有理数域上两方集合的高效保密计算

窦家维 刘旭红 王文丽
(陕西师范大学数学与信息科学学院 西安 710119)

摘 要 安全多方计算已经成为密码学的一个重要研究方向,是国际密码学界的一个研究热点.集合运算可以用来描述许多实际问题,因此研究集合的保密计算问题具有重要的理论与实际意义.目前,关于整数集上集合问题的保密计算已有很多重要成果,但在有理数域上集合问题的保密计算尚未见到有关研究报道.本文主要研究有理数域上集合的两方保密计算问题.首先,提出一种新的转化思想,将任意有理数编码为直角坐标系中一条过原点的直线,并结合三角形面积计算公式,将有理数域上元素与集合关系问题转化为整数范围内向量内积问题,进一步结合 Paillier 加密方案设计了集合运算的保密计算协议.其次,设计了将平面上的有理点编码为有理数的新编码方案,在此基础上设计了判定有理点是否属于有理点集合的保密判定协议.最后,应用模拟范例证明了所设计协议在半诚实模型下是安全的,并通过理论分析和实验测试说明协议是高效的.

关键词 保密计算;有理数;集合运算;编码方案;同态加密
中图法分类号 TP309 **DOI 号** 10.11897/SP.J.1016.2020.01397

Privacy Preserving Two-Party Rational Set Computation

DOU Jia-Wei LIU Xu-Hong WANG Wen-Li
(*School of Mathematics and Information Science, Shaanxi Normal University, Xi'an* 710119)

Abstract Secure multiparty computation is an important research area of cryptography and a key privacy-preserving technology in information society. It is also a focus of the international cryptographic community. Secure set operation is an important topic of secure multiparty scientific computation, an important field of secure multiparty computation. As many practical problems can be reduced to set problems, it is of both theoretically and practically important significance to study the secure multiparty computation of the set operation. There are many works to study secure integer set computation. To the best of our knowledge, there is little research investigating the secure rational set computation. The existing solutions to secure integer set computation can hardly be directly applied to solve secure rational set operation problems. This paper focuses on solving secure multiparty rational set operation problems including privately determining whether a rational number belongs to a rational set, privately computing the intersection or the union of two private sets, and privately computing the cardinality of the intersection/union of two private sets. The key to these problems is to privately determine whether a private rational number belongs to a private rational set, which can be reduced to privately determine whether two private numbers are equal. In this paper, we represent a rational number u by the line that passes the origin with slope u, and two rational numbers are equal if and only if the two lines to represent them coincide. Therefore, privately determine whether two rational numbers are equal is transformed into privately determining whether two private lines coincide. To privately determining whether two

收稿日期:2018-08-23;在线发布日期:2019-06-23.本课题得到国家自然科学基金(61272435)资助.**窦家维**,博士,副教授,主要研究方向为应用数学与应用密码学.E-mail:jiawei@snnu.edu.cn.**刘旭红**,硕士研究生,主要研究方向为应用数学与应用密码学.**王文丽**,硕士研究生,主要研究方向为应用数学与应用密码学.

private lines coincide, we can choose one point on a private line and two different points on another private line; these three points form a triangle; the area of this triangle is zero if and only if the two lines coincide. Then we utilize the Paillier additively homomorphic encryption cryptosystem to design a protocol to privately compute the area of such a triangle. Using this protocol as a building block and with some techniques, we further design protocols to privately determine the relationship between a private rational number and a private rational set, to privately compute the intersection set of two private sets, the cardinality of the intersection set of two private sets, the union of two private sets, etc. Second, we design an encoding scheme which transforms rational points into rational numbers. Based on this encoding scheme, we develop a secure multiparty computation protocol to privately determine the relationship between a rational point and a rational point set. We, using the well accepted simulation paradigm, prove that our protocols are secure in the semi-honest model. Finally, we theoretically analyze the efficiency of our protocols, and test the efficiency on a PC. The simulation verifies theoretical analysis conclusion that our protocols are of high efficiency. Our works solve some secure multiparty rational set operation problems which have not been investigated before.

Keywords secure computation; rational number; set operation; encoding scheme; homomorphic encryption

1 引 言

随着网络技术的迅速发展,多方协同计算已经成为计算机网络中越来越普遍的计算方式,在协同计算过程中对数据隐私的保护也越来越受到人们的关注.安全多方计算(Secure Multiparty Computation, SMC)使得一组互不信任的参与者能够在保护各自隐私数据的前提下进行协同计算.

自从姚期智教授在文献[1]中首次提出两方参与者的保密计算问题以来,安全多方计算问题的研究一直受到广泛关注,目前已发展成为密码学的一个重要研究方向[2-5].Goldreich 利用秘密共享和布尔电路理论证明了任意的保密计算问题都是可解的,并给出了通用解决方案[5].应用通用方案解决具体问题时计算效率与通信效率很低,实际中并不可行,近年来人们不断提出具有实际应用背景的安全多方计算问题并研究其高效的解决方案,推进了安全多方计算的发展.目前研究的问题主要有保密信息比较[6-7]、保密数据挖掘[8-9]、保密几何计算[10-11]、隐私入侵检测[12]、保密竞拍[13-14]等.

两方集合的保密计算问题作为保密计算的一类基本问题,在保密数据挖掘、保密选举、保密查询和模式匹配等方面有重要的应用.在许多实际应用中,参与者需要对他们的集合数据进行保密计算,包括保密判定元素与集合的关系[15-16]、保密计算集合的交集或交集的势[16-23]、保密计算集合的并集[24-27]及保密判定集合包含关系[27-31]等问题.文献[15-16]分别基于对称加密算法和公钥加密算法设计了元素与集合关系的保密判定协议,在协议中需要公布集合的势.文献[17]降低了文献[16]的复杂性,所设计协议的复杂性是集合势的准线性(almost linear)函数.文献[18]基于第三方硬件的安全设置设计了集合交集保密计算协议,如果两个参与者的集合元素均属于域 $D=\{0,1\}^t$,协议具有 $O(mt)$ 的计算复杂性.文献[19-22]应用更多的密码学工具设计了两方交集的高效计算协议.文献[23]应用 Bloom 过滤器设计了集合交集势的保密计算协议,但仅可得到近似结果.文献[24-27]分别设计了集合并集的保密计算协议,其中文献[24]将参与者的集合表示成多项式的形式,并利用翻转罗朗级数和秘密共享的方法给出了集合并集的保密计算方案;文献[25]通过对参与者集合中的元素进行不经意排序,根据排序结果得到集合并集;文献[26]借助全同态加密算法和多项式求值的方法给出计算集合并集的协议.这些并集协议的计算复杂性都较高,并且无法保证各参与者集合势的私密性.

文献[27]将集合以多项式形式表示,利用门限同态加密方案解决了集合包含问题,但此方案需要门限解密,复杂性较高;文献[28-30]分别基于不同

的密码学知识设计了集合包含关系的保密判定协议，效率比文献[27]有所提高. 文献[31]通过设定全集并利用编码方法设计了集合包含关系的保密判定协议，协议的计算复杂性与全集的势同阶，当全集元素不太多时，计算效率较高.

如上所述，很多研究者应用不同的密码学工具对于集合的各种基本运算提出了多种解决方案，这些方案所用方法和执行效率不尽相同，但这些已有解决方案都将集合元素限制在整数范围内.

目前关于有理数域上的保密计算研究还较少，文献[32]基于连分数编码方法和全同态加密算法设计了有理数加密方案，为研究有理数保密计算问题提供了新思路. 在很多应用场景中，需要应用有理数集合运算描述问题. 例如，(1) 要保密判断一个有理数是否是一个多项式的根，这个问题即可转化为保密判断一个有理数是否属于一个有理数集合的问题；(2) 在检验医学中常见到比值项目报告，较常规的如清蛋白/球蛋白(A/G)比值，由于 A/G 比值能反映肝功能状况，在临床医学中能协助诊断和鉴别. 假如有两家医学研究机构对各自患者的 A/G 比值进行统计分析，分别得到一组数据集(有理数集合). 两家医学机构希望合作统计分析患者群体的整体 A/G 比值，但医疗数据属于隐私数据，不能泄露具体的数据集，这就需要保密计算两个数据集的并集；(3) 假设在同一平面上，Alice 和 Bob 分别拥有私密直线集合 A 和 B，并假设 A(或 B)中直线互不平行且直线斜率均为有理数，记其直线斜率构成的集合为 X(或 Y). Alice 和 Bob 希望保密计算集合 A 中有哪些直线(或有几条直线)与 B 中某直线平行，这个问题即可转化为保密计算两个有理数集合 X 与 Y 的交集(或交集的势)问题. 由于有理数域上集合保密计算问题在实际中有广泛的应用，目前尚没有见到关于这类问题的研究报道. 本文主要对这类问题进行研究.

一般地，解决有理数域上集合保密计算问题最直接的想法是将其转化为相应的整数集合计算问题，这就需要首先将有理数编码为整数，再应用已有的整数集合的计算方案解决，但这种方法并不能很好地解决有理数域上集合的保密计算问题. 这是因为，一方面，保密计算要求所设计的编码方法不能泄露参与方的私密数据，而且目前适合于保密计算的将有理数编码为整数的编码方法还很少；另一方面，关于整数集合问题已有的解决方案中，效率较高的计算协议大都要求集合元素属于某个适当的全集. 对于有理数集合，由于有理数的稠密性，将其集合元素编码成整数后，可能需要全集的势很大或者可能根本无法确定全集，对这样的集合如果再应用已有的整数集合保密计算方案进行计算，计算复杂性将非常高. 因此已有的整数范围内集合问题保密计算协议很难直接推广应用于有理数域上的相关问题. 对于有理数集合保密计算问题，需要根据有理数本身的特点设计构造新的高效的解决方案.

为此，本文将有理数域上一维(二维)数据编码为直角坐标系中的直线，利用坐标系中三角形面积公式并结合 Paillier 加密方案，研究构造有理数域上各种集合运算的保密计算协议. 本文协议仅要求集合元素为有理数，对于数据的大小范围没有任何限制.

本文的主要贡献如下：

(1) 应用全新的方法解决有理数域上两方集合保密计算这类新问题. 对于有理数域上几类基本的集合运算设计保密计算协议，包括元素与集合关系的保密判定协议，集合交/并集和集合包含问题的保密计算协议，以及判定有理点是否属于有理点集合的保密计算协议.

(2) 提出了关于有理数以及有理点新的编码方法和转化技巧. 由于目前常用的公钥密码系统只能加密整数，为了应用这些密码系统解决有理数域上的保密计算问题，本文提出了关于有理数以及有理点新的编码方法和转化技巧，这些方法和技巧为解决有理数域上的保密计算问题提供了新的途径，有重要的实际意义.

(3) 所设计的协议更加安全高效. 目前已有的整数范围内集合运算的保密计算协议大多只保护集合元素的隐私性，本文协议能同时保证集合元素以及集合势的隐私性. 本文协议都具有线性或常数计算复杂性，与已有整数范围内相关集合计算协议相比较效率更高.

(4) 所设计协议具有广泛适用性. 本文协议适合于任何有理数集合，对这些协议进行适当修改或组合，可以解决更多的有理数保密计算问题.

2　预备知识

2.1　安全多方计算模型及安全性定义

本小节主要介绍下文中要用到的一些基本概念

与基础知识.本部分内容主要取自文献[5].

两方计算. 两方计算是一个将任意给定的输入对映射为输出对的随机过程,此过程用函数表示为 $f(x_1,x_2)\rightarrow(f_1(x_1,x_2),f_2(x_1,x_2))$. 即对于每一个输入对$(x_1,x_2)$,输出对是随机变量$(f_1(x_1,x_2),f_2(x_1,x_2))$. 记这样的函数为 $f=(f_1,f_2)$.

半诚实模型. 如果参与者能够按照协议要求执行协议,但在协议执行后参与者可能尝试根据其在执行协议时所收集到的信息,推算出其他参与者保密数据的额外信息,这样的计算模型称为半诚实模型.本文主要研究半诚实模型下的两方安全计算问题.

模拟范例. 模拟范例是两方计算协议安全性证明中广泛使用的证明方法.具体描述如下.

假设参与计算的两方参与者分别为 P_1 和 P_2. 设 $f=(f_1,f_2)$是一个概率多项式时间函数,π 表示计算函数 f 的一个两方协议.将 $P_i(i=1,2)$在执行协议 π 时获得的信息序列记为

$$view_i^{\pi}(x_1,x_2)=(x_i,r^i,m_1^i,\cdots,m_t^i,f_i(x_1,x_2)),$$

其中 x_i和r^i分别表示 P_i的输入及产生的随机数,m_j^i表示 P_i收到的第 j 个消息,$f_i(x_1,x_2)$表示 P_i获得的输出结果.

半诚实模型下协议的安全性定义如下.

定义 1. 对于函数 $f=(f_1,f_2)$和计算 f 的协议π,如果存在概率多项式时间算法 S_1和 S_2,使得

$$\{S_1(x_1,f_1(x_1,x_2))\}_{x_1,x_2}\overset{c}{\equiv}\{view_1^{\pi}(x_1,x_2)\}_{x_1,x_2} \quad (1)$$

$$\{S_2(x_2,f_2(x_1,x_2))\}_{x_1,x_2}\overset{c}{\equiv}\{view_2^{\pi}(x_1,x_2)\}_{x_1,x_2} \quad (2)$$

成立,则称 π 是计算 f 的保密计算协议,其中$\overset{c}{\equiv}$表示计算不可区分.

