u,c_2)$；

如果$r_1\neq -1$且$r_2\neq 1$，则表示u对应的明文m在查询的范围$[m_1,m_2]$内；

将满足条件的u所在的行号id添加到集合ID中；

输出集合ID作为查询结果。

6.根据权利要求5所述的一种基于比较索引的密文数据范围查询方法，其特征在于，基于查询结果获取对应的数据解密密钥，并从密文数据库表中获取对应的密文数据内容，基于数据解密密钥并通过第五算法对密文数据内容进行解密，输出明文查询结果，具体包括：

验证查询是否有效，计算$R(k_2,token)$，如果等于h，表示查询有效，进行后续解密，否则输出“查询无效”，算法退出；

预设需要输出的明文数据集合为PlainData，初始化PlainData为空，对每个$id\in ID$，执行：

获取id对应行的数据解密密钥DK，根据实际需要解密这一行中某一列的密文数据，得

到明文数据Plain=Dec(Cipher,DK),其中Cipher表示需要的那一列的密文数据内容,Dec表示对称解密算法;

将明文数据Plain添加到明文数据集合PlainData中;

输出明文数据集合PlainData。

7.根据权利要求1所述的一种基于比较索引的密文数据范围查询方法,其特征在于,在将生成的比较索引加入所述密文数据库表的比较索引列之后,所述方法还包括:

预设在预设时间段内有多个数据查询者提出查询请求信息,每个数据查询者的查询请求信息中包括各自的查询范围;

对多个查询范围进行取交集处理;

预设待查询列的数据依据对应的刻度线取值,统计在每个刻度参与交集的数据查询者数量;

将参与交集的数据查询者数量超过第一预设阈值的刻度判定为合并查询刻度;

将所有合并查询刻度进行并集处理,得到合并查询刻度的并集范围;

基于合并查询刻度的并集范围进行查询,得到合并明文查询结果;

基于每个数据查询者,判断其查询范围是否包含合并查询刻度,如果包括,则从合并明文查询结果中提取所有包含的合并查询刻度对应的明文查询结果,对所有包含的合并查询刻度以外的其它查询刻度,则执行第四算法查询过程;如果不包括,则基于对应数据查询者的查询范围直接执行第四算法查询过程。

8.一种基于比较索引的密文数据范围查询系统,用于实现上述权利要求1至7任意一项所述的基于比较索引的密文数据范围查询方法,其特征在于,所述系统包括:

密钥管理和授权模块,负责生成和管理数据加解密密钥和比较索引生成密钥,在数据查询者授权的前提下,提供比较索引生成密钥和数据解密密钥;

数据加密模块,负责从密钥管理和授权模块获取数据加密密钥,运行第一算法对数据库明文表中的明文数据内容进行加密,产生密文数据内容;

比较索引构造模块,负责从密钥管理和授权模块获取比较索引生成密钥,运行第二算法对数据库明文表中的待查询的明文列项数据生成比较索引;

范围查询陷门生成模块,负责接收数据查询者的查询范围,从密钥管理和授权模块获取陷门生成密钥,运行第三算法生成对应的范围查询陷门;

密文数据库查询模块,负责从范围查询陷门生成模块接收查询陷门,运行第四算法从密文数据库表中比较索引列进行查询,获得查询结果;

数据解密模块,负责从密文数据库查询模块接收查询结果,从密钥管理和授权模块获取比较索引生成密钥和数据解密密钥,运行第五算法验证本次查询的有效性,若有效则解密出明文数据内容,返回给数据查询者;

密文数据库存储模块,负责统一存储密文数据库表,供密文数据库查询模块进行查询以及为数据解密模块提供密文数据内容。

9.根据权利要求8所述的一种基于比较索引的密文数据范围查询系统,其特征在于,数据加密模块还执行以下步骤:

预设数据库明文表包括由多条明文数据内容,每条明文数据内容包括多个明文列项数据;

基于不同的明文数据内容分别生成对应的数据加密密钥；

采用数据加密密钥并按照第一算法对相应明文数据内容中的明文列项数据进行加密得到对应的密文列项数据，将密文列项数据置于密文数据库表的密文列中，同时将数据加密密钥与明文数据内容在数据库明文表中的位置标识信息进行关联存储在密钥库中；

其中，所述第一算法的执行过程为：

预设数据库明文表中的明文列项数据为M，数据加密密钥为K；

根据算法C=Enc(M,K)，计算出对应的密文列项数据C，其中Enc为对称加密算法，优选为AES或国密SM4算法。

10.根据权利要求8所述的一种基于比较索引的密文数据范围查询系统，其特征在于，比较索引构造模块还执行以下步骤：

预设数据库明文表中待查询列的明文列项数据为{M_1,M_2,…, M_n}；

预设安全参数为λ_1，随机选择λ_1比特长度的索引生成密钥k_1；

设伪随机函数$F:\{0,1\}^{\lambda_1}\times\{0,1\}^{m-1}\to\mathbb{Z}_3$，其中$\{0,1\}^{\lambda_1}$和$\{0,1\}^{m-1}$分别表示 λ_1比特和m-1比特长度的二进制字符串，$\mathbb{Z}_3$表示模3的整数剩余类环；

对每个明文列项数据M，M取自于集合{M_1,M_2,…, M_n}，设M的二进制表示为 $b_1b_2\cdots b_m$，令$[m]$表示整数集合{1, 2, …,m}，对$i\in[m]$，计算：

$u_i=F(k_1,b_1b_2\cdots b_{i-1}\,||\,0^{m-i})+b_i(\mathrm{mod}\ 3)$，符号||表示字符串级联，mod表示取模运算符；

计算出明文列项数据M的比较索引为$u=(u_1,u_2,\cdots,u_m)$；

输出索引生成密钥k和待查询列的所有明文列项数据的比较索引，并基于所有明文列项数据的比较索引形成比较索引列。

一种基于比较索引的密文数据范围查询方法和系统

技术领域

[0001] 本发明涉及数据安全存储与共享技术领域，尤其涉及一种基于比较索引的密文数据范围查询方法和系统。

背景技术

[0002] 目前许多云服务厂商为企业和个人提供了方便和实惠的数据外包存储服务，由于网络攻击导致数据泄露的事件时有发生，企业和个人更倾向于将自己的数据进行加密后存储到云服务器中。虽然加密给数据存储带来了安全保障，但同时也给数据的使用带来了困难。比如想要在这些密文数据上进行查询操作，这对于常规加密算法生成的密文数据是难以实现的。针对关键字查询，人们发明了可搜索加密这样的特殊加密算法，能够在密文数据中搜索出含有该关键字的数据。实际应用场景中，除了关键字查询，还有一类重要的查询就是范围查询。查询者通过提交查询范围的起点和终点，期望在密文数据库中搜索出满足该查询范围的数据。这其中的关键就是设计出高效和安全的密文数据范围查询算法。

[0003] 要在密文数据上进行数据的范围查询就要求能根据密文来比较出对应明文数据的大小，学者们提出了几类解决方案，如Stefanov等人提出的基于不经意随机访问内存方案，Gentry等人提出的全同态加密方案以及Agrawal等人提出的保序加密方案。目前最有效的一类方案是保序加密，即密文序列保持着相应明文序列的顺序，云服务器通过比较密文和查询者提供的查询范围的陷门就可以知道密文是否在要求的查询范围之内。我们可以将保序加密细分为三种：一种是直接的保序加密，Agrawal等人于2004年提出的方案就属于这一种，密文是直接根据明文的顺序来进行排序的，没有额外的索引结构，这种方案的安全性没有严格的定义和证明。Hore等人利用分桶的思想于2004年提出带索引结构的保序加密，Liu等人于2012年提出了基于线性索引结构的保序加密。Roche于2016年提出了基于部分保序编码的树型结构的保序加密。这些带索引结构的保序加密方案中，对应明文的大小关系可以通过索引的大小来得到，而明文内容的加密可以采用独立的加密算法来实现。这类型的保序加密安全性不高，而且基于分桶思想的方案返回的查询结果有一定的误差(因为同一个桶里的元素无法比较大小)。第三种保序加密同样基于索引结构，称为揭序加密，最早于2015年由Boneh等人提出。这种加密方式的密文具有一种能揭示明文大小的索引结构。这种索引并不能直接展示出大小关系，而是通过一个特殊设计的比较函数来计算出两个索引的大小。2016年，Chenette和Cash等人对揭序加密进行了效率上的改进，并且通过定义泄露函数来严格证明方案的安全性。

