

保护隐私的有理数科学计算

刘旭红, 孙晨

(上海体育学院经济管理学院, 上海 200438)

摘要: 安全多方计算作为密码学的基本组成部分, 是各种密码协议的基础, 是国际密码学界的研究热点。近年来, 许多学者研究了各种各样的安全多方计算问题, 包括保密的信息比较、保密的集合问题和保密的计算几何等, 并提出相应的解决方案。而在许多实际应用场景中, 安全多方计算问题需要应用有理数进行描述, 因此研究有理数域上的安全多方计算问题具有重要的理论与实际意义。但现有的安全多方计算问题的研究成果大多数局限于整数范围, 且研究的数据主要是单维度数据。关于有理数域上多维度数据安全多方计算问题的研究较少且无法推广应用。基于有理数的分数表示形式, 设计了新的编码方案(有理数编码方案和有理向量编码方案), 可将有理数域上任意维度的数据进行编码, 为研究有理数域上其他安全多方计算问题提供了新的解决思路。以该编码方案和单向哈希函数为基础, 分别设计了有理数相等、有理向量相等和集合问题的保密判定协议。所设计的协议仅采用基本算术运算和单向哈希函数进行计算, 不需要使用公钥加密算法, 使得协议的计算效率较高; 且协议对研究问题中的数据范围没有限制, 适用范围更广。进一步应用模拟范例严格证明了协议在半诚实模型下的安全性; 并通过理论分析和模拟实验验证了协议的高效性和适用性。通过具体实例说明协议具有广泛适用性, 可以推广应用于其他有理数域的安全多方计算几何问题。

关键词: 安全多方计算; 编码方法; 单向哈希函数; 有理数; 模拟范例

中图分类号: TP309

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2022038

Private-preserving scientific computation of the rational numbers

LIU Xuhong, SUN Chen

School of Economics and Management, Shanghai University of Sport, Shanghai 200438, China

Abstract: As a fundamental part of cryptography, secure multiparty computation (SMC) is a building block of various cryptographic protocols, and it is also a hot topic in the international cryptographic community. In recent years, many SMC problems, such as secret information comparison, secret set problems and secure multiparty computational geometry, have been widely studied. As many practical problems need to be described by rational numbers, it is both theoretically and practically important to study the SMC problems in the rational number field.

收稿日期: 2021-08-11; 修回日期: 2022-03-19

通信作者: 刘旭红, liuxuhong@sus.edu.cn

引用格式: 刘旭红, 孙晨. 保护隐私的有理数科学计算[J]. 网络与信息安全学报, 2022, 8(3): 97-110.

Citation Format: LIU X H, SUN C. Private-preserving scientific computation of the rational numbers[J]. Chinese Journal of Network and Information Security, 2022, 8(3): 97-110.

However, most of the existing researches focus on integers and the studied data are mainly one-dimensional data. There are few researches on secure multiparty computation of multi-dimensional data in the rational number field, but they can't be generalized. Based on the fractional representation of rational numbers, the new encoding schemes about rational numbers and rational number vectors were proposed, which could encode multi-dimensional data in the rational number field and provided new solutions for other SMC problems in the rational number field. Based on the encoding scheme and one-way hash function, some protocols were designed for equality problems and set problems in the rational number field. These protocols used basic arithmetic operation and hash operation to guarantee efficiency than existing related protocols. And these protocols didn't limit the range of research data and they were more widely applicable. It proves that these protocols are secure in the semi-honest model using simulation paradigm, and demonstrates the efficiency and the applicability of these protocols by theoretical analysis and experiment. A practical example was also given to illustrate that approaches are more versatile, and they could also be directly used to solve some secure multiparty computational geometry problems in the rational number field.

Keywords: secure multiparty computation, encoding scheme, one-way hash function, rational numbers, simulation paradigm

0 引言

安全多方计算 (SMC, secure multiparty computation) 是关于一组互不信任的参与者在保护各自隐私数据的前提下进行协同计算的问题, SMC 要确保各参与者输入数据的私密性以及计算结果的正确性。随着信息时代数据安全性越来越受到重视, 安全多方计算成为国际密码学界的研究热点。

图灵奖获得者姚期智教授在文献[1]中首次提出安全多方计算问题, 并引起国际密码学界的重视^[2-3], Goldreich 等^[4,5]对其进行了深入的研究, 从理论上证明了任意的安全多方计算问题都是可解的, 并给出了通用的解决方案。由于通用解决方案的效率很低, 利用通用方案解决具体的计算问题一般是不可行的, 考虑到实际应用场景和计算效率与通信效率的问题, 需要针对具体问题研究具体的解决方案。基于上述学者的研究理论, 许多密码学研究人员不断提出新的具有实际应用前景的安全多方计算问题及其解决方案, 推进了安全多方计算研究的发展。目前所研究的问题有保密信息比较^[6-7]、保密数据挖掘^[8-9]、保密几何计算^[10-11]、保密数据查询^[12-13]等。

保密比较相等问题和保密集合计算问题是基本的安全多方计算问题, 保密比较相等问题是保密信息比较中重要的基础问题并得到了广泛的研

究^[14-18], 其对集合问题的保密计算具有重要的基础作用。集合问题在保密数据挖掘、保密统计分析、保密几何计算等方面有重要的应用, 其中保密判定元素与集合的关系^[19-21]是集合保密计算中的基本问题。文献[14,15]基于第三方参与者设计了整数集上的保密比较相等协议, 计算复杂性较高。文献[16]提出了一个有理数阈值密码系统, 基于所提出的门限方案, 构造了分布式有理数相等测试协议, 协议中需要多个参与者共同参与保密判定两个有理数的相等问题, 且计算复杂性较高。文献[17]利用单项哈希函数解决了两个整数相等问题和整数向量相等问题, 计算效率较高, 但无法抵抗穷举攻击。文献[18]利用 Paillier 公钥加密算法解决了有理数相等问题, 相比文献[16]计算效率有所提高。文献[19]应用不经意多项式计算设计了整数集上元素与集合关系的保密判定协议, 该方案借助于 Paillier 公钥加密体制, 计算复杂性较高。文献[20]利用对称加密体制和异或运算设计了高效的整数集上元素与集合关系的保密判定协议, 但泄露了集合元素的取值范围, 具有一定的局限性。文献[21]应用 Paillier 加密算法首次解决了有理数域上元素与集合关系问题。

上述相关工作作为研究更高效更广泛的安全多方计算奠定了基础, 根据现有相关问题的多个解决方案, 分析每个解决方案的优缺点, 可以进一步设计更加安全高效且适用范围更广的解决方