2.2 三角形面积公式

在平面解析几何中给出了三角形面积公式.假设有一个三角形,三个顶点分别为 $p_1(x_1,y_1)$,$p_2(x_2,y_2)$和 $p_3(x_3,y_3)$,当 p_1,p_2,p_3以逆时针顺序排列时,三角形面积 $S_{\Delta p_1p_2p_3}$由下式计算:

$$S_{\Delta p_1p_2p_3}=\frac{1}{2}[y_1(x_3-x_2)+x_1(y_2-y_3)+x_2y_3-x_3y_2] \quad (3)$$

显然,$S_{\Delta p_1p_2p_3}=0$ 当且仅当 p_1,p_2,p_3共线.

2.3 Paillier 加密方案

Paillier 加密方案具体描述如下[33]:

密钥生成. 根据给定的安全参数 κ,选择两个大素数 p,q,计算 $N=pq$,$\omega=\mathrm{lcm}(p-1,q-1)$. 定义函数 $L(x)=(x-1)/N$,随机选择一个生成元 $g\in Z_{N^2}^*$,使得

$$\gcd(L(g^{\omega}\bmod N^2),N)=1,$$

其中 $\mathrm{lcm}(a,b)$及 $\gcd(a,b)$分别表示 a 和 b 的最小公倍数及最大公约数.则加密方案的公钥 $pk=(g,N)$,私钥 $sk=\omega$,下文中以 E 和 D 分别表示加密和解密算法.

加密. 为了加密明文 $m\in Z_N$,选择随机数 $r\in Z_N^*$,并按下式计算密文:

$$c=E(m)=g^m r^N\bmod N^2.$$

解密. 给定密文 $c\in Z_{N^2}^*$,计算明文

$$m=D(c)=\frac{L(c^{\omega}\bmod N^2)}{L(g^{\omega}\bmod N^2)}\bmod N.$$

加法同态性. 由于下面性质成立:

$$\begin{aligned}E(m_1)\cdot E(m_2)&=g^{m_1}r_1^N\cdot g^{m_2}r_2^N\bmod N^2\\&=g^{m_1+m_2}(r_1r_2)^N\bmod N^2\\&=E(m_1+m_2\bmod N),\end{aligned}$$

因此 Paillier 加密方案具有加法同态性.

Paillier 加密方案是语义安全的.即同一明文能够加密成许多不同的密文,所有密文都是计算不可区分的.

在下文中,以 Paillier 加密方案为基础所做的密文乘积运算和模乘运算均是在模 N^2 的意义下进行的.

3 有理数域上两方集合的保密计算

3.1 编码方法与转化原理

关于有理数域上集合问题的保密计算还未见有相关文献研究.为了研究这类新的保密计算问题,我们需要对有理数设计新的编码方法.

编码方法. 根据计算几何知识可知,对于任意给定的有理数 s,可将其编码为某平面直角坐标系中一条通过坐标原点,以 s 为斜率的直线.下文中所涉及到的有理数编码,均假设在某一确定直角坐标系中进行.

有理数相等问题转化原理. 根据上述编码方法,任意有理数均可编码为直角坐标系中过原点的直线.我们将采取下面方法将两个有理数是否相等的判定问题进行转化.

假设计算双方分别拥有非负有理数 u 和 v,他们首先将 u(或 v)编码为通过坐标原点以 u(或 v)为斜率的直线 l_u(或 l_v),其中一方在直线 l_u上选取一点 $p_1(x_1,y_1)$,另一方在直线 l_v上选择两个不同点 $p_2(x_2,y_2)$,$p_3(x_3,y_3)$,显然有理数 u,v 相等当且仅当 $\Delta p_1p_2p_3$的面积为零,如图 1 所示.

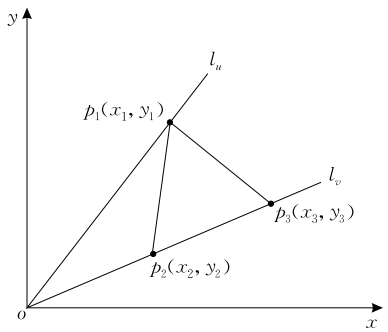

图 1　三点构造 $\Delta p_1 p_2 p_3$

为了将有理数相等判定问题转化为整数计算问题，要求计算双方分别将 p_1，p_2，p_3 选为整数点（(x,y)为整数点是指点的坐标 x，y 均为整数. 由于直线斜率是有理数，在直线上容易选到整数点）. 由于 $p_2(x_2,y_2)$，$p_3(x_3,y_3)$同在直线 l_v 上，因此有 $x_2 y_3 - x_3 y_2 = 0$. 如此，由面积公式(3)得到：

$$S_{\Delta p_1 p_2 p_3} = \frac{1}{2}|y_1(x_3 - x_2) + x_1(y_2 - y_3)| \quad (4)$$

因此两个有理数 u 和 v 是否相等的判定问题可转化为整数算式(4)是否为零的判定问题.

在下文中，假设所考虑的有理数（u，v 以及 U，V 的所有元素）均是正有理数. 如若不然，计算双方可事先商量一个充分大的有理数 M，对于 u，v 以及每个集合元素同加 M 进行变换，使变换后的有理数均是正的，如此做法即可把上面讨论的集合计算问题转化为正有理数范围内有关问题进行计算. 我们约定，本文所涉及的随机数都是随机整数，并以 Q^+ 表示正有理数集合. 在协议执行中如果某参与方没有获得任何输出，约定该参与方的输出为空串 λ.

3.2　有理数域上元素与集合关系的保密判定

问题描述. Alice 拥有一个私密的有理数集合 $U=\{u_1,\cdots,u_n\}$，Bob 拥有一个私密的有理数 v. 他们希望合作保密判定 v 是否属于集合 U，而不泄露 U 和 v 的任何额外信息（包括要保密集合 U 的势）.

计算原理. 由于 $v\in U$ 当且仅当 v 与 U 中某个元素相等. 我们应用上述编码方法和转化技巧进行下面操作：Bob 将 v 编码为过坐标原点以 v 为斜率的直线 L_0，并在 L_0 上随机选取两个不同的整数点 $q_1(x_{01},y_{01})$和 $q_2(x_{02},y_{02})$（假设 $x_{02}>x_{01}$）；对于每个 $i\in[1,n]=\{1,2,\cdots,n\}$，Alice 将 u_i 编码为过坐标原点以 u_i 为斜率的直线 L_i，并在直线 L_i 上至少选取一个整数点（为了隐藏集合 U 的势，在同一直线 L_i 上可以选取若干个不同的整数点），假设 Alice 在所有直线上共选取了 l 个不同的整数点，记为 $p_i=(x_i,y_i)$，$i\in[1,l]$.

根据面积公式(4)可得到：

$$S_{\Delta p_i q_1 q_2} = \frac{1}{2}|ay_i + bx_i|,$$

其中

$$a = x_{02} - x_{01} > 0,\ b = y_{01} - y_{02} < 0 \quad (5)$$

由于 $v\in U$ 当且仅当存在 $t\in[1,l]$，使得 $S_{\Delta p_t q_1 q_2}=0$，因此所有 l 个三角形面积的乘积也为零，即可得到：

$$\begin{aligned} v\in U &\Leftrightarrow \exists t\in[1,l],\ S_{\Delta p_t q_1 q_2}=0 \\ &\Leftrightarrow \exists t\in[1,l],\ ay_t + bx_t = 0 \\ &\Leftrightarrow \prod_{i=1}^{l}(ay_i + bx_i) = 0 \end{aligned} \quad (6)$$

下面主要根据等价条件(6)判定 v 是否属于集合 U. 为了不泄露 v 具体与 U 中哪个元素相同，需要进一步把式(6)中最后一个连乘积改写成两个向量内积形式：记 $Z=\prod_{i=1}^{l}(ay_i+bx_i)$，则有

$$Z = a^l c_l + a^{l-1} b c_{l-1} + \cdots + a b^{l-1} c_1 + b^l c_0,$$

其中

$$c_{l-k} = \sum_{i_1,\cdots,i_k\in[1,l]} \left(x_{i_1}\cdots x_{i_k}\prod_{j\neq i_1\cdots i_k} y_j\right),\ 0\le k\le l \quad (7)$$

定义向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 如下：

$$\begin{aligned} \boldsymbol{A} &= (c_l, c_{l-1}, \cdots, c_1, c_0), \\ \boldsymbol{B} &= (a^l, a^{l-1}b, \cdots, ab^{l-1}, b^l) \end{aligned} \quad (8)$$

其中 a，b 以及 $c_i(i=1,\cdots,l)$分别由式(5)和式(7)计算得到. 因此可知 $Z=\boldsymbol{A}\cdot\boldsymbol{B}$，即有

$$v\in U \Leftrightarrow \prod_{i=1}^{l}(ay_i+bx_i)=0 \Leftrightarrow \boldsymbol{A}\cdot\boldsymbol{B}=0 \quad (9)$$

例如，当 $l=3$ 时，

$$\begin{aligned} Z &= \prod_{i=1}^{3}(ay_i+bx_i) = \boldsymbol{A}\cdot\boldsymbol{B} \\ &= a^3 y_1 y_2 y_3 + a^2 b(x_1 y_2 y_3 + x_2 y_1 y_3 + x_3 y_1 y_2) + \\ &\quad ab^2(x_1 x_2 y_3 + x_1 x_3 y_2 + x_2 x_3 y_1) + b^3 x_1 x_2 x_3, \end{aligned}$$

此时，

$$\begin{aligned} \boldsymbol{A} &= (c_3, c_2, c_1, c_0) \\ &= (y_1 y_2 y_3, x_1 y_2 y_3 + x_2 y_1 y_3 + x_3 y_1 y_2, \\ &\quad x_1 x_2 y_3 + x_1 x_3 y_2 + x_2 x_3 y_1, x_1 x_2 x_3), \\ \boldsymbol{B} &= (a^3, a^2 b, ab^2, b^3). \end{aligned}$$

根据式(9)，即可将有理数 v 是否属于集合 U 的判定问题转化为两个整数向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的内积是否为零的问题. 下面将以式(9)为基础，结合应用 Paillier 加密方案设计保密判定协议.

为叙述方便，定义二元谓词：

$$P(v,U)=\begin{cases}1, & \text{如果 } v\in U\\ 0, & \text{如果 } v\notin U\end{cases}.$$

具体协议如下：

协议 1. 保密判定一个有理数是否属于一个有理数集合.

输入：Alice 和 Bob 分别输入有理数集合 $U=\{u_1,\cdots,u_n\}$ 和有理数 v

输出：Bob 输出 $P(v,U)$

1. Bob 生成 Paillier 加密方案的公钥/私钥(pk/sk). 并将公钥 $pk=(g,N)$ 发送给 Alice.

2. Alice 按照计算原理所述，根据有理数 $u_1,\cdots,u_n$ 构造相应的直线 $L_1,\cdots,L_n$，在这些直线上共取 l 个不同的整数点 (x_i,y_i)，$i\in[1,l]$，按照式(8)构造向量 $\boldsymbol{A}=(c_l,c_{l-1},\cdots,c_1,c_0)$. 并将 l 发送给 Bob.

3. Bob 按照计算原理所述，根据有理数 v 构造相应的直线 L_0，在其上任取两个整数点 (x_{01},y_{01})，(x_{02},y_{02})，并按照式(8)构造向量 $\boldsymbol{B}=(a^l,a^{l-1}b,\cdots,ab^{l-1},b^l)$. 进一步，Bob 选取随机数 r，使其满足 $r>\max\{|a^{l-1}b|,\cdots,|ab^{l-1}|,|b^l|\}$，使向量 $\boldsymbol{B}'=(a^l+r,a^{l-1}b+r,\cdots,ab^{l-1}+r,b^l+r)$ 中所有分量大于 0. Bob 加密随机数 r 和向量 $\boldsymbol{B}'$ 得到 $E(r)$ 和

$$E(\boldsymbol{B}')=(E(a^l+r),E(a^{l-1}b+r),\cdots,E(ab^{l-1}+r),E(b^l+r)),$$

Bob 将 $E(r)$ 和 $E(\boldsymbol{B}')$ 发送给 Alice.

4. Alice 任意选择随机数 $t_1,t_2>0$，计算

$$S=E(a^l+r)^{t_1c_l}E(a^{l-1}b+r)^{t_1c_{l-1}}\cdots E(ab^{l-1}+r)^{t_1c_1}E(b^l+r)^{t_1c_0}E(r)^{t_2}$$

以及 $z=t_1(c_l+c_{l-1}+\cdots+c_1+c_0)+t_2$，并把 S,z 发送给 Bob.

5. Bob 对 S 进行解密得到 s. 如果 $s=rz$，输出 $P(v,U)=1$；否则，输出 $P(v,U)=0$.

3.2.1 协议 1 的正确性

根据 Paillier 加密方案的加法同态性，在协议第 5 步中，Bob 解密得到

$$s=t_1\boldsymbol{A}\cdot\boldsymbol{B}+t_1r(c_l+c_{l-1}+\cdots+c_1+c_0)+rt_2,$$

并进一步计算 $rz=r[t_1(c_l+c_{l-1}+\cdots+c_1+c_0)+t_2]$. 由于 $t_1\neq0$，所以

$$s=rz\Leftrightarrow t_1\boldsymbol{A}\cdot\boldsymbol{B}=0\Leftrightarrow v\in U,$$

因此，协议 1 是正确的.