[0004] 上述现有方案存在以下问题：

1、密文索引尺寸扩展较大，方案需要多轮的交互。目前揭序加密方案虽然提高了保序加密的算法安全性(服务器只能知晓密文大小的比较结果)，但比较索引的尺寸扩展过大(达到明文比特长度的100倍以上)，并且方案需要多轮交互才能比较出两个密文值的大小，影响了方案的效率和实用性。

[0005] 2、范围查询结果的准确率低，比如效率较高的基于分桶思想和矩阵密钥的保序加

密方法,它们查询结果的准确性却比较低,误差甚至达到了40%以上,难以满足实际场景中的准确度需求。

[0006] 3、保序加密的安全性不足,泄露的明文比特较多或者攻击者能够从范围查询陷门中恢复出陷门生成密钥。另外,由于保序加密的很多方案并没有严格定义出明文数据的泄露信息,导致无法准确衡量方案的安全性。

[0007] 4、相同查询范围生成的查询陷门相同,不具备变化性。现有技术方案中范围查询陷门的生成是确定性的算法,攻击者可以利用这种确定性来发起重放攻击等非法查询。

发明内容

[0008] 为了解决上述至少一个技术问题,本发明提出了一种高效安全的基于比较索引的密文数据范围查询方法和系统。

[0009] 本发明第一方面提出了一种基于比较索引的密文数据范围查询方法,所述方法包括:

获取填有多条明文数据内容的数据库明文表以及数据加密密钥,采用数据加密密钥并利用第一算法对数据库明文表进行加密,得到对应的密文数据库表;

从数据库明文表中获取待查询列,利用第二算法对数据库明文表的待查询列生成比较索引,将生成的比较索引加入所述密文数据库表的比较索引列;

获取数据查询者提供的查询范围,基于查询范围并通过第三算法计算生成范围查询陷门;

基于范围查询陷门,并通过第四算法查询所述密文数据库表中的比较索引列,得到查询结果;

基于查询结果获取对应的数据解密密钥,并从密文数据库表中获取对应的密文数据内容,基于数据解密密钥并通过第五算法对密文数据内容进行解密,输出明文查询结果。

[0010] 本方案中,获取填有多条明文数据内容的数据库明文表以及数据加密密钥,采用数据加密密钥并利用第一算法对数据库明文表进行加密,得到对应的密文数据库表,具体包括:

预设数据库明文表包括由多条明文数据内容,每条明文数据内容包括多个明文列项数据;

基于不同的明文数据内容分别生成对应的数据加密密钥;

采用数据加密密钥并按照第一算法对相应明文数据内容中的明文列项数据进行加密得到对应的密文列项数据,将密文列项数据置于密文数据库表的密文列中,同时将数据加密密钥与明文数据内容在数据库明文表中的位置标识信息进行关联存储在密钥库中;

其中,所述第一算法的执行过程为:

预设数据库明文表中的明文列项数据为M,数据加密密钥为K;

根据算法C=Enc(M,K),计算出对应的密文列项数据C,其中Enc为对称加密算法,优选为AES或国密SM4算法。

[0011] 本方案中,从数据库明文表中获取待查询列,利用第二算法对数据库明文表的待查询列生成比较索引,具体包括:

预设数据库明文表中待查询列的明文列项数据为{M_1,M_2,…, M_n};

预设安全参数为λ_1，随机选择λ_1比特长度的索引生成密钥k_1；

设伪随机函数$F:\{0,1\}^{\lambda_1}\times\{0,1\}^{m-1}\to\mathbb{Z}_3$，其中$\{0,1\}^{\lambda_1}$和$\{0,1\}^{m-1}$分别表示$\lambda_1$比特和m-1比特长度的二进制字符串，$\mathbb{Z}_3$表示模3的整数剩余类环；

对每个明文列项数据M，M取自于集合{M_1，M_2，…，M_n}，设M的二进制表示为$b_1b_2\cdots b_m$，令$[m]$表示整数集合{1，2，…，m}，对$i\in[m]$，计算：

$u_i=F(k_1,b_1b_2\cdots b_{i-1}\,||\,0^{m-i})+b_i(\mathrm{mod}\,3)$，符号||表示字符串级联，mod表示取模运算符；

计算出明文列项数据M的比较索引为$u=(u_1,u_2,\cdots,u_m)$；

输出索引生成密钥k和待查询列的所有明文列项数据的比较索引，并基于所有明文列项数据的比较索引形成比较索引列。

[0012] 本方案中，获取数据查询者提供的查询范围，基于查询范围并通过第三算法计算生成范围查询陷门，具体包括：

预设安全参数为λ_2，随机选择λ_2比特长度的陷门生成密钥k_2；

获取查询请求信息$(m_1,m_2,token)$，所述查询请求信息包括查询范围和token信息，m_1,m_2表示查询范围的上下区间值，$token$本次数据查询者的请求令牌；

选择伪随机函数$R:\{0,1\}^{\lambda_2}\times T\to\mathbb{Z}_3^m$，其中$T$表示token的取值空间，$\mathbb{Z}_3^m$表示m维的$\mathbb{Z}_3$向量，$\mathbb{Z}_3$表示模3的整数剩余类环；

根据查询请求信息$(m_1,m_2,token)$，利用第三算法计算：

$c_1=TrapGen(m_1)$，$c_2=TrapGen(m_2)$；$TrapGen$表示查询陷门生成算法；

$h=R(k_2,token)$；

$c_{1t}=c_1\boxplus h$，$c_{2t}=c_2\boxplus h$，其中符号⊞表示m维向量按分量模3相加；

输出范围查询陷门$T=(c_{1t},c_{2t},h)$。

[0013] 本方案中，基于范围查询陷门，并通过第四算法查询所述密文数据库表中的比较索引列，得到查询结果，具体包括：

对两个索引值c和c'，定义索引比较函数$Compare(c,c')$：

设$c=(u_1,u_2,\cdots,u_m)$，$c'=(u'_1,u'_2,\cdots,u'_m)$，如果$c=c'$，输出0；

否则令i是$u_i\neq u'_i$的最小正整数，若满足$u'_i=u_i+1\,(mod\,3)$，则输出-1

若满足$u_i=u'_i+1\,(mod\,3)$，则输出1；

比较函数输出1表示c对应的明文数据值大于c'对应的明文数据值，-1表示c对应的明文数据值小于c'对应的明文数据值，0则表示两者相等；

对查询陷门$T=(c_{1t},c_{2t},h)$，初始化集合ID为空，进行下面的计算：

还原范围查询的端点值$c_1=c_{1t}\boxminus h$和$c_2=c_{2t}\boxminus h$，符号⊟表示m维向量按分量模3相减；

对密文数据库表中查询列的每个索引值u，利用比较函数计算函数值：

$r_1=Compare(u,c_1)$，$r_2=Compare(u,c_2)$；

如果$r_1\neq-1$且$r_2\neq1$，则表示u对应的明文m在查询的范围$[m_1,m_2]$内；

将满足条件的u所在的行号id添加到集合ID中；

输出集合ID作为查询结果。

[0014] 本方案中，基于查询结果获取对应的数据解密密钥，并从密文数据库表中获取对应的密文数据内容，基于数据解密密钥并通过第五算法对密文数据内容进行解密，输出明文查询结果，具体包括：