案。已有的关于保密比较相等问题和保密计算集合问题的研究大多局限于整数范围, 现有的有理数域上的解决方案计算效率低, 且无法应用于实际中有理数域上相关问题的保密计算, 限制了相等问题和集合问题的实际应用场景。

有理数不仅在现实生活中有广泛的应用, 在生物学、气候学、化学等领域中的应用更为广泛, 尤其这些领域的研究工作讲究准确性, 无误差性。例如, 在医疗保健方面, 不仅需要准确诊断病人的健康情况, 而且需要确保病人的健康状况和其他个人信息的安全性。由于大多数医疗数据(如血糖、胰岛素和c肽水平)是有理数^[22], 而传统的密码系统只能加密整数, 这将影响数据的准确性以及计算效率, 从而影响决策和导致错误诊断。基于上述相关问题, 文献[23]提出了允许用户将有理数外包给云服务器进行存储和处理的一种新的高效隐私保护外包计算框架, 为继续研究有理数域上安全多方计算问题提供了研究背景和意义。

在计算几何方面, 有理数的应用也较广泛^[24-25], 举例如下。① Alice 拥有一条直线 $l_1: y = k_1x + b_1$, Bob 拥有一条直线 $l_2: y = k_2x + b_2$, Alice 和 Bob 希望保密判定两条直线是否相交, 则只需保密判定两条直线的斜率 k_1 和 k_2 是否相等, 即保密判定两个有理数是否相等。② 在多边形相似的保密判定问题中, 需要保密判定对应边是否成比例, 通过归一化可以将问题转化为保密判定有理向量相等问题。③ 随着计算机网络的迅速发展, 坐标定位在多个行业中的应用越来越普遍(侦查活动、手机 App 等)。假设在某项侦查活动中, Alice 锁定了多个位置点, 记为点集合 U ; Bob 所在的位置坐标记为点 u , Alice 和 Bob 想要保密判定点 u 是否属于集合 U , 即保密判定二维空间的元素与集合关系。

上述相关问题中, 多个研究对象需要用有理数或有理数集合进行描述, 从而将问题抽象为有理数域上的安全多方计算问题。根据现有的安全多方计算协议, 由于加密体制的限制, 往往需要将计算几何问题转化到正整数集合上进行研究, 导致计算复杂性高、适用性差。研究有理数域上的安全多方计算问题, 通过有理数或有理数集合

直接解决遇到的问题, 或为解决问题提供新的途径, 对安全多方计算的研究具有重要意义。

本文主要研究构造有理数域上相等问题和集合问题的保密计算协议, 是有理数域上安全多方计算研究的基础。

本文的贡献如下。

1) 利用有理数的分数表示形式, 设计了关于有理数的一种全新的编码方法。该编码方案将任意有理数(数组)编码为整数, 为有理数与整数之间的转换提供了一种新的思路。

2) 将新的编码方案与单向哈希函数相结合, 解决了有理数域上的相等问题和集合问题。

3) 协议计算效率高, 且适用范围广(整数集合和有理数集合均适用), 可用于解决更多安全多方计算问题。

4) 通过实例说明对于所设计的协议进行适当修改或推广应用, 对更多的安全多方计算问题或实际应用问题可设计构造安全高效的解决方案。

1 预备知识

本节主要介绍下文研究中所需要的一些基本概念和基础知识^[5]。

1.1 基本定义

(1) 双方计算

双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, 这个过程用函数表示为 $f: (x, y) \rightarrow (f_1(x, y), f_2(x, y))$ 。即对于每一个输入对 (x, y) , 输出对是随机变量 $(f_1(x, y), f_2(x, y))$ 。记这样的函数为 $f = (f_1, f_2)$ 。

(2) 半诚实模型

安全多方保密计算问题的初期研究基本上是以半诚实模型为基础的。半诚实参与者是指在协议执行过程中按照协议要求忠实地履行协议的参与者, 不会欺骗和泄露信息, 但是会收集和保留在协议执行过程中获得的所有信息, 在协议执行后试图根据这些信息推算出额外的信息。如果所有参与者均为半诚实参与者, 这样的计算模型称为半诚实模型。由于半诚实参与者不对协议实施主动攻击, 所以半诚实模型又称为诚实但好奇(honest-but-curious)模型或被动模型。文献[5]给出了一个由半诚实模型下安全协议直接编译获得

恶意模型下安全协议的一般方法,因此半诚实模型下安全计算协议的设计是基本而重要的。本文主要研究半诚实模型下的两方计算问题。

(3) 模拟范例方法

模拟范例方法是由 Goldreich 提出的,在安全多方计算协议安全性证明中广泛使用的证明方法,其证明原理为:如果每个参与者在实际执行协议过程中所获得的信息都可以通过自己的输入和输出进行模拟,而且模拟所得到的消息序列与实际执行过程得到的消息序列计算不可区分,就说明协议是安全的。如果一个多方计算协议能够进行这样的模拟,即说明所有参与者都不可能从协议的执行过程中得到其他参与者任何有价值的信息。这种证明安全性的方法是研究安全多方计算问题时普遍接受的模拟范例方法。模拟范例方法的具体描述如下。

假设参与双方计算的参与者为 P_1 和 P_2 , 则

① 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 的一个双方协议;

② 当输入为 (x, y) 时,在执行协议 π 的过程中 $P_i (i=1, 2)$ 的 view 信息序列记为

$$\text{view}_i^\pi(x, y) = (x, r^i, m_1^i, \dots, m_j^i)$$

其中, r^i 表示参与者 P_i 产生的随机数, m_j^i 表示 P_i 收到的第 j 个消息;

③ 执行协议后,参与者 $P_i (i=1, 2)$ 获得的输出结果记为 $\text{output}_i^\pi(x, y)$ 。

定义 1 半诚实模型下协议的安全性。对于上述函数 f 和协议 π , 如果存在概率多项式时间算法 S_1 和 S_2 , 使

$$\{S_1(x, f_1(x, y))\}_{x, y} \stackrel{c}{=} \{\text{view}_1^\pi(x, y)\}_{x, y} \quad (1)$$

$$\{S_2(y, f_2(x, y))\}_{x, y} \stackrel{c}{=} \{\text{view}_2^\pi(x, y)\}_{x, y} \quad (2)$$