3.2.2 协议 1 的安全性

首先考虑 Bob 数据的安全性. 在协议中，Bob 仅给 Alice 发送了密文 $E(r)$ 和 $E(\boldsymbol{B}')$. 由于 Alice 没有私钥进行解密，根据 Paillier 加密方案的语义安全性，Alice 无法得到 Bob 数据的任何额外信息.

关于 Alice 集合的安全性分析如下. (1) 首先在协议的第 2 步，Alice 发送给 Bob 数据 l，由于 l 是 Alice 任意选取的随机数且 $l>n$，它的随机性可以保护集合 U 的势的私密性，且不会泄露集合元素的任何信息；(2) Alice 在协议第 4 步发送给 Bob 两个数据 S,z，Bob 进行解密得到 $s=D(S)$，这样 Bob 将获得下面两个算式：

$$s=t_1\boldsymbol{A}\cdot\boldsymbol{B}+t_1r(c_l+c_{l-1}+\cdots+c_1+c_0)+rt_2,$$

$$z=t_1(c_l+c_{l-1}+\cdots+c_1+c_0)+t_2.$$

如果 $v\notin U$，Bob 可求得 $t_1\boldsymbol{A}\cdot\boldsymbol{B}=s-rz$. 因 t_1，t_2 是 Alice 选择的随机数，Bob 根据 $z=t_1(c_l+c_{l-1}+\cdots+c_1+c_0)+t_2$ 和 $t_1\boldsymbol{A}\cdot\boldsymbol{B}=s-rz$ 无法获得随机数 t_1,t_2 以及 $c_l+c_{l-1}+\cdots+c_1+c_0$ 的任何信息，也无法获得向量 $\boldsymbol{A}$ 以及 Alice 集合元素的任何其他额外信息. 因此，协议 1 是安全的.

关于协议 1 的安全性有下面的定理 1.

定理 1. 在半诚实模型下协议 1 是安全的.

证明. 为了证明定理 1，需要构造模拟器 S_1(或 S_2)，使式(1)(或(2))成立.

首先构造 S_1. 接收到输入 (U,λ) 后，S_1 按以下方式运行：

(1) S_1 首先任意选择 v'，按照计算原理所述，根据 v' 构造直线 L'_0，在其上任取两个不同整数点 (x'_{01},y'_{01})，(x'_{02},y'_{02})，并按照式(8)构造向量

$$\hat{\boldsymbol{B}}=(\hat{a}^l,\hat{a}^{l-1}\hat{b},\cdots,\hat{a}\hat{b}^{l-1},\hat{b}^l),$$

其中 $\hat{a}=x'_{02}-x'_{01}$，$\hat{b}=y'_{01}-y'_{02}$.

(2) S_1 任意选取随机数 $\hat{r}$，使向量 $\hat{\boldsymbol{B}}'=(\hat{a}^l+\hat{r},\hat{a}^{l-1}\hat{b}+\hat{r},\cdots,\hat{a}\hat{b}^{l-1}+\hat{r},\hat{b}^l+\hat{r})$ 中所有分量大于 0. 加密随机数 $\hat{r}$ 和向量 $\hat{\boldsymbol{B}}'$，得到 $E(\hat{r})$ 和

$$E(\hat{\boldsymbol{B}}')=(E(\hat{a}^l+\hat{r}),E(\hat{a}^{l-1}\hat{b}+\hat{r}),\cdots,E(\hat{a}\hat{b}^{l-1}+\hat{r}),E(\hat{b}^l+\hat{r})).$$

(3) S_1 计算

$$\hat{S}=E(\hat{a}^l+\hat{r})^{t_1c_l}E(\hat{a}^{l-1}\hat{b}+\hat{r})^{t_1c_{l-1}}\cdots E(\hat{a}\hat{b}^{l-1}+\hat{r})^{t_1c_1}E(\hat{b}^l+\hat{r})^{t_1c_0}E(\hat{r})^{t_2}.$$

(4) S_1 计算

$$\hat{s}=t_1\boldsymbol{A}\cdot\hat{\boldsymbol{B}}+t_1\hat{r}(c_l+c_{l-1}+\cdots+c_1+c_0)+\hat{r}t_2$$

以及 $\hat{r}z=\hat{r}[t_1(c_l+c_{l-1}+\cdots+c_1+c_0)+t_2]$.

由于在协议执行中，

$$view_1^{\pi}(v,U)=(U,E(r),E(\boldsymbol{B}'),\lambda),$$

而 S_1 在模拟过程中产生的信息序列为

$$S_1(U,f_1(v,U))=(U,E(\hat{r}),E(\hat{\boldsymbol{B}}'),\lambda).$$

由于 Alice 没有解密密钥，由 Paillier 加密方案的语义安全性可知，对 Alice 来说，有 $E(r)\overset{c}{\equiv}E(\hat{r})$，$E(\boldsymbol{B}')\overset{c}{\equiv}E(\hat{\boldsymbol{B}}')$，因此

$\{S_1(U,f_1(v,U))\}_{v,u_i\in Q^+}\overset{c}{\equiv}\{view_1^{\pi}(v,U)\}_{v,u_i\in Q^+}$.

接收到输入$(v,f_2(v,U)=P(v,U))$后，S_2按以下方式运行：

(1) S_2任意选择有理数$u'_1,\cdots,u'_{n'}$，构成集合$U'=\{u'_1,\cdots,u'_{n'}\}$，并满足$P(v,U')=P(v,U)$. S_2按照计算原理所述，根据$u'_1,\cdots,u'_{n'}$构造相应的直线$L'_1,\cdots,L'_{n'}$，再任意选取随机数$l'>n'$，在n'条直线上共取l'个不同的整数点(x'_i,y'_i)，$i\in[1,l']$. 并类似于式(8)构造向量$\hat{\boldsymbol{A}}=(\hat{c}_{l'},\hat{c}_{l'-1},\cdots,\hat{c}_1,\hat{c}_0)$和$\hat{\boldsymbol{B}}=(a^{l'},a^{l'-1}b,\cdots,ab^{l'-1},b^{l'})$，其中

$$\hat{c}_{l'-k}=\sum_{i_1,\cdots,i_k\in[1,l']}\Big(x'_{i_1}\cdots x'_{i_k}\sum_{j\neq i_1,\cdots,i_k}y'_j\Big),\ 0\leq k\leq l'.$$

(2) S_2任意选择随机数$\hat{t}_1,\hat{t}_2>0$，计算

$$\hat{S}=E(a^{l'}+r)^{\hat{t}_1\hat{c}_{l'}}E(a^{l'-1}b+r)^{\hat{t}_1\hat{c}_{l'-1}}\cdots E(ab^{l'-1}+r)^{\hat{t}_1\hat{c}_1}E(b^{l'}+r)^{\hat{t}_1\hat{c}_0}E(r)^{\hat{t}_2},$$

以及$\hat{z}=\hat{t}_1(\hat{c}_{l'}+\hat{c}_{l'-1}+\cdots+\hat{c}_1+\hat{c}_0)+\hat{t}_2$.

(3) S_2计算

$$\hat{s}=\hat{t}_1\hat{\boldsymbol{A}}\cdot\hat{\boldsymbol{B}}+\hat{t}_1r(\hat{c}_{l'}+\hat{c}_{l'-1}+\cdots+\hat{c}_1+\hat{c}_0)+r\hat{t}_2,$$

以及$r\hat{z}=r[\hat{t}_1(\hat{c}_{l'}+\hat{c}_{l'-1}+\cdots+\hat{c}_1+\hat{c}_0)+\hat{t}_2]$.

由于在协议执行中，

$$view_2^{\pi}(v,U)=(v,l,S,z,P(v,U)),$$

而S_2在模拟过程中产生的信息序列为

$$S_2(v,f_2(v,U))=(v,l',\hat{S},\hat{z},P(v,U')).$$

首先，由于l为Alice任意选取的随机数，因此$l\overset{c}{\equiv}l'$；虽然Bob有密钥可以解密，但根据安全性分析过程可知，根据S和z的值，Bob无法得到关于向量$\boldsymbol{A}$的任何信息，故有$S\overset{c}{\equiv}\hat{S}$，$z\overset{c}{\equiv}\hat{z}$；又因为$P(v,U')=P(v,U)$，因此

$$\{S_2(v,f_2(v,U))\}_{v,u_i\in Q^+}\overset{c}{\equiv}\{view_2^{\pi}(v,U)\}_{v,u_i\in Q^+}.$$

证毕.

3.3　有理数域上集合交集及其势的保密计算

问题描述. Alice和Bob分别拥有私密有理数集合$U=\{u_1,\cdots,u_n\}$和$V=\{v_1,\cdots,v_m\}$，他们希望保密计算两个集合的交集，而不泄露集合U和V的任何额外信息(包括要保密集合U和V的势).

计算原理. 基本思想是Alice和Bob通过应用协议1逐个判断集合V中的元素是否属于集合U，从而得到两集合的交集.

为了保护集合U和V的势的私密性，Alice和Bob共同商议一个正整数l，满足$l>m,l>n$. 按照3.2节所述的计算原理，Alice在其集合元素所对应的直线上共取l个不同的整数点(每条直线上至少取一个点)，按照式(8)计算集合U对应的向量$\boldsymbol{A}=(c_l,c_{l-1},\cdots,c_1,c_0)$；Bob类似于3.2节的计算原理，对于每个元素$v_j\in V$，$j\in[1,m]$，根据式(8)构造向量$\boldsymbol{B}_j=(a_j^l,a_j^{l-1}b_j,\cdots,a_jb_j^{l-1},b_j^l)$. 则有

$$v_j\in U\Leftrightarrow\boldsymbol{A}\cdot\boldsymbol{B}_j=0.$$

下文中，记

$$w_j=a_j^l+a_j^{l-1}b_j+\cdots+a_jb_j^{l-1}+b_j^l\qquad(10)$$

与协议1类似，上述做法能够隐藏集合U的势.

为了隐藏集合V的势，Bob在协议中需要添加一些“假密文”，使得在保证计算结果正确性的同时，还要使Alice无法识别真假密文，以此来保护集合V的势的隐私性.

协议2.　有理数集合交集保密计算.

输入：Alice和Bob分别输入有理数集合$U=\{u_1,\cdots,u_n\}$和$V=\{v_1,\cdots,v_m\}$

输出：Bob输出$U\cap V$

准备：按照计算原理所述，Alice根据集合U构造向量$\boldsymbol{A}=(c_l,c_{l-1},\cdots,c_1,c_0)$；对于每一个$j\in[1,m]$，Bob构造$v_j$对应的向量$\boldsymbol{B}_j=(a_j^l,a_j^{l-1}b_j,\cdots,a_jb_j^{l-1},b_j^l)$.

1. Alice生成Paillier加密方案的公钥/私钥(pk/sk). 并选取随机数$r>0$，计算向量

$$\boldsymbol{A}'=(c_l+r,c_{l-1}+r,\cdots,c_1+r,c_0+r)=(c'_l,c'_{l-1},\cdots,c'_1,c'_0).$$

Alice加密随机数r，并将公钥$pk=(g,N)$以及向量$\boldsymbol{A}'$和密文$E(r)$发送给Bob.

2. (a) 对于每个$j\in[1,m]$，Bob选取随机数t_j互不相同，使得$t_j(\boldsymbol{A}'\cdot\boldsymbol{B}_j)>0$；并选取随机数$s>0$，计算$Z_j,Z'_j$如下：

$$Z_j=E[t_j(\boldsymbol{A}'\cdot\boldsymbol{B}_j)]E(r)^s,\ Z'_j=t_jw_j+s.$$

(b) 对于每个$j\in[m+1,l_1]$，Bob选取一些随机数对(r_j,s_j)，$r_j,s_j>0$互不相同，计算Z_j,Z'_j如下：

$$Z_j=E(r)^{r_j},\ Z'_j=s_j.$$

(c) 对于每个$j\in[l_1,l]$，Bob选取一些随机数对(α_j,α_j)，$\alpha_j>0$，计算Z_j,Z'_j如下：

$$Z_j=E(r)^{\alpha_j},Z'_j=\alpha_j.$$

Bob将l维数组$[(Z_1,Z'_1),\cdots,(Z_l,Z'_l)]$中的元素进行随机置换(记随机置换为$\pi$)，并将随机置换后的数组$[(Z_{\pi(1)},Z'_{\pi(1)}),\cdots,(Z_{\pi(l)},Z'_{\pi(l)})]$发送给Alice(这里有$\{\pi(1),\cdots,\pi(l)\}=\{1,\cdots,l\}$).

3. 对于每个$\pi(j)\in[1,l]$，Alice解密$Z_{\pi(j)}$得到$z_{\pi(j)}=D(Z_{\pi(j)})$，并计算$z'_{\pi(j)}=rZ'_{\pi(j)}$. Alice比较$z_{\pi(j)}$和$z'_{\pi(j)}$：如果$z_{\pi(j)}=z'_{\pi(j)}$，令$d_{\pi(j)}=1$；否则，令$d_{\pi(j)}=0$. Alice将向量$(d_{\pi(1)},\cdots,d_{\pi(l)})$发送给Bob.