验证查询是否有效，计算$R(k_2,token)$，如果等于h，表示查询有效，进行后续解密，否则输出“查询无效”，算法退出；

预设需要输出的明文数据集合为PlainData，初始化PlainData为空，对每个$id \in ID$，执行：

获取id对应行的数据解密密钥DK，根据实际需要解密这一行中某一列的的密文数据，得到明文数据Plain=Dec(Cipher,DK)，其中Cipher表示需要的那一列的密文数据内容，Dec表示对称解密算法；

将明文数据Plain添加到明文数据集合PlainData中；

输出明文数据集合PlainData。

[0015] 本方案中，在将生成的比较索引加入所述密文数据库表的比较索引列之后，所述方法还包括：

预设在预设时间段内有多个数据查询者提出查询请求信息，每个数据查询者的查询请求信息中包括各自的查询范围；

对多个查询范围进行取交集处理；

预设待查询列的数据依据对应的刻度线取值，统计在每个刻度参与交集的数据查询者数量；

将参与交集的数据查询者数量超过第一预设阈值的刻度判定为合并查询刻度；

将所有合并查询刻度进行并集处理，得到合并查询刻度的并集范围；

基于合并查询刻度的并集范围进行查询，得到合并明文查询结果；

基于每个数据查询者，判断其查询范围是否包含合并查询刻度，如果包括，则从合并明文查询结果中提取所有包含的合并查询刻度对应的明文查询结果，对所有包含的合并查询刻度以外的其它查询刻度，则执行第四算法查询过程；如果不包括，则基于对应数据查询者的查询范围直接执行第四算法查询过程。

[0016] 本发明第二方面还提出一种基于比较索引的密文数据范围查询系统，用于实现上述的基于比较索引的密文数据范围查询方法，所述系统包括：

密钥管理和授权模块，负责生成和管理数据加解密密钥和比较索引生成密钥，在数据查询者授权的前提下，提供比较索引生成密钥和数据解密密钥；

数据加密模块，负责从密钥管理和授权模块获取数据加密密钥，运行第一算法对数据库明文表中的明文数据内容进行加密，产生密文数据内容；

比较索引构造模块，负责从密钥管理和授权模块获取比较索引生成密钥，运行第二算法对数据库明文表中的待查询的明文列项数据生成比较索引；

范围查询陷门生成模块，负责接收数据查询者的查询范围，从密钥管理和授权模块获取陷门生成密钥，运行第三算法生成对应的范围查询陷门；

密文数据库查询模块，负责从范围查询陷门生成模块接收查询陷门，运行第四算法从密文数据库表中比较索引列进行查询，获得查询结果；

数据解密模块，负责从密文数据库查询模块接收查询结果，从密钥管理和授权模块获取比较索引生成密钥和数据解密密钥，运行第五算法验证本次查询的有效性，若有效则解密出明文数据内容，返回给数据查询者；

密文数据库存储模块，负责统一存储密文数据库表，供密文数据库查询模块进行查询以及为数据解密模块提供密文数据内容。

[0017] 本方案中，数据加密模块还执行以下步骤：

预设数据库明文表包括由多条明文数据内容，每条明文数据内容包括多个明文列项数据；

基于不同的明文数据内容分别生成对应的数据加密密钥；

采用数据加密密钥并按照第一算法对相应明文数据内容中的明文列项数据进行加密得到对应的密文列项数据，将密文列项数据置于密文数据库表的密文列中，同时将数据加密密钥与明文数据内容在数据库明文表中的位置标识信息进行关联存储在密钥库中；

其中，所述第一算法的执行过程为：

预设数据库明文表中的明文列项数据为M，数据加密密钥为K；

根据算法C=Enc(M,K)，计算出对应的密文列项数据C，其中Enc为对称加密算法，优选为AES或国密SM4算法。

[0018] 本方案中，比较索引构造模块还执行以下步骤：

预设数据库明文表中待查询列的明文列项数据为{M_1,M_2,…, M_n}；

预设安全参数为λ_1，随机选择λ_1比特长度的索引生成密钥 k_1；

设伪随机函数 $F:\{0,1\}^{\lambda_1}\times\{0,1\}^{m-1}\to\mathbb{Z}_3$，其中$\{0,1\}^{\lambda_1}$和$\{0,1\}^{m-1}$分别表示 λ_1比特和m-1比特长度的二进制字符串，$\mathbb{Z}_3$表示模3的整数剩余类环；

对每个明文列项数据M，M取自于集合{M_1,M_2,…, M_n}，设M的二进制表示为$b_1b_2\cdots b_m$，令$[m]$表示整数集合{1, 2, …,m}，对 $i\in[m]$，计算：

$u_i=F(k_1,b_1b_2\cdots b_{i-1}\,||\,0^{m-i})+b_i(\mathrm{mod}\,3)$，符号 ||表示字符串级联，mod表示取模运算符；

计算出明文列项数据M的比较索引为$u=(u_1,u_2,\cdots,u_m)$；

输出索引生成密钥k和待查询列的所有明文列项数据的比较索引，并基于所有明文列项数据的比较索引形成比较索引列。

[0019] 本发明的查询算法高效并能并行化地执行，算法泄露的明文信息比特数少，适用于大规模数据量下的范围查询业务场景。基于按比特的索引大小比较算法能够保证范围查询结果的准确性。利用token机制保证了范围查询请求的合法性和新鲜性，避免敌手发起非法查询或重放攻击，提高了范围查询业务的安全性。

[0020] 本发明的附加方面和优点将在下面的描述部分中给出，部分将从下面的描述中变得明显，或通过本发明的实践了解到。

附图说明

[0021] 图1示出了本发明一种基于比较索引的密文数据范围查询方法的流程图；

图2示出了本发明一种基于比较索引的密文数据范围查询系统的框图；

图3示出了本发明一种数据库明文表的示意图；

图4示出了本发明一种未加比较索引的密文数据库表的示意图；

图5示出了本发明一种已加比较索引的密文数据库表的示意图。

具体实施方式

[0022] 为了能够更清楚地理解本发明的上述目的、特征和优点，下面结合附图和具体实施方式对本发明进行进一步的详细描述。需要说明的是，在不冲突的情况下，本申请的实施例及实施例中的特征可以相互组合。

[0023] 在下面的描述中阐述了很多具体细节以便于充分理解本发明，但是，本发明还可以采用其他不同于在此描述的其他方式来实施，因此，本发明的保护范围并不受下面公开的具体实施例的限制。

[0024] 图1示出了本发明一种基于比较索引的密文数据范围查询方法的流程图。

[0025] 如图1所示，本发明第一方面提出一种基于比较索引的密文数据范围查询方法，所述方法包括：

S102，获取填有多条明文数据内容的数据库明文表以及数据加密密钥，采用数据加密密钥并利用第一算法对数据库明文表进行加密，得到对应的密文数据库表；

S104，从数据库明文表中获取待查询列，利用第二算法对数据库明文表的待查询列生成比较索引，将生成的比较索引加入所述密文数据库表的比较索引列；

S106，获取数据查询者提供的查询范围，基于查询范围并通过第三算法计算生成范围查询陷门；

S108，基于范围查询陷门，并通过第四算法查询所述密文数据库表中的比较索引列，得到查询结果；

S110，基于查询结果获取对应的数据解密密钥，并从密文数据库表中获取对应的密文数据内容，基于数据解密密钥并通过第五算法对密文数据内容进行解密，输出明文查询结果。

[0026] 本发明提出了一种基于比较索引的针对密文数据的范围查询方法，将待查询的明文数据转化为密文数据形式的比较索引，数据库中其他相关联的明文数据能灵活地进行独立加密，利用范围查询结果可以输出这些相关联的数据。技术方案主要包括明文数据的加密算法、比较索引构造算法、查询陷门生成算法、范围查询搜索算法以及密文数据解密算法等5个算法，可应用于密文数据库的范围查询业务场景。