成立,则协议 π 保密地计算 f , 其中 $\stackrel{c}{=}$ 表示计算上不可区分。

要证明一个多方计算方案是安全的,就必须构造满足式(1)和式(2)的模拟器 S_1 和 S_2 。

1.2 字母表与连接运算

字母表是对象的有穷非空集合 M , M 中符号的 n 元组称作 M 上的字或字符串, M 中所有字符串的全体记作 M^* 。例如,所有的自然数都

是字母表 $\{0, 1, 2, \dots, 9\}$ 上的字符串,所有的二进制数都是字母表 $\{0, 1\}$ 上的字符串,字符串的连接定义如下。

$$\begin{cases} \text{Concat}_n^1(u) = u \\ \text{Concat}_n^{m+1}(u_1, \dots, u_m, u_{m+1}) = zu_{m+1} \end{cases}$$

其中, $z = \text{Concat}_n^m(u_1, \dots, u_m)$ 。对于给定的字符串 $u_1, \dots, u_m \in M^*$, $\text{Concat}_n^m(u_1, \dots, u_m)$ 就是把字符串 $u_1, \dots, u_m$ 一个接着一个放在一起所得到的新的字符串。

1.3 单向哈希函数的性质

单向哈希函数(简称为 Hash 函数)是密码体制中一类常用的公开函数,该函数将任意长度的消息映射成某一固定长度消息。它具有如下性质。

1) 给定消息 m , 很容易计算 $h = \text{Hash}(m)$ 。

2) 给定 $h = \text{Hash}(m)$, 根据 h 计算其逆 $m = \text{Hash}^{-1}(h)$ 是困难的。

3) 给定 m , 要找一个消息 m' , 使 $\text{Hash}(m) = \text{Hash}(m')$ 是困难的。

4) 找出两个随机的消息 m, m' , 使 $\text{Hash}(m) = \text{Hash}(m')$ 是困难的。

5) 如果对 m 做微小的改变,即使只改变一个比特,变为 m' , $\text{Hash}(m)$ 的值与 $\text{Hash}(m')$ 的值相比也会发生惊人的变化,至少改变一半位数的值。

进一步,根据单向哈希函数的性质可知, $\text{Hash}(a) = \text{Hash}(b)$ 但 $a \neq b$ 的概率为 $1/2^{|\text{Hash}()|}$ 。例如,在 SHA-1 算法中 $|\text{Hash}()| = 160$, 如果利用 SHA-1 算法做单向哈希函数, $\text{Hash}(a) = \text{Hash}(b)$ 但 $a \neq b$ 的概率为 $1/2^{160}$ 。在一般应用中,如此小的出错概率是可以忽略不计的。

Hash 函数具有上述特性,所以常被用于消息的完整性检测和消息认证。

2 编码方案

本节设计一种有理数的编码方法,以此方法,任一有理数(或有理向量)与一个十进制整数相对应。

2.1 有理数的编码

(1) 有理数的编码方式

① 有理数的分数表示

对任一有理数 w , 首先按如下方式将其表示

为唯一的分数形式 $w := s \frac{w_1}{w_2}$ (其中 $w_1 \geq 0, w_2 \geq 1$ 均为整数, s 为符号位)。当 $w \neq 0$ 时, 要求 w_1, w_2 互素, 当 $w = 0$ 时, 约定 $w_1 = 0, w_2 = 1$; 当 $w \geq 0$ 时, 约定 $s = 1$, 当 $w < 0$ 时约定 $s = 2$ 。

② 有理数的向量表示

按照①中的约定, 有理数 $w := s \frac{w_1}{w_2}$ 与向量 (s, w_1, w_2) 对应。若向量的两个分量 w_1, w_2 的十进制位数最大为 k , 通过高位补零的方法使这两个分量的十进制位数均为 k 位, 并将如此转化后的向量记为 $V(w) = (s, \bar{w}_1, \bar{w}_2)$ 。以此方式, 有理数 w 与向量 $V(w)$ 一一对应。例如, $2/3$ 对应向量 $V(\frac{2}{3}) = (1, 2, 3)$, $23 = 23/1$ 对应向量 $V(23) = (1, 23, 01)$, $-0.048 = (-1) \cdot \frac{6}{125}$ 与向量 $V(-0.048) = (2, 006, 125)$ 对应。

③ 有理数的编码表示

根据②中的方式, 可以将 w 表示成向量 $V(w)$ 。若将 $V(w)$ 的分量 $s, \bar{w}_1, \bar{w}_2$ 看作字母表 $\{0, 1, 2, \dots, 9\}$ 上的字符串, 进一步应用连接运算得到一个长度为 $2k+1$ 的新的字符串 $\text{Concat}(s, \bar{w}_1, \bar{w}_2)$, 并记为 W , 显然 W 亦可看作一个正整数, 称其为 w 的编码整数。例如, $w_1 = 2/3$ 对应编码整数 $W_1 = 123, w_2 = 23$ 对应编码整数 $W_2 = 12301, w_3 = -0.048$ 对应于编码整数 $W_3 = 2006125$ 。

(2) 有理数编码方案的对应关系

设一个映射 $\text{BM}: Q \rightarrow \bar{Q}$, 其中 Q 为有理数集合, $\bar{Q}$ 为编码整数集合 (由集合 Q 中的有理数通过有理数编码方案得到的编码整数集合)。 BM 是一个双射, 即有理数集合 Q 与编码整数集合 $\bar{Q}$ 是一一对应的。

首先, 根据集合 $\bar{Q}$ 的定义知, $\bar{Q}$ 中的元素均为有理数的编码整数, 故 BM 是满射。其次, 对任意两个不同的有理数 $x = s_1 \frac{a_1}{b_1}, y = s_2 \frac{a_2}{b_2}$, 由有理数编码方案得到编码整数 $\text{BM}(x) = \text{Concat}(s_1, \bar{a}_1, \bar{b}_1), \text{BM}(y) = \text{Concat}(s_2, \bar{a}_2, \bar{b}_2)$ 。当 $s_1 \neq s_2$ 时, $\text{BM}(x) \neq \text{BM}(y)$; 当 $a_1 \neq a_2$ 时, 因为高位补零不影响数值, 所以有 $\bar{a}_1 \neq \bar{a}_2$, 即 $\text{BM}(x) \neq \text{BM}(y)$; 同理, 当 $b_1 \neq b_2$ 时, 有 $\text{BM}(x) \neq \text{BM}(y)$ 。即任意两个不相等的有理数 x