4. Bob根据随机置换π和向量$(d_{\pi(1)},\cdots,d_{\pi(l)})$构造集合$H$：$v_{\pi(j)}\in H$当且仅当$d_{\pi(j)}=1$且$j\in[1,m]$. Bob输出集合$H$.

3.3.1 协议 2 的正确性

我们只需证明 $H=U\cap V$ 成立即可. 根据 Paillier 加密方案的加法同态性,在协议第 3 步的计算中,Alice 得到:

当 $j\in[1,m]$ 时, $\begin{cases} z_j=D(Z_j)=t_j\boldsymbol{A}\cdot\boldsymbol{B}_j+r(t_jw_j+s) \\ z'_j=r(t_jw_j+s) \end{cases}$;

当 $j\in[m+1,l]$ 时, $z_j=D(Z_j)=rr_j$, $z'_j=rs_j$;或 $z_j=D(Z_j)=r\alpha_j$, $z'_j=r\alpha_j$.

由于 $j\in[1,m]$ 时, $t_j\neq0$,因此有

$$\begin{aligned} d_{\pi(j)}=1, j\in[1,m]&\Leftrightarrow z_j=z'_j, j\in[1,m]\\ &\Leftrightarrow t_j\boldsymbol{A}\cdot\boldsymbol{B}_j=0, j\in[1,m]\\ &\Leftrightarrow v_j\in U, j\in[1,m]. \end{aligned}$$

由此知对于 $j\in[1,m]$,使得 $z_j=z'_j$ 成立的 v_j 属于交集 $U\cap V$,这表明在 $d_{\pi(1)},\cdots,d_{\pi(l)}$ 中, $d_{\pi(j)}=1$ 且 $j\in[1,m]$ 当且仅当 $v_{\pi(j)}\in U\cap V$,故 $H=U\cap V$,协议 2 是正确的.

3.3.2 协议 2 的安全性

首先注意到,Alice 和 Bob 共同商议一个正整数 l,满足 $l>m$, $l>n$. 在协议执行中,Alice 在集合元素相对应的 n 条直线上共取 l 个不同点(每条直线上至少取一个点). 如此做法既不影响协议的正确性,又保证了集合 U 的势的私密性.

下面进一步分析 Alice 集合数据的安全性. Bob 收到 Alice 发送的向量 $\boldsymbol{A}'$ 和 $E(r)$,Bob 没有私钥解密,无法获得随机数 r 的任何消息. 根据向量 $\boldsymbol{A}'$ 的定义可知,Bob 可以得到向量 $\boldsymbol{A}$ 中任意两个分量 c_i,c_j 的关系式 $c_i-c_j=c'_i-c'_j$,又由向量 $\boldsymbol{A}$ 的定义可知,Alice 集合是安全的. 例如当 $n=3$ 时, $\boldsymbol{A}=(c_3,c_2,c_1,c_0)$,其中,

$$c_3=y_1y_2y_3,\ c_2=x_1y_2y_3+x_2y_1y_3+x_3y_1y_2,$$

$$c_1=x_1x_2y_3+x_1x_3y_2+x_2x_3y_1,\ c_0=x_1x_2x_3.$$

Bob 仅知道其中任意两个分量 c_i,c_j 之间的关系式,无法得到集合 U 的元素信息. 在协议第 3 步,Alice 将向量 $(d_{\pi(1)},\cdots,d_{\pi(l)})$ 发送给 Bob,Bob 据此也仅能够确定 $U\cap V$. 因此 Alice 集合数据是安全的.

接下来分析 Bob 集合的安全性. 在协议中 Bob 通过添加一些“假密文”数对: (Z_j,Z'_j), $j\in[m+1,l]$,共得到 l 个密文数对构成的数组 $[(Z_1,Z'_1),\cdots,(Z_l,Z'_l)]$,对其元素进行随机置换后将 $[(Z_{\pi(1)},Z'_{\pi(1)}),\cdots,(Z_{\pi(l)},Z'_{\pi(l)})]$ 发送给 Alice. Alice 对其解密及计算后仅能得到 $z_{\pi(j)}=D(Z_{\pi(j)})$ 以及 $z'_{\pi(j)}=rZ'_{\pi(j)}$. 当 $z_{\pi(j)}=z'_{\pi(j)}$ 时,Alice 无法区分 $\pi(j)\in[1,m]$ 满足 $\boldsymbol{A}\cdot\boldsymbol{B}_{\pi(j)}=0$,或是 $\pi(j)\in[l_1,l]$ 满足 $z_j=r\alpha_j=z'_j=r\alpha_j$. 当 $z_{\pi(j)}\neq z'_{\pi(j)}$ 时,Alice 也无法区分 $\pi(j)\in[1,m]$ 满足 $\boldsymbol{A}\cdot\boldsymbol{B}_{\pi(j)}\neq0$,亦或是 $\pi(j)\in[m+1,l_1]$ 满足 $z_j=rr_j\neq z'_j=rs_j$. 从而可以保护集合 V 的势的私密性.

下面进一步分析 Bob 集合数据的安全性. 由于 Alice 仅得到 Bob 随机置换后的 l 维数组 $[(Z_{\pi(1)},Z'_{\pi(1)}),\cdots,(Z_{\pi(l)},Z'_{\pi(l)})]$,对其进行解密及计算后得到 $z_{\pi(j)}=D(Z_{\pi(j)})$ 以及 $z'_{\pi(j)}=rZ'_{\pi(j)}$.

(1) 当 $z_{\pi(j)}=z'_{\pi(j)}$ 时,Alice 无法得到 Bob 集合的任何信息.

(2) 当 $z_{\pi(j)}\neq z'_{\pi(j)}$, $\pi(j)\in[1,m]$ 时, S_1 选取随机数 $t_j(A'\cdot\hat{B}_j)>0$;并选取随机数 $\hat{s}>0$,计算 $\hat{Z}_j$, $\hat{Z}'_j$ 如下:

$$z_j-z'_j=$$

$$t_ja_j^l\left[c_l+c_{l-1}\left(\frac{b_j}{a_j}\right)+\cdots+c_1\left(\frac{b_j}{a_j}\right)^{l-1}+c_0\left(\frac{b_j}{a_j}\right)^l\right] \quad (11)$$

因为 t_j 是 Bob 所选的私密数据,从式(11)中 Alice 无法获得 Bob 数据 a_j,b_j 的任何信息. 因此 Bob 集合是安全的.

综上所述,协议 2 是安全的. 关于协议 2 的安全性有下面的定理.

定理 2. 在半诚实模型下协议 2 是安全的.

证明. 下面应用模拟范例严格证明定理 2,构造模拟器 S_1(或 S_2),使式(1)(或(2))成立.

首先构造 S_1. 接收到输入 (U,λ) 后, S_1 按以下方式运行:

(1) S_1 任意选择集合 $V'=\{v'_1,\cdots,v'_{m'}\}$,并类似于协议 2 中向量 $\boldsymbol{B}_j$ 的构造方式,对于每一个 $j\in[1,m']$,构造 v'_j 的对应向量

$$\hat{\boldsymbol{B}}_j=(\hat{a}_j^l,\hat{a}_j^{l-1}\hat{b}_j,\cdots,\hat{a}_j\hat{b}_j^{l-1},\hat{b}_j^l),$$

并计算 $\hat{w}_j=\hat{a}_j^l+\hat{a}_j^{l-1}\hat{b}_j+\cdots+\hat{a}_j\hat{b}_j^{l-1}+\hat{b}_j^l$.

(2)(a) 对于每一个 $j\in[1,m']$, S_1 选取随机数 $\hat{t}_j$ 互不相同,使得 $\hat{t}_j(\boldsymbol{A}'\cdot\hat{\boldsymbol{B}}_j)>0$;并选取随机数 $\hat{s}>0$,计算 $\hat{Z}_j,\hat{Z}'_j$ 如下:

$$\hat{Z}_j=E[\hat{t}_j(\boldsymbol{A}'\cdot\hat{\boldsymbol{B}}_j)]E(r)^{\hat{s}},\ \hat{Z}'_j=\hat{t}_j\hat{w}_j+\hat{s}.$$

(b) 对于每个 $j\in[m'+1,l'_1]$, S_1 选取随机数对 $(\hat{r}_j,\hat{s}_j)$, $\hat{r}_j,\hat{s}_j>0$ 互不相同,计算 $\hat{Z}_j,\hat{Z}'_j$ 如下:

$$\hat{Z}_j=E(r)^{\hat{r}_j},\ \hat{Z}'_j=\hat{s}_j.$$

(c) 对于 $j\in[l'_1,l]$, S_1 选取随机数对 $(\hat{\alpha}_j,\hat{\alpha}_j)$, $\hat{\alpha}_j>0$,计算 Z_j,Z'_j 如下:

$$Z_j=E(r)^{\hat{\alpha}_j},\ Z'_j=\hat{\alpha}_j.$$

并对 l 维数组 $[(\hat{Z}_1,\hat{Z}'_1),\cdots,(\hat{Z}_l,\hat{Z}'_l)]$ 中的元素进行随机置换(记随机置换为 π'),记置换后的数组为 $[(\hat{Z}_{\pi'(1)},\hat{Z}'_{\pi'(1)}),\cdots,(\hat{Z}_{\pi'(l)},\hat{Z}'_{\pi'(l)})]$.

(3) S_1 计算：

$j\in[1,m']$ 时，$\begin{cases}\hat{z}_j=\hat{t}_j\boldsymbol{A}\cdot\hat{\boldsymbol{B}}_j+r(\hat{t}_j\hat{w}_j+\hat{s})\\ \hat{z}'_j=r(\hat{t}_j\hat{w}_j+\hat{s})\end{cases}$；

$j\in[m'+1,l]$ 时，$\hat{z}_j=r\hat{r}_j$，$\hat{z}'_j=r\hat{s}_j$；或 $\hat{z}_j=r\hat{\alpha}_j$，$\hat{z}'_j=r\hat{\alpha}_j$.

由于在协议执行中，

$$view_1^{\pi}(U,V)=(U,(Z_{\pi(j)},Z'_{\pi(j)}),j\in[1,l],\lambda),$$

而 S_1 在模拟过程中产生的信息序列为

$$S_1(U,f_1(U,V))=(U,(\hat{Z}_{\pi'(j)},\hat{Z}'_{\pi'(j)}),j\in[1,l],\lambda).$$

由于对所有 $j\in[1,l]$，t_j，s 和 r_j，s_j，α_j 均为随机数，所以 $t_j\overset{c}{\equiv}\hat{t}_j$，$s\overset{c}{\equiv}\hat{s}$，$r_j\overset{c}{\equiv}\hat{r}_j$，$s_j\overset{c}{\equiv}\hat{s}_j$，$\alpha_j\overset{c}{\equiv}\hat{\alpha}_j$. 虽然 Alice 有密钥可以解密，但由安全性分析过程可知，Alice 由 $z_{\pi(j)}$，$z'_{\pi(j)}$ 无法得到 Bob 数据的任何信息，对于 Alice 来说，$(Z_{\pi(j)},Z'_{\pi(j)})\overset{c}{\equiv}(\hat{Z}_{\pi'(j)},\hat{Z}'_{\pi'(j)})$. 因此 $\{S_1(U,f_1(U,V))\}_{u_i,v_i\in Q^+}\overset{c}{\equiv}\{view_1^{\pi}(U,V)\}_{u_i,v_i\in Q^+}$.

接收到输入 $(V,f_2(U,V))$ 后，S_2 按以下方式运行：

(1) S_2 任意选择集合 $U'=\{u'_1,\cdots,u'_{n'}\}$，使其满足 $f_2(U',V)=f_2(U,V)$. 并类似协议 2 中向量 $\boldsymbol{A}$ 的构造方式构造集合 U' 对应的向量 $\hat{\boldsymbol{A}}=(\hat{c}_l,\hat{c}_{l-1},\cdots,\hat{c}_1,\hat{c}_0)$.

(2) S_2 选取随机数 $\hat{r}>0$，计算向量

$$\begin{aligned}\hat{\boldsymbol{A}}'&=(\hat{c}_l+\hat{r},\hat{c}_{l-1}+\hat{r},\cdots,\hat{c}_1+\hat{r},\hat{c}_0+\hat{r})\\&=(\hat{c}'_l,\hat{c}'_{l-1},\cdots,\hat{c}'_1,\hat{c}'_0).\end{aligned}$$

并加密随机数 $\hat{r}$，得到密文 $E(\hat{r})$.