[0027] 需要说明的是，本发明针对多个数据提供者的数据共享场景设计了一种基于比较索引的密文数据范围查询的实用性方法，主要解决了以下技术问题：

1、有效控制了密文索引的扩展尺寸，扩展倍数仅为2倍。对常用的整型数据的范围查询，明文数据长度为4个字节，此时密文索引固定为8个字节。即使对于明文数据条数比较庞大的场景，密文索引存储量也是可接受的。

[0028] 2、范围查询结果具有较好的准确性。针对明文数据，按比特设计用于比较的密文索引和比较函数，可以准确地比较出两个索引密文的大小(也是对应明文的大小)，保证了查询结果的准确性。

[0029] 3、给出了明文数据的具体泄露信息，严格地度量了算法方案的安全性。

[0030] 4、每次范围查询的陷门不同，可以为查询进行授权和规避重放攻击。通过引入查

询令牌token(一个随机字符串),为每次查询生成不同的查询陷门,实现了查询请求的唯一性。

[0031] 根据本发明的实施例,获取填有多条明文数据内容的数据库明文表以及数据加密密钥,采用数据加密密钥并利用第一算法对数据库明文表进行加密,得到对应的密文数据库表,具体包括:

预设数据库明文表包括由多条明文数据内容,每条明文数据内容包括多个明文列项数据;

基于不同的明文数据内容分别生成对应的数据加密密钥;

采用数据加密密钥并按照第一算法对相应明文数据内容中的明文列项数据进行加密得到对应的密文列项数据,将密文列项数据置于密文数据库表的密文列中,同时将数据加密密钥与明文数据内容在数据库明文表中的位置标识信息进行关联存储在密钥库中;

其中,所述第一算法的执行过程为:

预设数据库明文表中的明文列项数据为M,数据加密密钥为K;

根据算法C=Enc(M,K),计算出对应的密文列项数据C,其中Enc为对称加密算法,优选为AES或国密SM4算法。

[0032] 根据本发明的实施例,从数据库明文表中获取待查询列,利用第二算法对数据库明文表的待查询列生成比较索引,具体包括:

预设数据库明文表中待查询列的明文列项数据为{M_1,M_2,…, M_n};

预设安全参数为λ_1,随机选择λ_1比特长度的索引生成密钥k_1;

设伪随机函数 $F:\{0,1\}^{\lambda_1}\times\{0,1\}^{m-1}\to\mathbb{Z}_3$,其中 $\{0,1\}^{\lambda_1}$和$\{0,1\}^{m-1}$分别表示 λ_1比特和m-1比特长度的二进制字符串,$\mathbb{Z}_3$表示模3的整数剩余类环;

对每个明文列项数据M,M取自于集合{M_1,M_2,…, M_n},设M的二进制表示为$b_1b_2\cdots b_m$,令 $[m]$表示整数集合{1, 2, …,m},对$i\in[m]$,计算:

$u_i=F(k_1,b_1b_2\cdots b_{i-1}\,||\,0^{m-i})+b_i(\mathrm{mod}\,3)$,符号||表示字符串级联,mod表示取模运算符;

计算出明文列项数据M的比较索引为$u=(u_1,u_2,\cdots,u_m)$。

[0033] 输出索引生成密钥k和待查询列的所有明文列项数据的比较索引,并基于所有明文列项数据的比较索引形成比较索引列。

[0034] 根据本发明的实施例,获取数据查询者提供的查询范围,基于查询范围并通过第三算法计算生成范围查询陷门,具体包括:

预设安全参数为λ_2,随机选择 λ_2比特长度的陷门生成密钥k_2;

获取查询请求信息$(m_1,m_2,token)$,所述查询请求信息包括查询范围和token信息,m_1,m_2表示查询范围的上下区间值, $token$本次数据查询者的请求令牌;

选择伪随机函数$R:\{0,1\}^{\lambda_2}\times T\to\mathbb{Z}_3^m$,其中$T$表示token的取值空间,$\mathbb{Z}_3^m$表示m维的 $\mathbb{Z}_3$向量,$\mathbb{Z}_3$表示模3的整数剩余类环;

根据查询请求信息 $(m_1,m_2,token)$,利用第三算法计算:

$c_1=TrapGen(m_1)$,$c_2=TrapGen(m_2)$;$TrapGen$表示查询陷门生成算法;

$h=R(k_2,token)$;

$c_{1t}=c_1 \boxplus h$，$c_{2t}=c_2 \boxplus h$，其中符号⊞表示m维向量按分量模3相加；

输出范围查询陷门$T=(c_{1t},c_{2t},h)$。

[0035] 根据本发明的实施例，基于范围查询陷门，并通过第四算法查询所述密文数据库表中的比较索引列，得到查询结果，具体包括：

对两个索引值c和c'，定义索引比较函数$Compare(c,c')$：

设$c=(u_1,u_2,\cdots,u_m)$，$c'=(u'_1,u'_2,\cdots,u'_m)$，如果$c=c'$，输出0；

否则令i是$u_i \neq u'_i$的最小正整数，若满足$u'_i=u_i+1\ (mod\ 3)$，则输出-1

若满足$u_i=u'_i+1\ (mod\ 3)$，则输出1；

比较函数输出1表示c对应的明文数据值大于c'对应的明文数据值，-1表示c对应的明文数据值小于c'对应的明文数据值，0则表示两者相等；

对查询陷门$T=(c_{1t},c_{2t},h)$，初始化集合ID为空，进行下面的计算：

还原范围查询的端点值$c_1=c_{1t} \boxminus h$和$c_2=c_{2t} \boxminus h$，符号⊟表示m维向量按分量模3相减；

对密文数据库表中查询列的每个索引值u，利用比较函数计算函数值：

$r_1=Compare(u,c_1)$，$r_2=Compare(u,c_2)$；

如果$r_1 \neq -1$且$r_2 \neq 1$，则表示u对应的明文m在查询的范围$[m_1,m_2]$内；

将满足条件的u所在的行号id添加到集合ID中；

输出集合ID作为查询结果。

[0036] 根据本发明的实施例，基于查询结果获取对应的数据解密密钥，并从密文数据库表中获取对应的密文数据内容，基于数据解密密钥并通过第五算法对密文数据内容进行解密，输出明文查询结果，具体包括：

验证查询是否有效，计算$R(k_2,token)$，如果等于h，表示查询有效，进行后续解密，否则输出“查询无效”，算法退出；

预设需要输出的明文数据集合为PlainData，初始化PlainData为空，对每个$id \in ID$，执行：

获取id对应行的数据解密密钥DK，根据实际需要解密这一行中某一列的密文数据，得到明文数据Plain=Dec(Cipher,DK)，其中Cipher表示需要的那一列的密文数据内容，Dec表示对称解密算法；