和 y , 对应的编码整数 $\text{BM}(x)$ 和 $\text{BM}(y)$ 也不相等。因此, BM 是单射。

综上所述, 映射 BM 是双射, 利用上述有理数编码方式得到的编码整数集合 $\bar{Q}$ 与有理数集合 Q 是一一对应的, 即任一有理数存在唯一与之对应的编码整数。因此, 两个有理数 x, y 相等当且仅当它们的编码整数相等。为叙述方便, 下文中将这种编码方案简记为 BMF 。

2.2 有理向量的编码

(1) 有理向量的编码方式

对于任意一个有理向量 $\bar{v} = (v_1, \dots, v_n)$, 首先对每个 $i \in [1, n] = \{1, \dots, n\}$, 应用 BMF 编码将分量 v_i 编码为正整数 V_i , 则向量 $(v_1, \dots, v_n)$ 与向量 $(V_1, \dots, V_n)$ 一一对应。假设 $V_1, \dots, V_n$ 中最长位分量的位数为 K , 若需要, 可通过高位补零的方式将 $V_1, \dots, V_n$ 均补齐到 K 位长度, 得到 $\bar{V}_1, \dots, \bar{V}_n$ 。进一步通过连接运算 $\text{Concat}(1, \bar{V}_1, \dots, \bar{V}_n)$, 将其记为 $\bar{V}$, 即 $\bar{V} = \text{Concat}(1, \bar{V}_1, \dots, \bar{V}_n)$ 。将 $\bar{V}$ 称为有理向量 $\bar{v}$ 的编码整数。

例如, 有理向量 $\bar{v} = (2/3, 23, -0.048)$, 则编码得到整数 $\bar{V} = \text{Concat}(1, 0000123, 0012301, 2006125) = 1000012300123012006125$ 。

(2) 有理向量编码方案的对应关系

根据上述有理向量的编码过程可知, 有理向量的编码方案是以有理数编码方案 BMF 为基础的, 类似于有理数编码方案的对应关系, 当向量位数 n 确定时, 按照上述有理向量编码方案得到的编码整数存在唯一与之对应的 n 维有理向量。 n 维有理向量 $\bar{v}$ 和正整数 $\bar{V}$ 形成一一对应关系, $\bar{V}$ 称为有理向量 $\bar{v}$ 的编码整数, 记该编码方案为 VBMF 。

下文中有理数组的编码方法与有理向量的编码方法类似, 在此不再重复叙述。

3 相等问题的保密判定方案

本节主要应用单向哈希函数和上节所介绍的编码方案设计有理数相等问题 (REq , rational number equality problem) 以及有理向量相等问题 (RVEq , rational number vector equality problem) 的保密判定协议。本文涉及的随机数

均为非零随机数。

3.1 有理数相等问题的

问题描述 Alice 有一个有理数 x , Bob 有一个有理数 y , 他们想保密判定两个有理数是否相等, 不泄露双方的任何信息。

基本思想 按照前面所述的 BMF 编码方案, 任意有理数 w 与一个编码整数 W 形成一一对应关系。如果有理数 x, y 对应的编码整数分别为 X, Y , 则 $x = y \Leftrightarrow X = Y$, 两个有理数相等的保密判定问题转化为两个整数相等的保密判定问题。

下面将结合编码方案 BMF 与单向哈希函数设计有理数相等问题的保密判定方案, 为叙述方便, 定义二元谓词。

$$P(x, y) = \begin{cases} 1, & x = y \\ 0, & x \neq y \end{cases}$$

具体协议如下。

协议 1 有理数相等保密判定协议

输入 Alice 输入有理数 x , Bob 输入有理数 y

输出 $P(x, y)$

准备 Alice 和 Bob 共同商定一个单向哈希函数 Hash()

1) Alice 和 Bob 分别随机选取足够大的有理数 r 和 t , 计算 $x_1 = x + r$ 和 $y_1 = y + t$, 并将结果发送给对方。

2) Alice 和 Bob 执行以下操作:

a) Alice 计算 $y_2 = y_1 + r$, Bob 计算 $x_2 = x_1 + t$;

b) 利用 BMF 编码方案, Alice 和 Bob 分别将有理数 y_2 和 x_2 编码为十进制正整数 $Y_2 = \text{BM}(y_2)$ 和 $X_2 = \text{BM}(x_2)$;

c) Alice 和 Bob 分别计算 $h_1 = \text{Hash}(Y_2)$ 和 $h_2 = \text{Hash}(X_2)$, 并将结果发给对方。

3) Alice 和 Bob 各自比较 h_1 与 h_2 。如果 $h_1 = h_2$, 则输出 $P(x, y) = 1$; 否则, 输出 $P(x, y) = 0$ 。

协议 1 的正确性

证明 $h_1 = h_2 \Leftrightarrow x = y$ 。首先, 由协议可知 $x_2 = x_1 + t = x + r + t$, $y_2 = y_1 + r = y + t + r$ 。

根据 BMF 编码方案可知, X_2 与 x_2 一一对应, Y_2 与 y_2 一一对应, 则 $X_2 = Y_2 \Leftrightarrow$

$$x_2 = y_2 \Leftrightarrow x = y。$$

其次, 根据单向哈希函数的性质可知 $h_2 = h_1 \Leftrightarrow X_2 = Y_2 \Leftrightarrow x = y$ 。

因此, 协议 1 是正确的。

协议 1 的安全性 关于协议 1 的安全性, 将从以下两部分分别进行详细的分析证明: ① 安全性分析, 包括理论分析和抵抗穷举攻击分析; ② 安全性证明, 对第①部分的理论证明。

安全性分析: 根据协议 1 可知, Alice 和 Bob 私有数据的安全性分析类似, 故在此仅对 Alice 的私有数据进行安全性分析。

首先, 在协议中 Bob 得到 Alice 发送的两个关系式

$$x_1 = x + r, h_1 = \text{Hash}(\text{BM}(y + t + r))$$

由于上面两个关系式含有两个未知数 x 和 r , Bob 只能设法从 $h_1 = \text{Hash}(\text{BM}(y + t + r))$ 中解出 r , 再代入 $x_1 = x + r$ 获得 x ; 或 Bob 试图根据 $h_1 = \text{Hash}(\text{BM}(y + t + x_1 - x))$ 直接求出 x 值。根据 Hash 函数性质, Bob 想由 $h_1 = \text{Hash}(\text{BM}(y + t + r))$ 求得 r , 或由 $h_1 = \text{Hash}(\text{BM}(y + t + x_1 - x))$ 求得 x , 只能试图通过穷举攻击方法实现。而由于 x, r 均为 Alice 的私密有理数, 即使 x 的范围限制在某有限区间内, 由于有理数的稠密性, Bob 不可能通过穷举攻击得到数据 x 的任何信息。因此, Alice 的私有数据 x 是安全的。