(3) (a) 对于 $j\in[1,m]$，S_2 选取随机数 t_j 互不相同，使得 $t_j(\hat{\boldsymbol{A}}'\cdot\boldsymbol{B}_j)>0$；并选取随机数 $s>0$，计算 $\hat{Z}_j$，$\hat{Z}'_j$ 如下：

$$\hat{Z}_j=E[t_j(\hat{\boldsymbol{A}}'\cdot\boldsymbol{B}_j)]E(\hat{r})^s,\ \hat{Z}'_j=t_jw_j+s.$$

(b) 对于 $j\in[m+1,l_1]$，S_2 选取一些随机数对 (r_j,s_j)，$r_j,s_j>0$ 互不相同，计算 $\hat{Z}_j$，$\hat{Z}'_j$ 如下：

$$\hat{Z}_j=E(\hat{r})^{r_j},\ \hat{Z}'_j=s_j.$$

(c) 对于 $j\in[l_1,l]$，S_2 选取随机数对 (α_j,α_j)，$\alpha_j>0$，计算 $\hat{Z}_j$，$\hat{Z}'_j$ 如下：

$$\hat{Z}_j=E(\hat{r})^{\alpha_j},\ \hat{Z}'_j=\alpha_j.$$

共得到 l 个数对 $(\hat{Z}_1,\hat{Z}'_1),\cdots,(\hat{Z}_l,\hat{Z}'_l)$.

(4) S_2 计算：

当 $j\in[1,m]$ 时，$\begin{cases}\hat{z}_j=t_j\hat{\boldsymbol{A}}\cdot\boldsymbol{B}_j+\hat{r}(t_jw_j+s)\\ \hat{z}'_j=\hat{r}(t_jw_j+s)\end{cases}$；

当 $j\in[m+1,l]$ 时，$\hat{z}_j=\hat{r}r_j$，$\hat{z}'_j=\hat{r}s_j$；或 $\hat{z}_j=\hat{r}\alpha_j$，$\hat{z}'_j=\hat{r}\alpha_j$.

由于在协议执行中，

$$view_2^{\pi}(U,V)=(V,\boldsymbol{A}',E(r),f_2(U,V)),$$

而 S_2 在模拟过程中产生的信息序列为

$$S_2(V,f_2(U,V))=(V,\hat{\boldsymbol{A}}',E(\hat{r}),f_2(U',V)).$$

首先，由于 r 为 Alice 任意选取的随机数，根据 Paillier 加密方案的语义安全性，$E(r)\overset{c}{\equiv}E(\hat{r})$；又因为 Bob 没有解密密钥，无法获得随机数 r 的任何信息，因此 $r\overset{c}{\equiv}\hat{r}$，$\boldsymbol{A}'\overset{c}{\equiv}\hat{\boldsymbol{A}}'$. 又由于 $f_2(U,V)=f_2(U',V)$，因此

$$\{S_2(V,f_2(U,V))\}_{u_i,v_i\in Q^+}\overset{c}{\equiv}\{view_2^{\pi}(U,V)\}_{u_i,v_i\in Q^+}.$$

证毕.

两个有理数集合交集势的保密计算. Alice 和 Bob 分别拥有私密有理数集合 $U=\{u_1,\cdots,u_n\}$ 和 $V=\{v_1,\cdots,v_m\}$，他们希望保密计算交集的势 $|U\cap V|$，并使对方无法获知自己集合的任何额外信息(包括要保密集合 U 和 V 的势).

根据协议 2 的设计原理，只需对协议 2 稍作修改即可得到 $|U\cap V|$ 的保密计算协议，简述如下：

协议 3.　有理数集合交集势的保密计算协议.

输入：Alice 输入有理数集合 $U=\{u_1,\cdots,u_n\}$，Bob 输入有理数集合 $V=\{v_1,\cdots,v_m\}$

输出：Alice 输出 $|U\cap V|$

协议的准备工作以及第 1,2 步与协议 2 相同(此时应取 $l_1=l$).

3. (a) 对于每个 $\pi(j)\in[1,l]$，Alice 解密 $Z_{\pi(j)}$ 得到 $z_{\pi(j)}=D(Z_{\pi(j)})$，并计算 $z'_{\pi(j)}=rZ'_{\pi(j)}$. Alice 比较 $z_{\pi(j)}$ 和 $z'_{\pi(j)}$：如果 $z_{\pi(j)}=z'_{\pi(j)}$，令 $d_{\pi(j)}=1$；否则，令 $d_{\pi(j)}=0$.

(b) Alice 计算 $d=\sum_{\pi(j)=1}^{l}d_{\pi(j)}$，并输出 d.

定理 3.　在半诚实模型下协议 3 是正确的和安全的.

定理 3 的证明类似于定理 2，故从略.

3.4　有理数域上集合并集的保密计算

问题描述. Alice 和 Bob 分别拥有私密有理数集合 $U=\{u_1,\cdots,u_n\}$ 和 $V=\{v_1,\cdots,v_m\}$，他们希望保密计算并集 $U\cup V$，并使对方无法获知自己集合的任何额外信息(包括要保密集合 U 和 V 的势).

计算原理. 在下文中，我们主要依据关系式 $U\cup V=U\cup(V\backslash U)$ 进行计算，基本思想是 Alice 设法将集合 V 中不属于 U 的元素添加到 U 中. 因此类似于 3.2 节的计算原理，只是在这里 Alice 和 Bob 需逐个求出属于集合 V 而不属于集合 U 的元素. 下面

协议 4 将沿用 3.2 节计算原理部分的叙述及记号进行描述.

协议 4. 两个有理数集合并集的保密计算协议.

输入:Alice 和 Bob 分别输入私密有理数集合 $U=\{u_1,\cdots,u_n\}$ 和 $V=\{v_1,\cdots,v_m\}$

输出:Alice 输出 $U\cup V$

准备:

(a) Alice 按照式(8)计算集合 U 对应的向量 $\boldsymbol{A}=(c_l,c_{l-1},\cdots,c_1,c_0)$;

(b)对于每一个 $j\in[1,m]$,Bob 类似于协议 1 中 v 对应向量 $\boldsymbol{B}$ 的构造方式,计算 v_j 对应的向量 $\boldsymbol{B}_j=(a_j^l,a_j^{l-1}b_j,\cdots,a_jb_j^{l-1},b_j^l)$ 以及 $w_j=a_j^n+a_j^{n-1}b_j+\cdots+a_jb_j^{n-1}+b_j^n$. 对于每个 $j\in[m+1,l]$,约定 $a_j=b_j=0$,并令 $\boldsymbol{B}_j$ 均为 $l+1$ 维零向量,以及 $w_j=0$.

(c) Alice 生成 Paillier 加密方案的公钥/私钥(pk/sk). 并秘密选取随机数 $r>1$,计算向量

$$\boldsymbol{A}'=(c_l+r,c_{l-1}+r,\cdots,c_1+r,c_0+r)=(c'_l,c'_{l-1},\cdots,c'_1,c'_0).$$

1. Alice 加密 r 和 $r-1$,并将公钥 $pk=(g,N)$,向量 $\boldsymbol{A}'$ 以及密文 $E(r),E(r-1)$ 发送给 Bob.

2. 对于每个 $j\in[1,l]$,Bob 选取互不相同的随机数 $t_j>0$,使得

$$\boldsymbol{A}'\cdot(a_j\boldsymbol{B}_j)+t_j>0,\ a_jw_j+t_j>0,$$
$$\boldsymbol{A}'\cdot(b_j\boldsymbol{B}_j)+t_j>0,\ b_jw_j+t_j>0,$$

并计算:

$$Z_{j1}=E[\boldsymbol{A}'\cdot(b_j\boldsymbol{B}_j)+t_j]E(r)^{a_jw_j+t_j}E(r-1)^{t_j},$$
$$Z_{j2}=E[\boldsymbol{A}'\cdot(a_j\boldsymbol{B}_j)+t_j]E(r)^{b_jw_j+t_j}E(r-1)^{t_j},$$
$$S_j=a_jw_j+b_jw_j+2t_j,$$

Bob 把 l 个数组$(Z_{11},Z_{12},S_1),\cdots,(Z_{l1},Z_{l2},S_l)$随机置换后发送给 Alice.

3. 对于每一个 $j\in[1,l]$,

(a) Alice 解密得到:$z_{j1}=D(Z_{j1}),z_{j2}=D(Z_{j2})$;

(b) Alice 计算 $s_j=rS_j$,并判断 z_{j2} 和 s_j 是否相等,如果 $z_{j2}\neq s_j$,则进一步计算

$$\hat{v}_j=\frac{s_j-z_{j1}}{z_{j2}-s_j}.$$

(c) Alice 按照如下方法构造集合 H:

```
Set H=∅;
For (j=1; j≤l; j++){
    If (z_{j2}≠s_j) H=H∪{v̂_j}}
```

并计算集合 $W=U\cup H$. Alice 输出 W.

3.4.1 协议 4 的正确性

根据计算原理,如果能够证明 $H=V\backslash U$ 即可.

首先,在协议第 3 步,对于每一个 $j\in[1,l]$,根据 Paillier 加密方案的加法同态性,Alice 解密得到

$$z_{j1}=b_j\boldsymbol{A}\cdot\boldsymbol{B}_j+r(b_jw_j+a_jw_j+2t_j),$$
$$z_{j2}=a_j\boldsymbol{A}\cdot\boldsymbol{B}_j+r(a_jw_j+b_jw_j+2t_j),$$

并计算 $s_j=rS_j=r(a_jw_j+b_jw_j+2t_j)$.

当 $z_{j2}\neq s_j$ 时,进一步计算:

$$\hat{v}_j=\frac{s_j-z_{j1}}{z_{j2}-s_j}=-\frac{b_j}{a_j},$$

根据式(5)中 a,b 表达式的含义(这里 a_j,b_j 有同样的含义),$\hat{v}_j$ 也表达了 v_j 对应直线的斜率,因此 $\hat{v}_j=v_j$.

下面只需再证明:对于 $j\in[1,m]$,当且仅当 $z_{j2}\neq s_j$ 成立时,$v_j\in V\backslash U$.

(1) 当 $z_{j2}\neq s_j$ 时,可知 $a_j\boldsymbol{A}\cdot\boldsymbol{B}_j\neq 0$,即 $a_j\neq 0$,且 $\boldsymbol{A}\cdot\boldsymbol{B}_j\neq 0$,因此有 $j\in[1,m]$. 进一步由 $\boldsymbol{A}\cdot\boldsymbol{B}_j\neq 0$ 可知,$v_j\notin U$. 因此 $v_j\in V\backslash U$.

(2) 当 $z_{j2}=z_{j1}=s_j$ 时,有 $b_j\boldsymbol{A}\cdot\boldsymbol{B}_j=a_j\boldsymbol{A}\cdot\boldsymbol{B}_j=0$. 如果 $j\in[1,m]$,由于 $a_j\neq 0$,则必有 $\boldsymbol{A}\cdot\boldsymbol{B}_j=0$,即有 $v_j\in U$. 如果 $j\in[m+1,l]$,按照约定 $a_j=b_j=0$,不存在相应的 $v_j\in V$.

(3) 最后证明 $z_{j2}=s_j$ 与 $z_{j1}\neq s_j$ 不可能同时成立. 若不然,假设它们同时成立,则必有

$$a_j\boldsymbol{A}\cdot\boldsymbol{B}_j=0,\ b_j\boldsymbol{A}\cdot\boldsymbol{B}_j\neq 0,$$

显然,这需要 $a_j=0,b_j\neq 0,\boldsymbol{A}\cdot\boldsymbol{B}_j\neq 0$ 同时成立. 当 $j\in[1,m]$ 时,$a_j=0$ 与 a_j,b_j 的定义要求矛盾(a_j,b_j 与式(5)中 a,b 的定义性质相同);当 $j\in[m+1,l]$ 时,根据前面的约定有 $a_j=0,b_j=0$ 与 $a_j=0,b_j\neq 0$ 相矛盾. 因此,$z_{j2}=s_j$ 与 $z_{j1}\neq s_j$ 不可能同时成立.

综上所述,已经证明了当 $z_{j2}\neq s_j$ 时,有 $v_j\in V\backslash U$,即 $H=V\backslash U$,$W=U\cup H=U\cup V$,因此协议 4 是正确的.

3.4.2 协议 4 的安全性

首先,Alice 和 Bob 共同商议一个正整数 l,满足 $l>m,l>n$. 类似于协议 2 中关于 Alice 集合势的安全性分析,可知协议 4 中 Alice 集合势 $|U|=n$ 是安全的.

下面考虑 Alice 集合元素的安全性. 在协议执行中,Alice 仅给 Bob 发送了向量 $\boldsymbol{A}'$ 及密文 $E(r-1)$,$E(r)$,由于 Bob 没有私钥,因此无法获得随机数 r 的任何消息. 又根据向量 $\boldsymbol{A}'$ 的定义可知,Bob 仅能得到向量 $\boldsymbol{A}$ 中任意两个分量 c_i,c_j 间的关系式 $c_i-c_j=c'_i-c'_j$,完全类似于协议 2 中关于 Alice 数据的安全性分析,可知协议 4 中 Alice 集合元素是安全的.