将明文数据Plain添加到明文数据集合PlainData中；

输出明文数据集合PlainData。

[0037] 根据本发明的实施例，在将生成的比较索引加入所述密文数据库表的比较索引列之后，所述方法还包括：

预设在预设时间段内有多个数据查询者提出查询请求信息，每个数据查询者的查询请求信息中包括各自的查询范围；

对多个查询范围进行取交集处理；

预设待查询列的数据依据对应的刻度线取值，统计在每个刻度参与交集的数据查询者数量；

将参与交集的数据查询者数量超过第一预设阈值的刻度判定为合并查询刻度；

将所有合并查询刻度进行并集处理，得到合并查询刻度的并集范围；

基于合并查询刻度的并集范围进行查询，得到合并明文查询结果；

基于每个数据查询者，判断其查询范围是否包含合并查询刻度，如果包括，则从合并明文查询结果中提取所有包含的合并查询刻度对应的明文查询结果，对所有包含的合并查询刻度以外的其它查询刻度，则执行第四算法查询过程；如果不包括，则基于对应数据查询者的查询范围直接执行第四算法查询过程。

[0038] 可以理解，本发明通过合并查询能够减少部分刻度的重复查询，减少了算法计算次数，进而提升了多个数据查询者的查询效率。

[0039] 根据本发明的具体实施例，所述方法还包括：

对数据库明文表中的所有数据提供者的每一条明文数据内容进行摘要值计算，得第一摘要值，并存入摘要值数据库中；

当后续接收到其他数据库明文表的明文数据内容时，则对每条明文数据内容进行摘要值计算，得到每条明文数据内容的第二摘要值；

将每条明文数据内容的第二摘要值分别遍历摘要值数据库，并查找摘要值数据库中是否有相同的第一摘要值，如果有，则中止对该明文数据内容的存储处理。

[0040] 根据本发明的具体实施例，在获取对应的数据解密密钥之前，所述方法还包括：

预设数据库明文表中带有数据提供者设定的共享信息，并规定拥有该共享信息身份的数据查询者有权限获得数据解密密钥；

预设某数据查阅者有共享信息，且共享信息包括按照顺序排列的p个字符，且p为偶数；

将共享信息的p个字符进行两两配对，形成p/2个配对组，每个配对组包括前字符和后字符；

数据查阅者将共享信息的p/2个配对组的前字符分别作为调制光子串的各个调制基随机选取的第一选取源，并将每个字符串中的p/2个配对组的后字符作为调制光子串的调制初始信号的第二选取源；

数据查阅者从第一选取源中，随机选取对应的前字符，作为调制基，并从第二选取源中选取与前字符相对应的后字符，作为调制初始信号，由各个调制基分别将对应的调制初始信号调制成光子的偏振态；

将光子的偏振态与共享信息对应调制基的随机选取方式一并通过量子通信发送给查询系统；

查询系统接收到光子的偏振态与共享信息对应调制基的随机选取方式，基于共享信息对应调制基的随机选取方式从共享信息的p/2个配对组的前字符中找出对应的前字符作为测量基，并采用测量基对光子的偏振态进行测量，得到测量结果，将获得的测量结果与共享信息中的后字符进行比对，若都一致，则通过对数据查阅者的认证。

[0041] 需要说明的是，数据加解密密钥存储在密钥数据库中，且不同数据提供者的数据采用不同的密钥进行加密，即数据库明文表的不同位置标识的明文数据内容采用对应的密钥进行加密。然而在数据查询者从密钥数据库中获取对应的数据解密密钥时，则需要进行身份验证，本发明则基于共享信息进行对数据查询者授权身份进行验证。

[0042] 本发明的密钥数据库中包括每个密钥对应有一个认证共享信息，则在比对时，则遍历密钥数据库中的全部共享信息并进行比对，待有比对成功时，即可认证通过。

[0043] 本发明第二方面还提出一种基于比较索引的密文数据范围查询系统，用于实现上述的基于比较索引的密文数据范围查询方法，所述系统包括：

密钥管理和授权模块，负责生成和管理数据加解密密钥和比较索引生成密钥，在数据查询者授权的前提下，提供比较索引生成密钥和数据解密密钥；

数据加密模块，负责从密钥管理和授权模块获取数据加密密钥，运行第一算法对数据库明文表中的明文数据内容进行加密，产生密文数据内容；

比较索引构造模块，负责从密钥管理和授权模块获取比较索引生成密钥，运行第二算法对数据库明文表中的待查询的明文列项数据生成比较索引；

范围查询陷门生成模块，负责接收数据查询者的查询范围，从密钥管理和授权模块获取陷门生成密钥，运行第三算法生成对应的范围查询陷门；

密文数据库查询模块，负责从范围查询陷门生成模块接收查询陷门，运行第四算法从密文数据库表中比较索引列进行查询，获得查询结果；

数据解密模块，负责从密文数据库查询模块接收查询结果，从密钥管理和授权模块获取比较索引生成密钥和数据解密密钥，运行第五算法验证本次查询的有效性，若有效则解密出明文数据内容，返回给数据查询者；

密文数据库存储模块，负责统一存储密文数据库表，供密文数据库查询模块进行查询以及为数据解密模块提供密文数据内容。

[0044] 根据本发明的实施例，数据加密模块还执行以下步骤：

预设数据库明文表包括由多条明文数据内容，每条明文数据内容包括多个明文列项数据；

基于不同的明文数据内容分别生成对应的数据加密密钥；

采用数据加密密钥并按照第一算法对相应明文数据内容中的明文列项数据进行加密得到对应的密文列项数据，将密文列项数据置于密文数据库表的密文列中，同时将数据加密密钥与明文数据内容在数据库明文表中的位置标识信息进行关联存储在密钥库中；

其中，所述第一算法的执行过程为：

预设数据库明文表中的明文列项数据为M，数据加密密钥为K；

根据算法C=Enc(M,K)，计算出对应的密文列项数据C，其中Enc为对称加密算法，优选为AES或国密SM4算法。

[0045] 根据本发明的实施例，比较索引构造模块还执行以下步骤：

预设数据库明文表中待查询列的明文列项数据为{M_1,M_2,…, M_n}；

预设安全参数为λ_1，随机选择λ_1比特长度的索引生成密钥k_1；

设伪随机函数$F:\{0,1\}^{\lambda_1}\times\{0,1\}^{m-1}\to\mathbb{Z}_3$，其中 $\{0,1\}^{\lambda_1}$和$\{0,1\}^{m-1}$分别表示 λ_1比特和m-1比特长度的二进制字符串，$\mathbb{Z}_3$表示模3的整数剩余类环；

对每个明文列项数据M，M取自于集合{M_1,M_2,…, M_n}，设M的二进制表示为$b_1b_2\cdots b_m$，令$[m]$表示整数集合{1, 2, …,m}，对$i\in[m]$，计算：

$u_i=F(k_1,b_1b_2\cdots b_{i-1}\,||\,0^{m-i})+b_i(\mathrm{mod}\,3)$，符号||表示字符串级联，mod表示取模运算符；

计算出明文列项数据M的比较索引为$u=(u_1,u_2,\cdots,u_m)$。

[0046] 输出索引生成密钥k和待查询列的所有明文列项数据的比较索引，并基于所有明文列项数据的比较索引形成比较索引列。

[0047] 为了进一步说明本发明的技术方案，下面以具体实施例进行说明。

[0048] 本发明提出了一种基于比较索引的针对密文数据的范围查询方法，将待查询的明文数据转化为密文数据形式的比较索引，数据库中其他相关联的明文数据能灵活地进行独立加密，利用范围查询结果可以输出这些相关联的数据。技术方案主要包括明文数据的加密算法、比较索引构造算法、查询陷门生成算法、范围查询搜索算法以及密文数据解密算法等5个算法，可应用于密文数据库的范围查询业务场景。

[0049] 算法1：明文数据的加密算法PlainEnc

明文数据的加密算法对数据库中的明文内容进行加密，生成密文数据。这里的明文内容不限于要进行范围查询的数据(也可以对数据库中的其他关联数据进行加密)。

[0050] 算法输入：数据库中的明文数据M，加密密钥K(密钥K可以根据加密强度要求设置为对数据库中的每行或每列都不同，甚至每个数据都不同)。

[0051] 算法输出：M的密文数据C。

[0052] 算法描述：计算C=Enc(M,K)，其中Enc为一个对称加密算法(AES、SM4等)。

[0053] 算法2：比较索引构造算法BuildIndex

比较索引构造算法是对待查询的数据库表每列的明文数据进行索引构造，这些索引用于后续的大小比较，我们称这些索引为比较索引。

[0054] 算法输入：数据库表中待查询列的明文数据{M_1,M_2,…, M_n}.