安全性证明: 通过前面的理论分析和抵抗穷举攻击的分析可知, 协议 1 理论上是安全的。为了进一步证明协议 1 的安全性, 通过严格的模拟范例方法进行证明, 证明如下。

定理 1 有理数相等保密判定协议 1 是安全的。

证明 在半诚实模型下, 利用模拟范例方法严格证明协议 1 的安全性。构造模拟器 S_1 (或 S_2), 使式(1) (或式(2))成立。仅构造 S_2 (S_1 的构造方法完全类似), 对于输入 $(y, f_2(x, y) = P(x, y))$, 模拟器 S_2 按以下方式运行。

1) S_2 先任意选择 x' , 使 $P(x', y) = P(x, y)$ 。

2) S_2 选择随机数 r' , 并计算 $x'_1 = x' + r'$ 。

3) S_2 计算 $y'_2 = y_1 + r'$ 和 $x'_2 = x'_1 + t$, 并应用 BMF 编码方案将 x'_2, y'_2 分别编码为十进制正整数 $X'_2 = \text{BM}(x'_2), Y'_2 = \text{BM}(y'_2)$ 。

4) S_2 计算 $h'_2 = \text{Hash}(X'_2)$ 和 $h'_1 = \text{Hash}(Y'_2)$ 。

在协议执行中, $\text{view}_2^\pi(x, y) = (y, x_1, h_1, P(x, y))$, 而 S_2 在模拟过程中产生的信息序列为

$$S_2(y, f_2(x, y)) = (y, x'_1, h'_1, P(x', y))$$

由于 r 是 Alice 选的随机数, 通过前面的安全性分析可知, Bob 由 $x_1 = x + r, h_1 = \text{Hash}(\text{BM} \cdot (y + t + r))$ 无法获得 Alice 数据 x 和 r 的任何信息, 因此有

$$x'_1 \stackrel{c}{=} x_1, h'_1 \stackrel{c}{=} h_1$$

又由于 $P(x', y) = P(x, y)$, 因此

$$\{S_2(y, f_2(x, y))\}_{x, y} \stackrel{c}{=} \{\text{view}_2^\pi(x, y)\}_{x, y}$$

类似地, 可以构造模拟器 S_1 使

$$\{S_1(x, f_1(x, y))\}_{x, y} \stackrel{c}{=} \{\text{view}_1^\pi(x, y)\}_{x, y}$$

证毕。

3.2 有理数相等问题-优化的 REq 协议

协议 2 有理数相等保密判定协议

输入 Alice 输入有理数 x , Bob 输入有理数 y

输出 $P(x, y)$

准备 Alice 和 Bob 共同商定一个单向哈希函数 $\text{Hash}()$

1) Bob 执行以下操作:

a) 随机选取足够大的有理数 t , 计算 $y_1 = y + t$ 和 $y_2 = 2y + t$;

b) 利用 BMF 编码方案, 把有理数 y_2 编码为整数 $Y_2 = \text{BM}(y_2)$ 并计算 $h_2 = \text{Hash}(Y_2)$, 将 y_1 和 h_2 发送给 Alice。

2) Alice 执行以下操作:

a) 计算 $x_1 = x + y_1$, 并将有理数 x_1 编码为整数 $X_1 = \text{BM}(x_1)$;

b) 计算 $h_1 = \text{Hash}(X_1)$, 并比较 h_1 与 h_2 , 如果 $h_1 = h_2$, 则输出 $P(x, y) = 1$; 否则, 输出 $P(x, y) = 0$ 。

分析 协议 2 的正确性和安全性完全类似于协议 1, 故详细的分析在此省略。协议 2 只需要半轮通信, 通信效率高于协议 1, 但降低了协议的公平性。因此, 根据不同的场景, 可以选择不同的协议进行计算。关于协议 2 的安全性有下面定理 2。

定理 2 有理数相等保密判定协议 2 是安全的。

协议 2 的证明过程完全类似于协议 1, 故在此省略。

3.3 有理向量相等问题

问题描述 Alice 拥有有理向量 $\bar{x} = (x_1, \dots, x_n)$, Bob 拥有有理向量 $\bar{y} = (y_1, \dots, y_n)$ (这里有理向量指向量的所有分量都是有理数), 双方想要保密判定两个向量 $\bar{x}$ 与 $\bar{y}$ 是否相等。如果不相等, 不应向对方泄露自己私密向量的任何信息。

计算原理 根据有理向量的编码方法 VBMF, 分别将有理向量编码为正整数, 则可将有理向量相等的保密判定问题转化为整数集上数值相等问题。

为叙述方便, 定义二元谓词 $P(\bar{x}, \bar{y})$: 如果 $\bar{x} = \bar{y}$, $P(\bar{x}, \bar{y}) = 1$; 否则, $P(\bar{x}, \bar{y}) = 0$ 。具体协议如下。

协议 3 有理向量相等保密判定协议

输入 Alice 输入有理向量 $\bar{x} = (x_1, \dots, x_n)$, Bob 输入有理向量 $\bar{y} = (y_1, \dots, y_n)$

输出 $P(\bar{x}, \bar{y})$

1) Alice 和 Bob 分别选取随机有理向量 $R = (r_1, \dots, r_n)$ 和 $T = (t_1, \dots, t_n)$ (其中 r_i, t_i ($i \in [1, n]$) 均为随机有理数), 计算 $\bar{x}_1 = \bar{x} + R$ 和 $\bar{y}_1 = \bar{y} + T$, 并将结果发送给对方。

2) Alice 和 Bob 执行如下步骤:

a) Alice 和 Bob 分别计算 $\bar{y}_2 = \bar{y}_1 + R$ 和 $\bar{x}_2 = \bar{x}_1 + T$;

b) Alice 和 Bob 根据向量编码方法 VBMF 分别将向量 $\bar{y}_2$ 和 $\bar{x}_2$ 编码为整数 $\bar{Y}_2$ 和 $\bar{X}_2$;

c) Alice 和 Bob 分别计算 $\text{Hash}(\bar{Y}_2)$ 和 $\text{Hash}(\bar{X}_2)$, 并将结果发给对方。

3) Alice 和 Bob 各自比较 $\text{Hash}(\bar{X}_2)$ 与 $\text{Hash}(\bar{Y}_2)$ 。如果 $\text{Hash}(\bar{X}_2) = \text{Hash}(\bar{Y}_2)$, 则输出 $P(\bar{x}, \bar{y}) = 1$; 否则, 输出 $P(\bar{x}, \bar{y}) = 0$ 。