关于 Bob 集合的安全性分析如下. 首先分析 Bob 集合元素的安全性. 在协议中 Bob 仅给 Alice 发送了 l 个经过随机置换的数组$(Z_{11},Z_{12},S_1),\cdots,(Z_{l1},Z_{l2},S_l)$. 对于每个 $j\in[1,l]$,Alice 解密并计算可得到:

$$\begin{aligned}&z_{j1}=b_j\boldsymbol{A}\cdot\boldsymbol{B}_j+r(b_jw_j+a_jw_j+2t_j),\\&z_{j2}=a_j\boldsymbol{A}\cdot\boldsymbol{B}_j+r(a_jw_j+b_jw_j+2t_j),\\&s_j=r\,S_j=r(a_jw_j+b_jw_j+2t_j)\end{aligned}\quad(12)$$

由于 t_j 是 Bob 选取的秘密随机数，Alice 由上面单个独立式子均无法获得关于 a_j, b_j, w_j 的任何信息. 如果利用上面三式进行联立求解，可得到

$$b_j\boldsymbol{A}\cdot\boldsymbol{B}_j=z_{j1}-s_j,\ a_j\boldsymbol{A}\cdot\boldsymbol{B}_j=z_{j2}-s_j.$$

如果 $a_j\boldsymbol{A}\cdot\boldsymbol{B}_j\neq 0$，由上面两式进一步可求得

$$v_j=\frac{s_j-z_{j1}}{z_{j2}-s_j}=-\frac{b_j}{a_j}.$$

根据协议 4 的正确性分析可知，此时 $v_j\notin U$，且 $v_j\in U\cup V$，即 Alice 根据协议的输出也可获得 v_j，因此无额外的信息泄露.

如果 $a_j\boldsymbol{A}\cdot\boldsymbol{B}_j=0$，这时也必有 $b_j\boldsymbol{A}\cdot\boldsymbol{B}_j=0$，这意味着有下面两种情形：(1) $a_j=b_j=0$，(2) $a_j\neq 0$，$\boldsymbol{A}\cdot\boldsymbol{B}_j=0$之一将会发生. 在情形(1)发生时，$\boldsymbol{B}_j$ 为添加的零向量，并无相应的 $v_j\in V$ 与之对应；而当情形(2)发生时，有 $v_j\in U\cap V$ 与之对应. 由于在协议中 Bob 是将 l 个数组随机置换后发送给 Alice 的，因此 Alice 无法区分到底是两种情形中哪种情形发生，故也无法获知 Bob 数据的任何额外信息.

进一步分析，由于 $t_1,\cdots,t_l$ 是 Bob 选择的 l 个互不相同的独立随机数，对于得到的形如式(12)的所有 $3l$ 个关系式，Alice 无论以何种方式(将不同的 $j=1,\cdots,l$ 对应的式子)组合求解，同样无法得到 Bob 数据的任何额外信息. 所以 Bob 集合元素是安全的.

协议 4 也保证了 Bob 集合势的私密性，具体分析如下：

在协议执行中 Bob 针对每一个 $j\in[1,l]$，对 $\boldsymbol{B}_j$，w_j 进行了完全相同的运算，并对 l 个数组 $(Z_{11},Z_{12},S_1),\cdots,(Z_{l1},Z_{l2},S_l)$ 经过随机置换后发送给 Alice，当 $z_{j2}=s_j$ 时，Alice 无法区分两种情形((1) $a_j=b_j=0$，不存在 $v_j\in V$ 与之对应以及(2) $a_j\neq 0$，$\boldsymbol{A}\cdot\boldsymbol{B}_j=0$，有 $v_j\in U\cap V$ 与之对应)中哪种情形发生. 因此在协议中 Bob 构造 l 个数组，并对其进行随机置换完全保证了集合 V 的势 $|V|$ 的私密性(如果不进行随机置换，Alice 顺序解密及计算各数组 $(Z_{11},Z_{12},S_1),\cdots,(Z_{l1},Z_{l2},S_l)$，由于此时 $j\in[m+1,l]$ 时均有 $z_{j2}=s_j$，这样 Alice 可能猜出 V 的势为 m).

定理 4.　在半诚实模型下协议 4 是安全的.

定理 4 的证明类似于定理 2，在此省略.

3.5　有理数域上集合包含关系的保密判定

问题描述. Alice 和 Bob 分别拥有私密有理数集合 $U=\{u_1,\cdots,u_n\}$ 和 $V=\{v_1,\cdots,v_m\}$. 他们希望保密判定集合 V 是否包含于集合 U 中，而不泄露 U 和 V 的任何额外信息(包括要保密 U 和 V 的势).

计算原理. 关于集合包含问题的解决思路和计算原理与计算集合交集非常类似，参看 3.3 节，这里不再赘述. 仅有的区别在于保护 V 的势的隐私性方面做法有所不同.

关于集合包含问题，为了隐藏集合 V 的势，对于每一个 $j\in[m+1,l]$，约定 $a_j=b_j=0$，并令 $\boldsymbol{B}_j$ 均为 $l+1$ 维零向量，以及 $w_j=0$. 因此有

$$\begin{aligned}V\subset U&\Leftrightarrow\forall j\in[1,m],\ \boldsymbol{A}\cdot\boldsymbol{B}_j=0\\&\Leftrightarrow\forall j\in[1,l],\ \boldsymbol{A}\cdot\boldsymbol{B}_j=0\\&\Leftrightarrow(\boldsymbol{A}\cdot\boldsymbol{B}_1,\cdots,\boldsymbol{A}\cdot\boldsymbol{B}_l)=0.\end{aligned}$$

这里 $l,\boldsymbol{A},\boldsymbol{B}_j\ (j\in[1,m])$ 等符号的含义与 3.3 节相同.

根据上述分析可知，保密判定集合包含问题即可转化为保密判定向量 $(\boldsymbol{A}\cdot\boldsymbol{B}_1,\cdots,\boldsymbol{A}\cdot\boldsymbol{B}_l)$ 是否为零向量. 为了解决两个向量是否相等的判定问题，给出下面命题(证明很简单，故省略).

命题 1.　对于任意两个向量 $\boldsymbol{X}=(x_1,\cdots,x_l)$ 和 $\boldsymbol{Y}=(y_1,\cdots,y_l)$，$\boldsymbol{X}=\boldsymbol{Y}$ 的充要条件是下面等式成立：

$$|\boldsymbol{X}|^2+|\boldsymbol{Y}|^2=2\boldsymbol{X}\cdot\boldsymbol{Y}\tag{13}$$

其中 $|\boldsymbol{X}|^2=x_1^2+\cdots+x_l^2$，$|\boldsymbol{Y}|^2=y_1^2+\cdots+y_l^2$.

为叙述方便，定义二元谓词：

$$P(U,V)=\begin{cases}1, & \text{如果 } V\subset U\\ 0, & \text{如果 } V\not\subset U\end{cases}.$$

协议 5.　有理数域上集合包含关系的保密判定协议.

输入：Alice 和 Bob 分别输入私密有理数集合 $U=\{u_1,\cdots,u_n\}$ 和 $V=\{v_1,\cdots,v_m\}$

输出：Alice 输出 $P(U,V)$

准备：Alice 按照式(8)计算集合 U 对应的向量 $\boldsymbol{A}=(c_l,c_{l-1},\cdots,c_1,c_0)$. 而 Bob 类似于式(8)中向量 $\boldsymbol{B}$ 的构造方式，对于每一个 $j\in[1,m]$，构造元素 v_j 的对应向量 $\boldsymbol{B}_j=(a_j^l,a_j^{l-1}b_j,\cdots,a_jb_j^{l-1},b_j^l)$，并记 $w_j=a_j^l+a_j^{l-1}b_j+\cdots+a_jb_j^{l-1}+b_j^l$. 对于每个 $j\in[m+1,l]$，约定 $a_j=b_j=0$，并令 $\boldsymbol{B}_j$ 为 $l+1$ 维零向量，$w_j=0$.

1. Alice 生成 Paillier 加密方案的公钥/私钥(pk/sk). 并选取随机数 $r>0$，计算向量

$$\boldsymbol{A}'=(c_l+r,c_{l-1}+r,\cdots,c_1+r,c_0+r)=(c_l',c_{l-1}',\cdots,c_1',c_0').$$

Alice 加密随机数 r. 将公钥 $pk=(g,N)$，向量 $\boldsymbol{A}'$ 和密文 $E(r)$ 发送给 Bob.

2. (a) Bob 选取互不相同的正随机数 $t_1,\cdots,t_l$，构造向量 $\boldsymbol{Z}=(Z_1,\cdots,Z_l)$ 和 $\boldsymbol{S}=(S_1,\cdots,S_l)$，其中

$$Z_j=t_j(\boldsymbol{A}'\cdot\boldsymbol{B}_j),\ S_j=t_jw_j.$$

(b) Bob 选取随机数 t，使得 $-2\boldsymbol{Z}\cdot\boldsymbol{S}+t>0$，计算 $H=E(|\boldsymbol{Z}|^2)E(r)^{-2\boldsymbol{Z}\cdot\boldsymbol{S}+t}$. Bob 把 $t,|\boldsymbol{S}|^2$ 以及 H 发送给 Alice.

3. Alice 解密得到 $h=D(H)$，并计算

$$h'=h+r^2|\boldsymbol{S}|^2-rt.$$

如果 $h'=0$，Alice 输出 $P(U,V)=1$；否则，Alice 输出 $P(U,V)=0$.

3.5.1 协议 5 的正确性

由协议第 2(a)步可知

$$Z_j=t_j(\boldsymbol{A}'\cdot\boldsymbol{B}_j)=t_j\boldsymbol{A}\cdot\boldsymbol{B}_j+rS_j.$$

根据 Paillier 加密方案的加法同态性，在协议第 3 步的计算中，Alice 得到

$$h=D(H)=|\boldsymbol{Z}|^2-2\boldsymbol{Z}\cdot(r\boldsymbol{S})+rt,$$

$$h'=h+r^2|\boldsymbol{S}|^2-rt=|\boldsymbol{Z}|^2+|r\boldsymbol{S}|^2-2\boldsymbol{Z}\cdot(r\boldsymbol{S}),$$

根据命题 1，对每一个 $j\in[1,l]$，由于 $t_j\neq0$，因此

$$\begin{aligned}h'=0&\Leftrightarrow\boldsymbol{Z}=r\boldsymbol{S}\\&\Leftrightarrow\forall j\in[1,l],\ \boldsymbol{A}\cdot\boldsymbol{B}_j=0\\&\Leftrightarrow V\subset U.\end{aligned}$$

正确性得证.

3.5.2 协议 5 的安全性

关于 Alice 集合 U 的势的安全性分析与协议 2 完全相同. 下面考虑 Alice 集合数据的安全性.

协议中 Bob 仅收到 Alice 发送的向量 $\boldsymbol{A}'$ 和 $E(r)$，由于 Bob 没有私钥解密，无法获得随机数 r 的任何信息. 根据向量 $\boldsymbol{A}'$ 的定义可知，Bob 仅能得到向量 $\boldsymbol{A}$ 中任意两个分量 c_i,c_j 的关系式 $c_i-c_j=c_i'-c_j'$，完全类似于协议 2 中关于 Alice 数据的安全性分析，可知 Alice 集合元素是安全的.

关于集合 V 的势的安全性分析如下：

Bob 对每个 $j\in[m+1,l]$，约定 $a_j=b_j=0$，并令 $\boldsymbol{B}_j$ 均为 $l+1$ 维零向量，这样即使得向量 $(\boldsymbol{A}\cdot\boldsymbol{B}_1,\cdots,\boldsymbol{A}\cdot\boldsymbol{B}_m)=\boldsymbol{0}$ 与向量 $(\boldsymbol{A}\cdot\boldsymbol{B}_1,\cdots,\boldsymbol{A}\cdot\boldsymbol{B}_l)=\boldsymbol{0}$ 完全等价，如此既保证了协议的正确性，又隐藏了 $|V|=m$ 的私密信息.

最后考虑 Bob 集合数据的安全性. 在协议第 2 步 Alice 得到 Bob 发送的 $t,|\boldsymbol{S}|^2$ 以及 H，Alice 解密得到 $h=D(H)$，进一步计算可到下面信息：

$$h-rt+r^2|\boldsymbol{S}|^2=(t_1\boldsymbol{A}\cdot\boldsymbol{B}_1)^2+\cdots+(t_l\boldsymbol{A}\cdot\boldsymbol{B}_l)^2,$$

$$|\boldsymbol{S}|^2=(t_1w_1)^2+\cdots+(t_lw_l)^2,$$

由于 $t_1,\cdots,t_l$ 是 Bob 的私有数据，Alice 从上面两式得不到 Bob 集合元素的任何额外信息. 可知 Bob 集合元素是安全的. 因此，协议 5 是安全的.

关于协议 5 的安全性有下面的定理 5，定理 5 的严格证明类似于定理 2，在此省略.

定理 5. 有理数域上集合包含关系的保密判定协议 5 是安全的.

4 有理点集合问题的保密计算

随着信息技术的日益发展，与坐标定位有关的应用软件得到广泛使用，这一方面给人们生活带来很多便利，另一方面也对私密位置的信息泄露等安全性问题提出挑战. 因此，研究与点的位置相关的保密计算问题具有重要意义.

如果点 $P=(x,y)$ 的两个坐标 x,y 均为有理数，称 $P=(x,y)$ 为有理点（特殊地，如果 x,y 均为整数，称 $P=(x,y)$ 为整数点）. 目前，还未见到关于有理点和有理点集合关系问题的保密计算研究. 本节针对这个新的保密计算问题，提出一个全新的编码方案以及相应的转化思想，并以此为基础设计有理点和有理点集合关系的保密判定协议.

4.1 有理点编码方案

对于任意一个有理点 $p=(b/a,d/c)$，其中 $\gcd(a,b)=1,\gcd(c,d)=1$，利用哥德尔编码方法进行编码，即将有理点 $p=(b/a,d/c)$ 直接编码为有理数 $\bar{p}=(2^b7^d)/(5^a3^c)$.