算法输出：每个数据的比较索引index。

[0055] 算法描述：

(a)对安全参数为λ，随机选择 λ比特长度的索引生成密钥k；

(b)设伪随机函数 $F:\{0,1\}^{\lambda}\times\{0,1\}^{m-1}\to\mathbb{Z}_3$，其中$\{0,1\}^{\lambda}$和$\{0,1\}^{m-1}$分别表示 λ比特和m-1比特长度的二进制字符串，$\mathbb{Z}_3$表示模3的整数剩余类环。对每个明文数据M(M取自于集合{M_1,M_2,…, M_n})，设M的二进制表示为$b_1b_2\cdots b_m$，令$[m]$表示整数集合{1, 2, …,m}，对$i\in[m]$，计算：

$u_i=F(k,b_1b_2\cdots b_{i-1}\,||\,0^{m-i})+b_i \pmod 3$，符号||表示字符串级联。

[0056] 明文M的比较索引为 $u=(u_1,u_2,\cdots,u_m)$。

[0057] (c)输出索引生成密钥 k和所有明文数据的比较索引。

[0058] 算法3：查询陷门生成算法TrapGen

查询陷门生成算法是利用陷门生成密钥k对要查询的范围 $[m_1,m_2]$和本次查询请求令牌token(每次的token是不同的随机值)生成查询陷门T。

[0059] 算法输入：查询范围和token信息$(m_1,m_2,token)$。

[0060] 算法输出：范围查询陷门T。

[0061] 算法描述：

(a)选择伪随机函数$R:\{0,1\}^{\lambda}\times T\to\mathbb{Z}_3^m$，其中$T$表示token的取值空间，$\mathbb{Z}_3^m$表示m维的$\mathbb{Z}_3$向量。根据查询请求$(m_1,m_2,token)$，利用算法3(TrapGen算法)计算：

1) $c_1 = TrapGen(m_1)$， $c_2 = TrapGen(m_2)$；

2) $h = R(k,token)$；

3) $c_{1t} = c_1 \boxplus h$， $c_{2t} = c_2 \boxplus h$，其中符号⊞表示m维向量按分量模3相加。

[0062] (b)输出范围查询陷门 $T = (c_{1t}, c_{2t}, h)$.

算法4:范围查询搜索算法Search

范围查询搜索算法是利用范围查询陷门T来比对密文数据库表中比较索引列，通过设计的索引比较函数得到满足查询范围的数据库表的对应行序号id。

[0063] 算法输入:范围查询陷门T，密文数据库DB；

算法输出:满足查询范围的数据库表行号的集合ID。

[0064] 算法描述：

(a)对两个索引值c和c'，定义索引比较函数$Compare(c,c')$：

设$c = (u_1, u_2, \cdots, u_m)$， $c' = (u'_1, u'_2, \cdots, u'_m)$，

如果 $c = c'$，输出0；

否则令i是$u_i \neq u'_i$的最小正整数，若满足 $u'_i = u_i + 1 \ (mod\ 3)$，则输出-1；

若满足 $u_i = u'_i + 1 \ (mod\ 3)$，则输出1；

容易看出，比较函数输出1表示 c对应的明文数据值大于c'对应的明文数据值，-1表示 c对应的明文数据值小于 c'对应的明文数据值，0则表示两者相等。

[0065] (b)对查询陷门 $T = (c_{1t}, c_{2t}, h)$，初始化集合ID为空，进行下面的计算：

1)还原范围查询的端点值$c_1 = c_{1t} \boxminus h$和 $c_2 = c_{2t} \boxminus h$，符号⊟表示m维向量按分量模3相减；

2)对密文数据库表中查询列的每个索引值u，利用比较函数计算函数值:

$r_1 = Compare(u, c_1)$，$r_2 = Compare(u, c_2)$；

3)如果$r_1 \neq -1$且 $r_2 \neq 1$，则表示u对应的明文 m在查询的范围 $[m_1, m_2]$内。准确地说，明文m处于范围$m_1 \leq m \leq m_2$.如果要求$m_1 < m < m_2$，则可以修改判断条件为$r_1 = 1$并且$r_2 = -1$即可。将满足条件的 u所在的行号id添加到集合ID中(用于后续输出该行对应的所需要的数据信息)。

[0066] (c)输出集合ID。

[0067] 算法5:密文数据解密算法CipherDec

密文数据解密算法是对密文数据内容进行解密，首先根据数据表行号集合ID获取到对应的解密密钥，然后解密出该数据的明文值。注意这里我们可以根据实际需要，通过行号灵活地输出数据库表中任意关联数据列的值，而不是局限于查询列的明文值。我们在实施例中会给出一个简单的例子。

[0068] 算法输入:查询请求和查询结果信息 $(token, h, ID)$；

算法输出:相关需要的明文数据集合PlainData。

[0069] 算法描述：

(a)验证查询是否有效，计算$R(k,token)$，如果等于h，表示查询有效，进行后续解密，否则输出“查询无效”，算法退出；

(b)初始化PlainData为空,对每个 $id \in ID$,执行:

1)获取id对应行的解密密钥DK,根据实际需要解密这一行中某一列的的密文数据,得到明文数据内容Plain=Dec(Cipher,DK),其中Cipher表示需要的那一列的密文数据内容。

[0070] 2)将明文数据Plain添加到集合PlainData中。

[0071] (c)输出明文数据集合PlainData。

[0072] 利用上面给出的5个算法,本发明设计了基于比较索引的密文数据范围查询系统,系统由密钥管理和授权模块、数据加密模块、比较索引构造模块、范围查询陷门生成模块、密文数据库查询模块、数据解密模块以及密文数据库存储模块等7个模块构成,系统如图2所示。

[0073] (1)密钥管理和授权模块:负责生成和管理数据加解密密钥和比较索引生成密钥,在数据查询者授权的前提下,提供比较索引生成密钥和数据解密密钥。

[0074] (2)数据加密模块:负责从密钥管理和授权模块获取数据加密密钥,运行算法1(PlainEnc算法)对数据库表中明文数据内容进行加密,产生密文数据内容。

[0075] (3)比较索引构造模块:负责从密钥管理和授权模块获取比较索引生成密钥,运行算法2(BuildIndex算法)对数据库表中待查询的数据列生成比较索引,并存储到密文数据库存储模块。这里也可以看出我们给出的算法2能够应对数据动态变化的情况(新增或删除),新增的数据只需要计算新增数据的比较索引并添加到密文数据库即可。数据删除时只需删除对应数据行即可,对数据库表的其他数据没有影响。

[0076] (4)范围查询陷门生成模块:负责接收数据查询者的查询范围,从密钥管理和授权模块获取陷门生成密钥,运行算法3(TrapGen算法)生成对应的范围查询陷门。

[0077] (5)密文数据库查询模块:负责从范围查询陷门生成模块接收查询陷门,运行算法4(Search算法)从密文数据库中比较索引列进行查询,获得查询结果ID。

[0078] (6)数据解密模块:负责从密文数据库查询模块接收查询结果ID,从密钥管理和授权模块获取比较索引生成密钥和解密密钥,运行算法5(CipherDec算法)验证本次查询的有效性,若有效则解密出明文数据内容,返回给数据查询者。