协议 3 的正确性和安全性 协议 3 的设计思想与协议 1 类似, 协议 3 的正确性和安全性证明过程与协议 1 也相似, 这里仅叙述下面定理。

定理 3 有理向量相等的保密判定协议 3 是安全的。

协议 3 的安全性证明完全类似于协议 1，故在此省略。

4 集合问题的保密判定方案

本节主要以协议 1 和协议 3 的构造思想为基础，研究设计有理数域上元素与集合关系问题（ReES, the relationship problem between element and set）、坐标系中的点集合问题（RePS, point set problem of coordinate system）的保密计算协议。

4.1 元素与集合关系问题

问题描述 Alice 拥有私密的有理数集合 $X = \{x_1, \dots, x_n\}$ ，Bob 拥有有理数 y ，他们希望合作保密判定 y 是否属于集合 X ，不泄露双方的任何信息（包括保密集合 X 的势）。

计算原理

1) 根据协议 1 的构造思想，Alice（或 Bob）利用 BMF 编码对自己的有理数集合元素 $x_1, \dots, x_n$ （或 y ）进行编码，Alice 得到编码整数集合 $U = \{u_1, \dots, u_n\}$ （其中 $u_i = \text{BM}(x_i), i \in [1, n]$ ），Bob 得到 $v = \text{BM}(y)$ 。那么，判定 y 是否属于集合 X 的问题可转化为判定 v 是否属于集合 U 的问题。

2) 为了隐藏 Alice 集合的势，Alice 随机选择整数 $l > n$ ，构造 $l - n$ 个形如 $1b_1 \dots b_k 0 \dots 0$ 或 $2b_1 \dots b_k 0 \dots 0$ 的“伪编码数”（这里 k 为任意选取的正整数， $b_j (j \in [1, k])$ 在 $\{0, \dots, 9\}$ 中随机选取，最后 k 位设为零），记这些整数为 $u_{n+1}, \dots, u_l$ 。将这些整数添加到集合 U 中，得到一个势为 l 的新集合，将其记为 $\hat{U} = \{u_1, \dots, u_l\}$ 。

3) 根据“伪编码数” $u_{n+1}, \dots, u_l$ 的构造方式，由于任何有理数的编码整数都不可能是 $u_{n+1}, \dots, u_l$ 的形式，可知 $y \in X$ 当且仅当 $v \in \hat{U}$ 。

下面根据上述计算原理设计协议，定义二元谓词 $P(y, X)$ ：当 $y \in X$ 时， $P(y, X) = 1$ ；否则， $P(y, X) = 0$ 。

协议 4 元素与集合关系的保密判定协议

输入 Alice 输入有理数集合 $X = \{x_1, \dots, x_n\}$ ，Bob 输入有理数 y

输出 $P(y, X)$

1) Bob 执行以下操作：

a) Bob 随机选取足够大的有理数 t ，计算 $y_1 = y + t$ 和 $y_2 = 2y + t$ ，并将 y_1 发送给 Alice；

b) Bob 按照 BMF 编码方案，将 y_2 编码为整数 $v = \text{BM}(y_2)$ ，并计算 $z = \text{Hash}(v)$ 。

2) Alice 执行以下操作：

a) Alice 计算 $W = \{x_1 + y_1, \dots, x_n + y_1\} := \{w_1, \dots, w_n\}$ ；

b) Alice 按照 BMF 编码方案，将 w_i 编码为整数 $u_i = \text{BM}(w_i) (i \in [1, n])$ ，并随机添加“伪编码数” $u_{n+1}, \dots, u_l$ 得到集合 $\hat{U} = \{u_1, \dots, u_l\} (l > n)$ ；

c) Alice 计算 $Z = \{\text{Hash}(u_1), \dots, \text{Hash}(u_l)\} := \{z_1, \dots, z_l\}$ ，并将 Z 发送给 Bob。

3) Bob 比较 z 与 Z ，若存在 $i \in [1, l]$ ，使 $z = z_i$ ，则输出 $P(y, X) = 1$ ；否则，输出 $P(y, X) = 0$ 。

协议 4 的正确性 根据协议 4 的过程和 BMF 编码方案可知

$$\begin{aligned} y \in X &\Leftrightarrow y_2 \in W \Leftrightarrow v \in \hat{U} \\ &\Leftrightarrow \exists i \in [1, l], z = z_i \\ &\Leftrightarrow \exists i \in [1, l], v = u_i \end{aligned}$$

因此，协议 4 是正确的。

协议 4 的安全性 关于协议 4 的安全性，从以下两部分进行详细的分析证明：① 安全性分析，包括理论分析和抵抗穷举攻击分析；② 安全性证明，对第①部分的理论证明。

安全性分析：首先分析 Alice 集合 X 的安全性。在协议执行中，Bob 得到 Alice 发送的集合 $Z = \{z_1, \dots, z_l\}$ 。由于 l 是 Alice 选取的随机数，保证了 Alice 的集合势 n 的安全性，并且 Bob 无法区分出集合 Z 中的真假元素。根据单向哈希函数的性质，Bob 无法从 $z_i = \text{Hash}(u_i) (i \in [1, l])$ 中得到 u_i 的任何信息，因此集合元素 x_i 是安全的，集合 X 是安全的。

接着分析 Bob 数据 y 的安全性。Alice 仅得到 Bob 发送的 $y_1 = y + t$ ，由于 Alice 不知道随机数 t ，所以 Bob 的数据 y 是安全的。

下面说明协议 4 可以抵抗穷举攻击。协议 4 抵抗穷举攻击的过程和协议 1 类似，根据协议 1 的分析可知，由于有理数的稠密性，任意有理数

的编码整数均匀分布在整个数轴上, 所以编码整数的范围是无限的, 故协议 4 可以抵抗穷举攻击。

安全性证明: 通过前面的理论分析和抵抗穷举攻击的分析可知, 协议 4 理论上是安全的。为了进一步证明协议 4 的安全性, 通过严格的模拟范例方法进行证明, 证明如下。