例如，对于有理点 $(7/4,3/5)$，根据上述编码方案，可以将其编码为有理数 $(2^77^3)/(5^43^5)$.

命题 2. 有理点 $p=(b/a,d/c)$ 与有理数 $\bar{p}=(2^b7^d)/(5^a3^c)$ 一一对应.

证明. 要证明有理点 p 与编码有理数 $\bar{p}$ 一一对应，显然只需证明：如果编码有理数相同，则其对应的两个有理点也相同.

现假设存在另一个有理点 $p_1=(b_1/a_1,d_1/c_1)$，其中 $\gcd(a_1,b_1)=1,\gcd(c_1,d_1)=1$，其对应的编码有理数为 $\bar{p}_1=(2^{b_1}7^{d_1})/(5^{a_1}3^{c_1})$. 我们需要证明如果 $\bar{p}_1=\bar{p}$，则有 $p_1=p$，即 $(b_1/a_1,d_1/c_1)=(b/a,d/c)$.

由 $\bar{p}_1=\bar{p}$ 得到 $(2^{b_1}7^{d_1})/(5^{a_1}3^{c_1})=(2^b7^d)/(5^a3^c)$. 因为 2,3,5,7 均为素数，所以 $\gcd(2^{b_1}7^{d_1},5^{a_1}3^{c_1})=1$，$\gcd(2^b7^d,5^a3^c)=1$，即有 $2^{b_1}7^{d_1}=2^b7^d,5^{a_1}3^{c_1}=5^a3^c$. 进一步，根据 $2^{b_1}7^{d_1}=2^b7^d$ 易知 $b_1=b,d_1=d$；同理可知 $a_1=a,c_1=c$. 故有 $(b_1/a_1,d_1/c_1)=(b/a,d/c)$.

综上所述，有理点编码方案保证了任意有理点 $p=(b/a,d/c)$ 与唯一有理数 $\bar{p}=(2^b7^d)/(5^a3^c)$ 对应. 证毕.

以上述有理点编码方案为基础，下面研究有理点与有理点集合关系的保密判定问题.

4.2 有理点与有理点集合关系的保密判定

问题描述. Alice 拥有一个私密的有理点集合 $Z=\{z_1,\cdots,z_n\}$，Bob 拥有一个私密的有理点 z_0，其中 $z_i=(x_i,y_i),i\in[0,n]$. Alice 和 Bob 希望保密判定 z_0 是否属于 Z，同时不泄露 z_0 和 Z 的任何额外信息.

计算原理. 首先，应用有理点编码方案，Alice 将 Z 编码为有理数集合 $\bar{Z}$（$\bar{Z}$ 中的有理数均由集合

Z 中的有理点编码得到），Bob 将有理点 z_0 编码为有理数 $\bar{z}_0$，根据命题 2，有理点与有理点集合关系的保密判定问题即可转化为有理数域上元素与集合关系的保密判定问题，即有

$$z_0 \in Z \Leftrightarrow \bar{z}_0 \in \bar{Z},$$

因此，可以通过调用协议 1 设计协议.

定义函数 $u=f(z_0,Z)$：当 $z_0 \in Z$ 时，令 $u=1$；否则，令 $u=0$.

协议 6.　有理点与有理点集合关系判定协议.

输入：Alice 输入集合 $Z=\{z_1,\cdots,z_n\}$，Bob 输入 z_0，其中 $z_i=(x_i,y_i)(x_i,y_i \in Q^+, i \in [0,n])$

输出：$u=f(z_0,Z)$

准备：根据计算原理，Alice 和 Bob 分别将有理点集合 Z 和有理点 z_0 编码为有理数集合 $\bar{Z}$ 和有理数 $\bar{z}_0$.

1. Alice 和 Bob 将 $\bar{Z}$ 和 $\bar{z}_0$ 作为协议 1 的输入，调用协议 1，得到 $P(\bar{z}_0,\bar{Z})$.

2. 如果 $P(\bar{z}_0,\bar{Z})=1$，输出 $u=1$；否则，输出 $u=0$.

4.2.1　协议 6 的正确性和安全性

根据协议 6 的计算原理以及协议 1 的正确性，易知协议 6 是正确的.

在协议的准备阶段，Alice 和 Bob 分别将有理点集合 Z 及有理点 z_0 编码为有理数集合 $\bar{Z}$ 和有理数 $\bar{z}_0$. 在这个编码过程中没有信息传递，因此，$\bar{Z}$ 和 $\bar{z}_0$ 分别为 Alice 和 Bob 的私密数据.

协议 6 的执行过程完全是以 Alice 和 Bob 的私密数据 $\bar{Z}$ 和 $\bar{z}_0$ 作为协议 1 的输入，调用协议 1 完成的，根据协议 1 的安全性，Alice 和 Bob 的数据 $\bar{Z}$ 和 $\bar{z}_0$ 是完全安全的，因此，相应的有理点集合 Z 以及有理点 z_0 也是安全的.

关于协议 6 的安全性，有下面的定理 6，其证明过程完全类似于定理 1，在此省略.

定理 6.　有理点与有理点集合关系的保密判定协议 6 是安全的.

注解 1. 命题 2 证明了本节给出的有理点编码方案能够将有理点 p 与有理数 $\bar{p}$ 一一对应，进而可以将有理数域内关于点与点集关系问题转化为相应的有理数与集合关系问题，从而调用协议 1 得到解决，这表明该编码方案在理论上是可行的. 由于选取小素数 2,3,5,7 将有理点 $p=(b/a,d/c)$ 编码成有理数 $\bar{p}=(2^b7^d)/(5^a3^c)$，当 a,b,c,d 较小时（比如不超过 10^3），直接计算 $\bar{p}$ 较简单；当 a,b,c,d 较大时，可对 $2^b7^d,5^a3^c$ 加上适当的模运算以简化计算. 因此本节的有理点编码方式在实际中也是可行的.

注解 2. 在前面以协议 1 为基础设计构造了关于有理数集合的各种基本运算协议（协议 2～协议 5）. 基于协议 6 可以类似构造关于有理点集合的各种基本运算协议.

5　性能分析

下面主要对本文所设计协议的性能和执行效率进行分析. 本文关于元素与集合关系的保密判定协议（协议 1）和两集合并集的保密计算协议（协议 4）保证了参与者集合的势是私密的，而目前已有的相关协议基本上都没有考虑集合势的保密性（如文献[15-16,24-27]），本文协议 6 首次研究了有理点与有理点集合关系的保密判定问题，目前未见到相关研究. 因此下面对协议 1、协议 4 与协议 6 仅进行复杂性分析，而将本文协议 2、协议 3（集合交集及其势的保密计算协议）以及协议 5（集合包含关系的保密判定协议）与目前已有的效率较高的相关协议从效率以及适用范围等方面进行分析比较. 本文协议主要以 Paillier 加密方案为基础，所比较的有关协议主要是基于离散对数、哈希运算以及 ElGamal 加密方案进行设计. 为了方便分析，在分析计算复杂性时，只考虑协议执行中最费时的模指数运算，其他花费忽略不计，并应用执行协议所需要的通信轮数来衡量协议的通信复杂性. 在下面分析中，假设 Alice 集合的势为 n，Bob 集合的势为 m，若需要全集时，设定全集的势为 η.

5.1　效率分析与比较

计算复杂性分析. 本文的协议 1～协议 6 主要以 Paillier 加密方案为基础进行设计，每进行一次 Paillier 加密或解密都需要 2 次模指数运算.

协议 1（协议 6）中 Bob 加密 $r,\boldsymbol{B}'$ 以及解密 S 需要进行 $2\times(l+3)$ 次模指数运算，Alice 计算密文 S 需要进行 $l+2$ 次模指数运算，协议 1（协议 6）共需要 $3l+8$ 次模指数运算. 这里 l 是为隐藏集合 U 的势选取的大于 $|U|$ 的一个数值，不需要取得太大，能够保证协议 1 具有线性复杂性.

协议 2（协议 3）中 Alice 执行了 1 次加密及 l 次解密运算，需要 $2\times(l+1)$ 次模指数运算，Bob 计算密文需要 $2m+l$ 次模指数运算，所以协议 2（协议 3）需要 $3l+2m+2$ 次模指数运算.

协议 4 中 Alice 执行了 2 次加密及 $2l$ 次解密运算，需要 $4\times(l+1)$ 次模指数运算，Bob 计算密文需要 $7l$ 次模指数运算，所以协议 4 需要 $11l+4$ 次模指数运算. 与协议 1 类似，l 的取值能保证协议 4 具有线性复杂性.

协议 5 中 Alice 需进行 1 次加密及 1 次解密，需要 4 次模指数运算，Bob 计算密文 H 需 3 次模指数运算，协议 5 共需要 7 次模指数运算.

文献[21]基于 Paillier 加密方案设计了集合交集的保密计算协议，其中 Alice 和 Bob 分别需要执行 $2(n+m+1)$次和 nm 次模指数运算，文献[21]共需要执行 $2(n+m+1)+nm$ 次模指数运算.

文献[22]基于离散对数和单向散列函数设计了集合交集势的保密计算协议，其中 Alice 和 Bob 分别需要执行 $2n+1$ 次和 $n+m+2$ 次模指数运算，共需执行 $3n+m+3$ 次模指数运算.

文献[31]通过设定全集进行编码，基于 ElGamal 加密方案设计了集合包含关系的保密判定协议，其中 Alice 和 Bob 分别需要执行 $2\eta+1$ 次和 2 次模指数运算，共需执行 $2\eta+3$ 次模指数运算.

通信复杂性分析. 本文协议 1 和协议 6 各需要 2 轮通信，协议 2～协议 5 均只需要 1 轮通信. 文献[21-22,31]各需要 1 轮通信.

本文所设计协议与目前已有的较好协议的效率以及适用性比较如表 1 所示.

表 1 协议的效率分析和适用范围比较

文献	计算功能	计算复杂性	通信复杂性	适用范围
本文协议 2	集合交集	$3l+2m+2$	2	有理数
本文协议 3	集合交集势	$3l+2m+2$	1	有理数
本文协议 5	集合包含关系	7	1	有理数
文献[21]	集合交集	$2(n+m+1)+nm$	1	整数
文献[22]	集合交集势	$3n+m+3$	1	整数
文献[31]	集合包含关系	$2\eta+3$	1	整数

在表 1 中，计算功能一栏为各协议研究的具体内容，计算复杂性是以协议所需的模指数运算次数表示，通信复杂性为协议所需的通信轮数. 文献[31]需要设定全集，η 表示全集的势，为了隐私数据的保密性，一般要求 η 取值较大. 根据分析可知，本文协议 1～协议 4 及协议 6 均具有线性计算复杂性，协议 5 的计算复杂性仅为常数 7. 因此，本文协议与所提到的整数集合的保密计算协议相比计算效率都有所提高，同时本文协议能够解决有理数集合的相关问题，且不需要集合元素取自某个全集，由此知本文协议是高效的且具有广泛的应用范围.

5.2 协议效率实验

本小节对有关协议进行实验测试，并将本文协议的执行结果与已有效率较高的协议的执行结果进行比较(本文协议 2、协议 3 以及协议 5 分别与文献[21]、[22]以及文献[31]进行比较).

(1)实验平台. 计算机的配置如下：操作系统为 Windows10 企业版，Intel(R) Core(TM) i5-6600 CPU@3.30 GHz，安装内存 8.00 GB，64 位操作系统. 采用 Java 编程语言在 MyEclipse 上对协议分别进行了编程实现，在此约定本文所做模拟实验均在此环境下进行.

(2)实验结果. 本文协议 2、协议 3 和协议 5 及文献[21]基于 Paillier 加密方案进行设计，文献[22]和[31]分别基于离散对数和 ElGamal 加密方案进行设计，下面分别在相应环境下进行仿真实验.

实验设定 Paillier 加密方案中使用的大素数 p 和 q 的位数为 512 bits，即设定模数 $N=pq$ 的位数为 1024 bits，为了保证实验结果的公平性，同样设定离散对数运算和 ElGamal 加密方案中的模数 N 的位数为 1024 bits，且统一限定保密数据的范围为(0,50]. 下面分别对本文协议 2、3、5 和文献[21-22,31]的协议进行实际计算，并保持集合的势一致，对每个协议进行多次实验，实验结果随机抽取 10 次数据求取平均值. 实验结果如表 2 所示.

表 2 实验结果分析

文献	实验耗时/ms
协议 2	189.13
协议 3	176.97
协议 5	18.21
文献[21]	619.30
文献[22]	211.36
文献[31]	229.58

由表 2 可知，本文协议 2、3、5 的效率较高，优势明显.

为了测试本文协议的执行效率与集合势的关系，我们对本文协议进行了多维度的实验测试. 由于协议 2 至协议 5 均是以协议 1 为基础进行设计，在此仅对协议 2 进行多维度分析. 下面对协议 2 和文献[21]进行实际计算，并对不同维度(集合势)进行多次实验求取平均值，实验结果如图 2 所示.