[0079] (7)密文数据库存储模块:负责统一存储密文数据库,供密文数据库查询模块进行查询以及为数据解密模块提供密文数据内容。

[0080] 图2中也依次标识了整个系统的运行流程(用流程箭头上数字表示),从数据提供者提供的数据及待查询列数据开始处理,首先生成密文数据内容和待查询列的比较索引,存储在密文数据库存储模块中。然后数据查询者根据查询范围发起查询,范围查询陷门生成模块产生出查询陷门,密文数据库查询模块利用查询陷门进行密文范围查询获得数据库表行号集合ID。数据解密模块解密ID中对应行所需的密文数据列,最终将明文结果返回给数据查询者。

[0081] 方案效率分析:

(1)方案中涉及的算法全部采用对称密码原语进行设计,保证了算法执行的高效性。

[0082] (2)范围查询算法中比较两个索引时,只需要查找索引分量的不同值并验证是否有相差1的关系,比较算法非常高效。并且可以采用并行计算方式进一步提高比较效率,这

样即使对数据库表比较庞大的情况也能兼顾范围查询的速度。

[0083] 系统安全性分析：

(1)数据库表中对不同数据提供者的数据行可以采用不同的加密密钥,数据库表中同一行的不同列理论上能采用不同的密钥(但实际中为了减少密钥管理的复杂度可以不用如此高强度的加密方式),由此保证数据内容的安全性。

[0084] (2)密文数据库存储模块中存储的都是密文形式的比较索引和数据内容,即使密文存储模块被攻破,敌手也无法解密数据或根据比较索引反推查询列的明文(这是由加密算法和伪随机函数提供的安全性)。

[0085] (3)数据查询者提供的查询范围会被范围查询陷门生成模块进行变换,密文数据库查询模块和密文存储模块都不能知道查询范围的明文(只能比较出大小),保证了查询范围的安全性。

[0086] (4)相比现有技术中可能泄露一半以上的明文比特,我们方案中泄露的明文信息仅有1比特。这是因为在比较函数Compare中,当两个索引不完全相同时,对应的明文总会有1比特不同,而比较函数会返回两个明文第一个不同的比特位置。虽然在安全性上可以通过设计交互式方案来进一步减少泄露的信息量,但多轮的交互会严重影响方案的执行效率。

[0087] 下面以一个具体实施例来展示本发明中给出的范围查询算法的执行过程。

[0088] (1)利用算法1(明文数据的加密算法PlainEnc)对数据库表的明文进行加密。这里以图3中的数据库表为例,范围查询列为年龄列,假设资产列也是需要加密的数据列,因此我们将年龄列和资产列都进行加密。这里我们使用AES加密算法,前两行和后两行分别使用不同的密钥加密,用来模拟两个数据提供者各自提供的两条数据,得到的密文数据库表如图4所示。

[0089] (2)利用算法2(比较索引构造算法BuildIndex)对数据库明文表的年龄列生成比较索引,其中伪随机函数F使用SHA256哈希算法实现,将生成的比较索引一起放入密文数据库表中,如图5所示。

[0090] (3)设查询者现在提出的查询范围[30, 36],要求输出满足该范围的人员姓名和资产信息。设本次查询的token为abcdef-123456-789012,利用算法3(查询陷门生成算法TrapGen)生成范围查询陷门T=(01001121, 01111200, 12100011). 注意这里为了方便表示,我们将陷门T中的每个向量的分量进行级联成字符串,比如其中起点的陷门向量为(0, 1, 0, 0, 1, 1, 2, 1)就直接表示为01001121。

[0091] (4)利用范围查询陷门T和算法4(范围查询搜索算法Search)查询密文数据库的比较索引列,得到查询结果为ID={1,2},表示数据库表中的第1行和第2行的数据在要求的查询范围之内。

[0092] (5)利用查询结果ID,输出要求的姓名和资产信息,其中资产信息是密文数据需要解密。最终的查询结果为:(张三,100w) 和 (李四,50w)。

[0093] 本发明设计了一种高效安全的基于比较索引的密文数据范围查询方法,查询算法高效并能并行化地执行,算法泄露的明文信息比特数少。索引尺寸扩展小,扩展倍数为明文比特长度的log_2^3倍(约为1.58倍,具体编码实现时为2倍),非常适用于大规模数据量下的范围查询业务场景。

[0094] 基于按比特的索引大小比较算法能够保证范围查询结果的准确性,对于数据动态

变化的场景,方案也能够支持方便地进行新数据比较索引的增加和已有数据的删除。

[0095] 利用token机制保证了范围查询请求的合法性和新鲜性,避免敌手发起非法查询或重放攻击,提高了范围查询业务的安全性。

[0096] 在本申请所提供的几个实施例中,应该理解到,所揭露的设备和方法,可以通过其它的方式实现。以上所描述的设备实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,如:多个单元或组件可以结合,或可以集成到另一个系统,或一些特征可以忽略,或不执行。另外,所显示或讨论的各组成部分相互之间的耦合、或直接耦合、或通信连接可以是通过一些接口,设备或单元的间接耦合或通信连接,可以是电性的、机械的或其它形式的。

[0097] 上述作为分离部件说明的单元可以是、或也可以不是物理上分开的,作为单元显示的部件可以是、或也可以不是物理单元;既可以位于一个地方,也可以分布到多个网络单元上;可以根据实际的需要选择其中的部分或全部单元来实现本实施例方案的目的。

[0098] 另外,在本发明各实施例中的各功能单元可以全部集成在一个处理单元中,也可以是各单元分别单独作为一个单元,也可以两个或两个以上单元集成在一个单元中;上述集成的单元既可以采用硬件的形式实现,也可以采用硬件加软件功能单元的形式实现。

[0099] 本领域普通技术人员可以理解:实现上述方法实施例的全部或部分步骤可以通过程序指令相关的硬件来完成,前述的程序可以存储于计算机可读取存储介质中,该程序在执行时,执行包括上述方法实施例的步骤;而前述的存储介质包括:移动存储设备、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0100] 或者,本发明上述集成的单元如果以软件功能模块的形式实现并作为独立的产品销售或使用时,也可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明实施例的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机、服务器、或者网络设备等)执行本发明各个实施例所述方法的全部或部分。而前述的存储介质包括:移动存储设备、ROM、RAM、磁碟或者光盘等各种可以存储程序代码的介质。

[0101] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到变化或替换,都应涵盖在本发明的保护范围之内。因此,本发明的保护范围应以所述权利要求的保护范围为准。