定理 4 元素与集合关系的保密判定协议 4 是安全的。

证明 下面应用模拟范例严格证明定理 4, 为此需要构造模拟器 S_1 (或 S_2), 分别使式(1)(或式(2))成立。

先构造 S_1 。对输入 $(X, f_1(y, X) = P(y, X))$, S_1 按以下方式运行。

- 1) S_1 任意选择 y' , 使 $P(y', X) = P(y, X)$ 。
- 2) S_1 随机选取足够大的有理数 t' 计算 $y'_1 = y' + t'$ 和 $y'_2 = 2y' + t'$ 。将 y'_2 编码为整数 $v' = \text{BM}(y'_2)$, 并计算 $h' = \text{Hash}(v')$ 。
- 3) S_1 计算 $W' = \{x_1 + y'_1, \dots, x_n + y'_1\} := \{w'_1, \dots, w'_n\}$ 。并将 w'_i 编码为整数 $u'_i = \text{BM}(w'_i) (i \in [1, n])$ 。根据计算原理, S_1 随机添加“伪编码数”得到集合 $\hat{U}' = \{u'_1, \dots, u'_l\} (l > n)$ 。

在协议执行中, $\text{view}_1^*(y, X) = (X, y_1, P(y, X))$, 而 S_1 在模拟过程中产生的信息序列为 $S_1(X, f_1(y, X)) = (X, y'_1, P(y', X))$ 。

由于 t 和 t' 为随机数, 有 $t' \stackrel{c}{=} t$, 因此 $y'_1 \stackrel{c}{=} y_1$; 进一步, 由于 $P(y', X) = P(y, X)$, 因此 $S_1(X, f_1(y, X)) \stackrel{c}{=} \text{view}_1^*(y, X)$ 。

对于输入 $(y, f_2(y, X) = P(y, X))$, 模拟器 S_2 按以下方式运行。

- 1) S_2 构造含 n' 个元素的集合 $X' = \{x'_1, \dots, x'_{n'}\}$, 使 $P(y, X') = P(y, X)$ 。
- 2) S_2 计算 $W' = \{x'_1 + y_1, \dots, x'_{n'} + y_1\} := \{w'_1, \dots, w'_{n'}\}$, 并且按照 BMF 编码方案, 将 w'_i 编码为整数 $u'_i = \text{BM}(w'_i) (i \in [1, n'])$ 。由计算原理, S_2 随机添加“伪编码数”得到集合 $\hat{U}' = \{u'_1, \dots, u'_l\} (l > n')$ 。
- 3) S_2 计算 $Z' = \{\text{Hash}(u'_1), \dots, \text{Hash}(u'_{n'})\} := \{z'_1, \dots, z'_{n'}\}$ 。

在协议执行中, $\text{view}_2^*(y, X) = (y, Z, P(y, X))$, 而 S_2 在模拟的过程中产生的信息序列为

$S_2(y, f_2(y, X)) = (y, Z', P(y, X'))$ 。根据 Hash 函数的安全性可知, $Z' \stackrel{c}{=} Z$; 又 $P(y, X') = P(y, X)$, 因此 $S_2(y, f_2(y, X)) \stackrel{c}{=} \text{view}_2^*(y, X)$ 。

证毕。

4.2 坐标系中的点集问题

目前, 在 SMC 中对集合问题的研究仅限于元素为单一数据的集合。当集合元素是一个数组时(如集合元素是坐标点 (x, y)), 现有的解决方案并不适用。因此, 本节研究坐标点集问题(进一步可以扩展到更高维的集合问题)。

问题描述 Alice 有一个点集合 $A = \{a_1, \dots, a_n\}$, Bob 有一点 a_0 , 其中 $a_i = (x_i, y_i)$ 且 $x_i, y_i (i \in [0, n])$ 均为有理数。他们想保密判定是否 $a_0 \in A$, 不泄露双方的任何信息 A, a_0 。

计算原理 根据 VBMF 编码方案, Alice 将点 $a_i = (x_i, y_i)$ 编码为整数 $\bar{a}_i (i \in [1, n])$, 得到一个新的集合 $\bar{A} = \{\bar{a}_1, \dots, \bar{a}_n\}$; Bob 将点 $a_0 = (x_0, y_0)$ 编码为整数 $\bar{a}_0$ 。那么

$$a_0 \in A \Leftrightarrow \bar{a}_0 \in \bar{A}$$

为方便叙述, 定义 $P(a_0, A)$: 若 $a_0 \in A$, $P(a_0, A) = 1$; 否则, $P(a_0, A) = 0$ 。

协议 5 点与点集合关系的保密判定协议

输入 Alice 输入一个点集合 $A = \{a_1, \dots, a_n\}$, Bob 输入点 $a_0 = (x_0, y_0)$

输出 $P(a_0, A)$

1) Bob 随机选取点 $T = (t_1, t_2)$ 计算 $\hat{a}_0 = a_0 + T$, 其中 t_1, t_2 均为有理数, 并将 $\hat{a}_0$ 发送给 Alice。

2) Alice 执行以下操作:

a) Alice 随机选取点集合 $R = \{R_1, \dots, R_n\}$ 计算 $W = \{a_1 + R_1, \dots, a_n + R_n\} := \{w_1, \dots, w_n\}$ 以及 $U = \{\hat{a}_0 + R_1, \dots, \hat{a}_0 + R_n\} := \{u_1, \dots, u_n\}$, 其中 $R_i = (r_{i1}, r_{i2})$, r_{i1}, r_{i2} 均为有理数。

b) Alice 对点集合 W 和 U 执行相同的随机置换得到 $\hat{W} = (\hat{w}_1, \dots, \hat{w}_n)$ 和 $\hat{U} = (\hat{u}_1, \dots, \hat{u}_n)$ 。根据 VBMF 编码方案, Alice 将点 $\hat{u}_i (i \in [1, n])$ 编码为整数 $\bar{u}_i$, 并得到向量 $\bar{U} = (\bar{u}_1, \dots, \bar{u}_n)$ 。

c) Alice 计算 $Z = (\text{Hash}(\bar{u}_1), \dots, \text{Hash}(\bar{u}_n)) := (z_1, \dots, z_n)$, 并将 $\hat{W}, Z$ 发送给 Bob。

3) Bob 执行以下操作:

a) Bob 计算 $V = (\hat{w}_1 + T, \dots, \hat{w}_n + T) := (v_1, \dots, v_n)$, 根据 VBMF 编码方案, Bob 将点 $v_i (i \in [1, n])$ 编码为整数 $\bar{v}_i$, 得到向量 $\bar{V} = (\bar{v}_1, \dots, \bar{v}_n)$ 。

b) Bob 计算 $Z' = (\text{Hash}(\bar{v}_1), \dots, \text{Hash}(\bar{v}_n)) := (z'_1, \dots, z'_n)$ 。

c) Bob 比较 Z 和 Z' 。若存在 $i \in [1, n]$, 使 $z_i = z'_i$, 则输出 $P(a_0, A) = 1$; 否则, 输出 $P(a_0, A) = 0$ 。