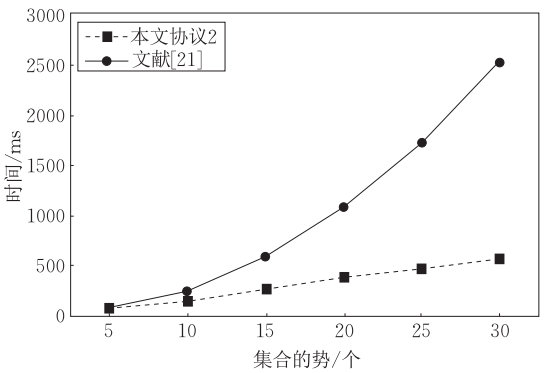

图 2 不同维度下的实验结果分析

由图 2 可知，本文协议 2 和文献[21]的执行时间都随参与方集合势的增长而增长，本文协议 2 趋于线性增长，增长幅度缓慢；文献[21]的增长幅度很大，当集合的势很大时，文献[21]的协议需要的时间变得非常大，本文协议 2 与之相比要小得多.

根据前面的理论分析，如果两方集合的势均为 m，在协议 2 中 l 不是太大时(和 m 线性相关即可)，协议 2 具有线性计算复杂性，而文献[21]中协议具有二次计算复杂性，实验结果与理论分析结果是一致的. 因此，协议 2 的计算效率较高，优势明显.

6　协议的推广应用

下面举例说明本文协议的一些推广应用.

坐标点与直线位置关系的保密判定

问题描述. 假设 Alice 有一个直线集合 U，它由 n 条直线 $L_i(i\in[1,n])$ 构成(直线斜率均为有理数)，Bob 有一个有理点 $p(x,y)$，Alice 和 Bob 要保密判定点 $p(x,y)$ 与 n 条直线具有何种位置关系：(1)点 p 至少位于 U 中一条直线上，这时称 p 与 U 相关；(2)点 p 不在 U 中任一条直线上，这时称 p 与 U 不相关.

计算原理. 类似于协议 1 的基本思想，首先，对于每一个 $i\in[1,n]$，Alice 在直线 L_i 上选两个不同的点 $q_{i1}=(x_{i1},y_{i1})$，$q_{i2}=(x_{i2},y_{i2})(x_{i2}>x_{i1})$. 当 p 与 U 相关时，$\exists i\in[1,n]$ 使得 $S_{\Delta pq_{i1}q_{i2}}=0$；当 p 与 U 不相关时，$\forall i\in[1,n]$ 都有 $S_{\Delta pq_{i1}q_{i2}}\neq 0$. 由于点 q_{i1}，q_{i2} 在同一条直线上，根据面积式(3)计算得到：

$$S_{\Delta pq_{i1}q_{i2}}=\frac{1}{2}|a_iy+b_ix|,$$

上式中 $a_i=x_{i2}-x_{i1}>0$，$b_i=y_{i1}-y_{i2}$.

因此，p 与 U 相关$\Leftrightarrow\prod_{i=1}^{n}(a_iy+b_ix)=0$.

下文中，记

$$\begin{aligned}Z&=\prod_{i=1}^{n}(a_iy+b_ix)\\&=y^nc_n+y^{n-1}xc_{n-1}+\cdots+yx^{n-1}c_1+x^nc_0,\end{aligned}$$

其中 $c_k=\sum_{i_1,\cdots,i_k\in[1,n]}\left(a_{i_1}\cdots a_{i_k}\prod_{j\neq i_1,\cdots,i_k}b_j\right)$，$0<k\leq n$.

进一步，定义向量 $\boldsymbol{A}$ 和 $\boldsymbol{B}$：

$$\boldsymbol{A}=(c_n,c_{n-1},\cdots,c_1,c_0),$$

$$\boldsymbol{B}=(y^n,y^{n-1}x,\cdots,yx^{n-1},x^n),$$

因此得到 $Z=\prod_{i=1}^{n}(a_iy+b_ix)=\boldsymbol{A}\cdot\boldsymbol{B}$，即有

$$\begin{aligned}p\text{ 与 }U\text{ 相关}&\Leftrightarrow\prod_{i=1}^{n}(a_iy+b_ix)=0\\&\Leftrightarrow\boldsymbol{A}\cdot\boldsymbol{B}=0.\end{aligned}$$

根据上述等价关系可将点 p 与集合 U 的相关性问题转化为向量 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的内积是否为 0 的判定问题，利用协议 1 的设计思路可设计保密计算协议.

7　结　论

为了解决关于有理数集合的保密计算问题，本文首先提出一种新的转化思想，将任意有理数编码为坐标系中一条过原点的直线，并结合三角形面积公式，将有理数域上元素与集合关系问题转化为整数范围内向量内积问题，设计构造了有理数与有理数集合关系的保密判定协议 1. 以协议 1 为基础，结合各种有理数集合运算的特点，解决了有理数域上各种集合保密计算问题. 其次，结合哥德尔编码设计了新的有理点编码方案，应用该编码方案，设计构造了关于有理点集合问题的保密计算协议. 应用模拟范例严格证明了这些协议的安全性，理论分析和实验测试表明本文方案的高效性. 本文协议是在半诚实模型下设计的，未来将进一步研究恶意模型下集合问题的保密计算.

参　考　文　献

[1] Yao A C. Protocols for secure computations//Proceedings of the 23rd IEEE Symposium on Foundations of Computer Science. Chicago, USA, 1982: 160-164

[2] Goldwasser S. Multi-party computations: Past and present//Proceedings of the 16th Annual ACM Symposium on Principles of Distributed Computing. Santa Barbara, USA, 1997: 1-6

[3] Cramer R, Damgard I B, Nielsen J B. Secure Multiparty Computation. London, UK: Cambridge University Press, 2015

[4] Goldreich O, Micali S, Wigderson A. How to play any mental game//Proceedings of the 19th Annual ACM Symposium on Theory of Computing. New York, USA, 1987: 218-229

[5] Goldreich O. The Fundamental of Cryptography: Basic Applications. London, UK: Cambridge University Press, 2004

[6] Tang C M, Shi G H, Yao Z A. Secure multi-party computation protocol for sequencing problem. Science China Information Sciences, 2011, 54(8): 1654-1662

[7] Toft T. Sub-linear, secure comparison with two non-colluding parties//Proceedings of the 14th International Conference on Practice and Theory in Public Key Cryptography. Taormina, Italy, 2011: 174-191

[8] Fong P K, Weber-Jahnke J H. Privacy preserving decision tree learning using unrealized data sets. IEEE Transactions on Knowledge and Data Engineering, 2012, 24(2): 353-364

[9] Yang Jing, Zhao Jia-Shi, Zhang Jian-Pei. A privacy preservation method for high dimensional data mining. Acta Electronica Sinica, 2013, 41(11): 2187-2192(in Chinese)
(杨静, 赵家石, 张健沛. 一种面向高维数据挖掘的隐私保护方法. 电子学报, 2013, 41(11): 2187-2192)

[10] Li S D, Wu C Y, Wang D S, et al. Secure multiparty computation of solid geometric problems and their applications. Information Sciences, 2014, 282: 401-413

[11] Guo Yi-Min, Zhou Shu-Fang, Dou Jia-Wei, et al. Efficient privacy-preserving interval computation and its applications. Chinese Journal of Computers, 2017, 40(7): 1-16(in Chinese)
(郭奕旻, 周素芳, 窦家维等. 高效的区间保密计算及应用. 计算机学报, 2017, 40(7): 1-16)

[12] Niksefat S, Sadeghiyan B, Mohassel P, et al. ZIDS: A privacy-preserving intrusion detection system using secure two-party computation protocols. Computer Journal, 2014, 57(4): 494-509

[13] Huang H, Li X Y, Sun Y, et al. PPS: Privacy-preserving strategyproof social-efficient spectrum auction mechanisms. IEEE Transactions on Parallel and Distributed Systems, 2015, 26(5): 1393-1404

[14] Li M J, Juan J S T, Tsai J H C. Practical electronic auction scheme with strong anonymity and bidding privacy. Information Sciences, 2011, 181(12): 2576-2586

[15] Li S D, Wang D S, Dai Y Q. Symmetric cryptographic protocols for extended millionaires' problem. Science China Information Sciences, 2009, 52(6): 974-982

[16] Freedman M J, Nissim K, Pinkas B. Efficient private matching and set intersection//Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques. Interlaken, Switzerland, 2004: 1-19

[17] Hazay C, Nissim K. Efficient set operations in the presence of malicious adversaries//Proceedings of the IACR International Workshop on Public Key Cryptography. Paris, France, 2010: 312-331

[18] Fischlin M, Pinkas B, Sadeghi A R, et al. Secure set intersection with untrusted hardware tokens//Proceedings of the Cryptographers' Track at the RSA Conference. San Francisco, USA, 2011: 1-16

[19] De Cristofaro E, Kim J, Tsudik G. Linear-complexity private set intersection protocols secure in malicious model//Proceedings of the International Conference on the Theory and Application of Cryptology and Information Security. Singapore, 2010: 213-231

[20] Jarecki S, Liu X. Fast secure computation of set intersection//Proceedings of the International Conference on Security and Cryptography for Networks. Amalfi, Italy, 2010: 418-435

[21] Freedman M J, Hazay C, Nissim K, et al. Efficient set intersection with simulation-based security. Journal of Cryptology, 2016, 29(1): 115-155

[22] De Cristofaro E, Gasti P, Tsudik G. Fast and private computation of cardinality of set intersection and union//Proceedings of the International Conference on Cryptology and Network Security. Darmstadt, Germany, 2012: 218-231

[23] Egert R, Fischlin M, Gens D, et al. Privately computing set-union and set-intersection cardinality via bloom filters//Proceedings of the Australasian Conference on Information Security and Privacy. Brisbane, Australia, 2015: 413-430

[24] Seo J H, Cheon J H, Katz J. Constant-Round Multi-Party Private Set Union Using Reversed Laurent Series. Berlin, Germany: Springer-Verlag, 2012

[25] Blanton M, Aguiar E. Private and oblivious set and multiset operations. International Journal of Information Security, 2016, 15(4): 493-518

[26] Chun J Y, Hong D, Jeong I R, et al. Privacy-preserving disjunctive normal form operations on distributed sets. Information Sciences, 2013, 231(9): 113-122

[27] Kissner L, Song D. Privacy-preserving set operations//Proceedings of the Annual International Cryptology Conference. Santa Barbara, USA, 2005: 241-257

[28] Li Rong-Hua, Wu Chuan-Kun, Zhang Yu-Qing. Secure computation protocol for testing the inclusion relation of sets. Chinese Journal of Computers, 2009, 32(7): 1337-1345(in Chinese)
(李荣花, 武传坤, 张玉清. 判断集合包含关系的安全计算协议. 计算机学报, 2009, 32(7): 1337-1345)

[29] Guo F, Mu Y, Susilo W. Subset membership encryption and its applications to oblivious transfer. IEEE Transactions on Information Forensics and Security, 2014, 9(7): 1098-1107

[30] Chen Zhen-Hua, Li Shun-Dong, Huang Qiong, et al. Protocols for secure computation of two set-relationships with the unencrypted method. Journal of Software, 2017, 28(3): 473-482(in Chinese)
(陈振华, 李顺东, 黄琼等. 非加密方法安全计算两种集合关系. 软件学报, 2017, 28(3): 473-482)

[31] Zhou S F, Li S D, Dou J W, et al. Efficient secure multiparty subset computation. Security and Communication Networks, 2017, 2017(3): 1-11

[32] Chung H W, Kim M S. Encoding rational numbers for FHE-based applications. IACR Cryptology ePrint Archive, 2016, 2016: 344

[33] Paillier P. Public-key cryptosystems based on composite degree residuosity classes//Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques. Prague, Czech Republic, 1999: 223-238

DOU Jia-Wei, Ph. D., associate professor. Her main research interests include applied mathematics and applied cryptography.

LIU Xu-Hong, M. S. candidate. Her main research interests include applied mathematics and applied cryptography.

WANG Wen-Li, M. S. candidate. Her main research interests include applied mathematics and applied cryptography.

Background

Secure multiparty computation (SMC) plays an important role in information security, and is a pivotal privacy preserving technology both in cyberspaces and in cooperative computation. It is a research focus in the international cryptographic community in recent years. Many cryptographic scholars have explored various secure multiparty computation problems arising in various fields and proposed their solutions. However, there are many difficult problems needed to be further studied in the future.

Set theory is the most important base of modern mathematics, and many mathematical fields are based on set theory. Since many practical problems can be abstract as set problems, it is of great theoretical and practical significance to study secure multiparty set computation. There are many important achievements in the secure multiparty integer set computation. But in the rational number field, the secure multiparty set computation has not yet been reported. Many practical problems can be abstract as rational set problems which have important practical significance and extensive application prospect in information security. In this paper, we mainly study the secure multiparty rational set computation.

We transform the determination of the relation between an element and a set into secure vector product computation by using the triangle area formula, and propose a secure and efficient protocol for determining the relation between an element and a set based on the Paillier encryption system. Then we design protocols for privately computing rational set intersection, rational set union and secure computation of set-inclusion based on first protocol. Finally, we prove that these schemes are secure in the semi-honest model using the simulation paradigm; the experiment demonstrates that our protocols are efficient.

This study is sponsored by the National Natural Science Foundation of China under Grant No. 61272435. The aim of these projects is to address the privacy problems in private data applications. We have been engaged in designing general cryptographic protocols over 10 years. Our members have published more than 80 papers on this topic. More than 30 papers have been indexed by SCI.