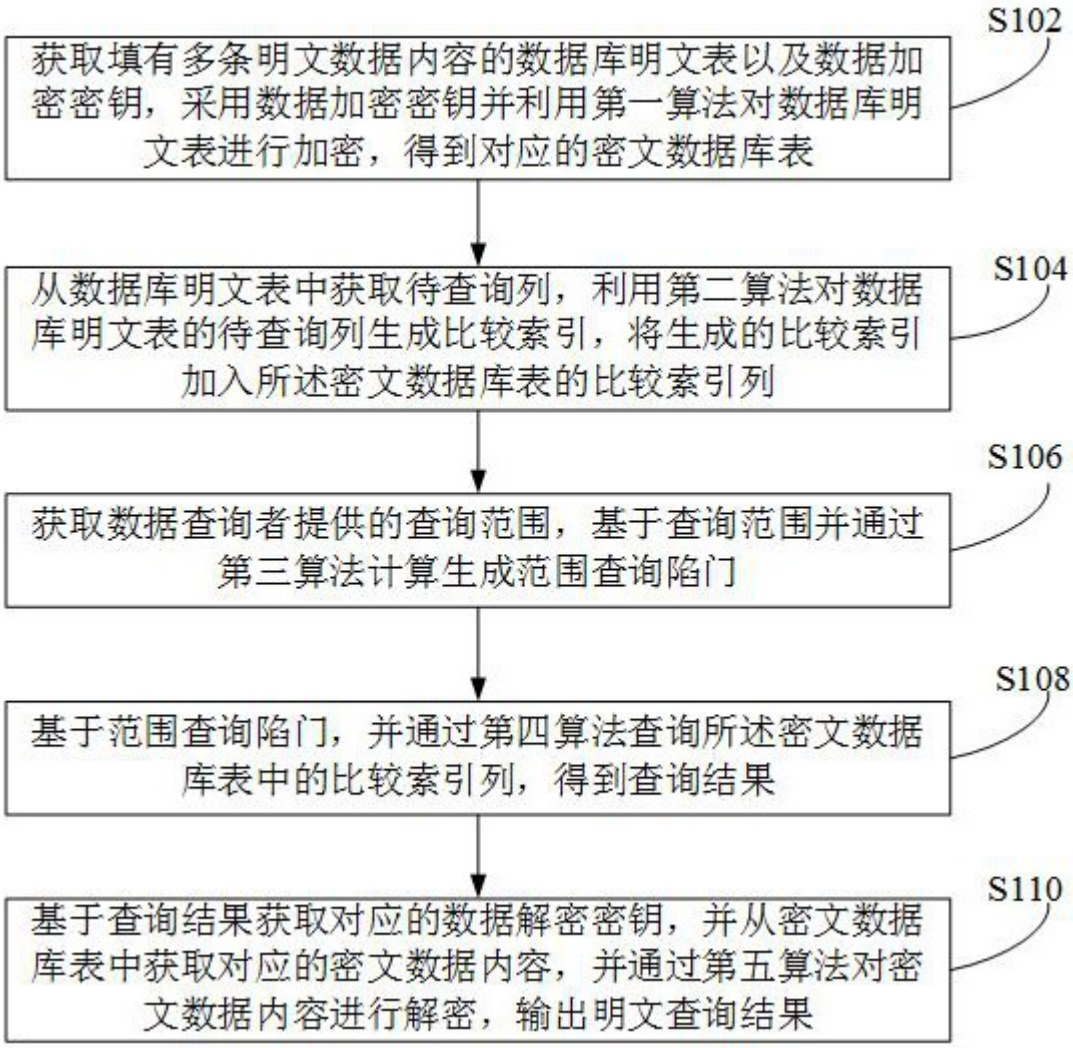

S102
获取填有多条明文数据内容的数据库明文表以及数据加密密钥，采用数据加密密钥并利用第一算法对数据库明文表进行加密，得到对应的密文数据库表
S104
从数据库明文表中获取待查询列，利用第二算法对数据库明文表的待查询列生成比较索引，将生成的比较索引加入所述密文数据库表的比较索引列
S106
获取数据查询者提供的查询范围，基于查询范围并通过第三算法计算生成范围查询陷门
S108
基于范围查询陷门，并通过第四算法查询所述密文数据库表中的比较索引列，得到查询结果
S110
基于查询结果获取对应的数据解密密钥，并从密文数据库表中获取对应的密文数据内容，并通过第五算法对密文数据内容进行解密，输出明文查询结果

图1

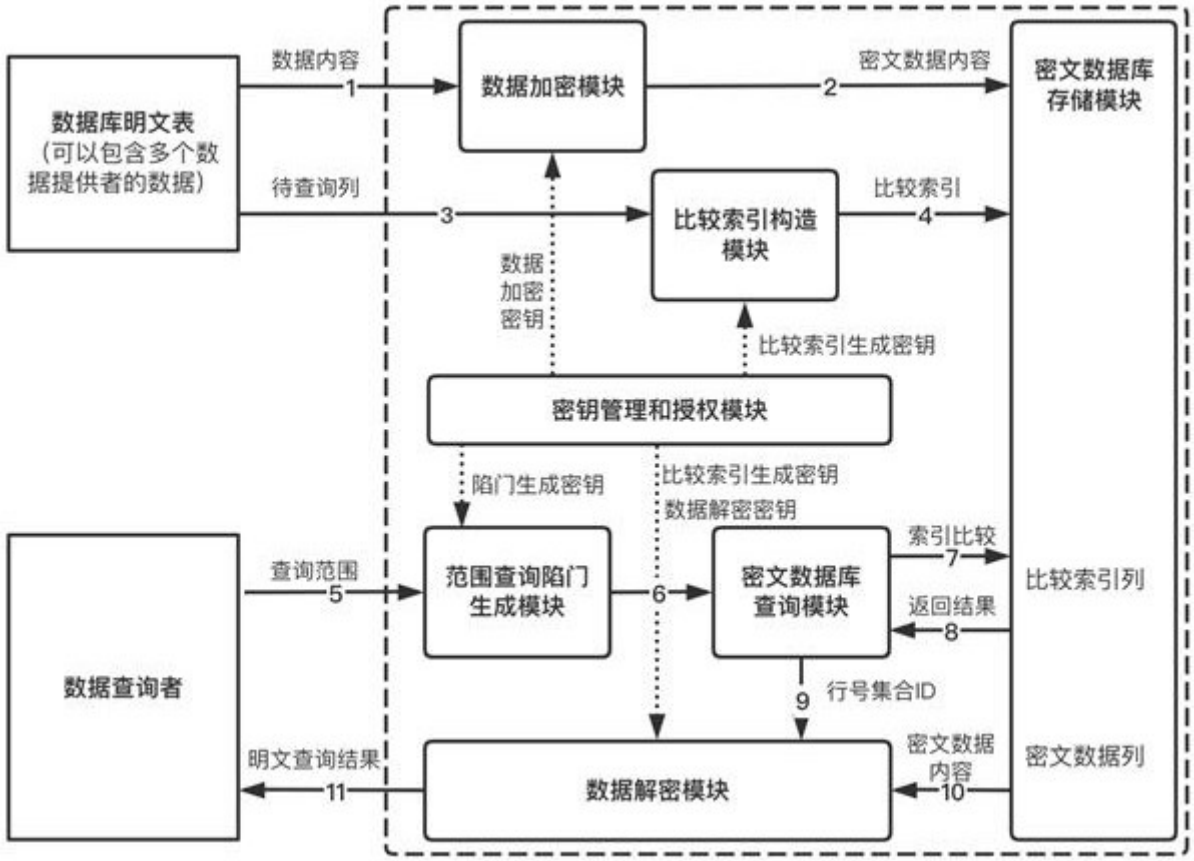


图2

序号	姓名	年龄	资产
1	张三	35	100w
2	李四	32	50w
3	王五	40	200w
4	赵六	45	30w

图3

序号	姓名	年龄	资产
1	张三	e20259e3f1800a14f3a92573f3feba0c	c62cfb72401a792f0e4eafab5ba9e8f3
2	李四	004ba7dd79b9bfe370320e7a0fbaf43f	c5e492b4b9b74c28ec4ab8f7f97d46a4
3	王五	090e0f3b0c2ff8bfb52964a572b57a9d	7218242faf65e9feb56f1878de849367
4	赵六	386783dfe920bfda1276e42f96a1518d	154825c5572850e26977e5a0cb91c08f

图4

序号	姓名	年龄	资产	比较索引（年龄）
1	张三	e20259e3f1800a14f3a92573f3feba0c	c62cfb72401a792f0e4eafab5ba9e8f3	22011121
2	李四	004ba7dd79b9bfe370320e7a0fbaf43f	c5e492b4b9b74c28ec4ab8f7f97d46a4	22011111
3	王五	090e0f3b0c2ff8bfb52964a572b57a9d	7218242faf65e9feb56f1878de849367	22012222
4	赵六	386783dfe920bfda1276e42f96a1518d	154825c5572850e26977e5a0cb91c08f	22012020

图5