协议 5 的正确性 根据协议 5 的过程以及 VBMF 编码方案可知

$$\begin{aligned} a_0 \in A &\Leftrightarrow \exists i \in [1, n], a_0 = a_i \\ &\Leftrightarrow \exists i \in [1, n], \hat{a}_0 + R_i = w_i + T \\ &\Leftrightarrow \exists i \in [1, n], \hat{u}_i = \hat{w}_i + T \\ &\Leftrightarrow \exists i \in [1, n], \bar{u}_i = \bar{v}_i \\ &\Leftrightarrow \exists i \in [1, n], z_i = z'_i \end{aligned}$$

因此, 协议 5 是安全的。

协议 5 的安全性 关于协议 5 的安全性, 本文从以下两部分分别进行详细的分析证明: ① 安全性分析, 包括理论分析和抵抗穷举攻击分析; ② 安全性证明, 对第①部分的理论证明。

安全性分析: 下面分析协议 5 的安全性, 并表明协议 5 可以抵抗穷举攻击。首先, 分析 Alice 集合 A 的安全性。在协议执行中, Bob 得到 Alice 发送的 $Z = (z_1, \dots, z_l)$ 和 $\hat{W} = (\hat{w}_1, \dots, \hat{w}_n)$ 。根据单向哈希函数的性质可知, Bob 根据 $z_i = \text{Hash}(u_i) (i \in [1, n])$ 得不到关于 u_i 或者 R_i 的任何信息。进一步, 对于 $\hat{w}_i$ 中的有理数组 a_i 和 R_i 均为 Alice 的私有数据, 因此 Bob 根据 $\hat{w}_i$ 得不到关于 a_i 的任何信息, 即点 a_i 是安全的。因此集合 A 是安全的。

其次, 分析 Bob 的点 a_0 的安全性。Alice 仅得到 Bob 发送的 $\hat{a}_0 = a_0 + T$, 因为随机数组 T 是 Bob 的私密数据, Alice 根据 $\hat{a}_0$ 得不到 a_0 的任何信息。因此, Bob 的点 a_0 是安全的。

最后, 说明协议 5 可以抵抗穷举攻击。协议 5 抵抗穷举攻击的过程类似于协议 1 和协议 4, 故在此省略。

安全性证明: 通过前面的理论分析和抵抗穷举攻击的分析可知, 协议 5 理论上是安全的。为了进一步证明协议 5 的安全性, 下面通过严格的模拟范例方法进行证明。

定理 5 点与点集合关系的保密判定协议 5 是安全的。

协议 5 的安全性证明过程完全类似于协议 4, 故在此省略。

5 性能分析

下面对本文所设计协议的性能和适用范围进行分析, 并与现有相关协议分析进行比较。文献[17,18]分别研究了整数范围和有理数域上的数据相等问题, 主要以 Hash 函数以及 Paillier 同态加密算法为基础设计协议, 文献[17]的协议计算效率较高但仅适用于整数范围; 文献[18]解决了有理数相等问题但计算效率相对较低。文献[21]以 Paillier 同态加密算法为基础设计了有理数域上元素与集合关系的保密判定协议, 但计算效率低且应用范围小。

下面主要通过比较各协议所需要进行的 Hash 运算次数以及模指数运算的次数来衡量计算复杂性, 而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。为了便于分析比较, 假设 Alice 集合的势为 n , Bob 集合的势为 m 。

5.1 理论分析与比较

5.1.1 计算复杂性分析

本文协议 1~协议 5 主要运用了基本算术运算(普通实数的加减乘除运算)和 Hash 运算, 由于基本算术运算与 Hash 运算相比计算复杂性可以忽略, 在此只考虑 Hash 运算次数。

协议 1 (协议 2) 中 Alice 和 Bob 分别执行了 1 次 Hash 运算, 因此协议 1 (协议 2) 共执行了 2 次 Hash 运算。协议 3 应用协议 1 的基本思想, 将向量相等问题转化为数值相等问题, 协议 3 也需要执行 2 次 Hash 运算。协议 4 中 Alice 执行了 l 次 Hash 运算, Bob 执行了 1 次 Hash 运算, 即协议 4 共执行了 $l+1$ 次 Hash 运算(其中 l 是为隐藏集合 X 的势选取的大于 $|X| = n$ 的一个数值, 不需要取值太大, 能够保证协议 4 具有线性复杂性)。基于协议 4 的设计思想, 协议 5 首次解决

了坐标系中点与点集合关系的保密计算问题。在协议 5 中, Alice 和 Bob 均需要执行 n 次 Hash 运算, 即协议 5 共执行了 $2n$ 次 Hash 运算。

文献[17]基于 Hash 函数设计了正整数范围内相等问题的保密判定协议 1 和向量相等协议 2, 协议 1 和协议 2 均需执行 2 次 Hash 运算。文献[18]基于 Paillier 同态加密算法设计了有理数相等保密判定协议, 协议共需要 6 次模指数运算。文献[21]基于 Paillier 同态加密算法设计了有理数域上元素与集合关系的保密判定协议, 其中 Alice 需要执行 $l+2$ 次模指数运算, Bob 需要执行 $2l+6$ 次模指数运算, 协议共需执行 $3l+8$ 次模指数运算。

5.1.2 通信复杂性分析

本文协议 1 和协议 3 各需要 2 轮通信; 协议 2 需要半轮通信; 协议 4~协议 5 各需要 1 轮通信。

文献[17]中的协议 1 和协议 2 均需 2 轮通信; 文献[18]需要 2 轮通信; 文献[21]需要 2 轮通信。

根据上述理论分析, 现将本文协议的效率及适用范围分析的结果在表 1 中给出。在表 1 中, 其中计算功能一栏为所列文献研究的具体内容; 计算复杂性一栏, H (或 M) 表示 Hash (或模指数) 运算, 通信复杂性为通信轮数。

表 1 协议的效率分析与适用范围比较
Table 1 Efficiency analysis and scope analysis of the protocols

协议	计算功能	计算复杂性	通信复杂性	适用范围	抵抗穷举攻击
协议 1	REq	2(H)	2	有理数	抵抗
协议 3	RVEq	2(H)	2	有理数	抵抗
协议 4	ReES	$l+1$ (H)	1	有理数	抵抗
文献[17]协议 1	IEq	2(H)	2	正整数	不抵抗
文献[17]协议 2	IVEq	2(H)	2	正整数	不抵抗
文献[18]	REq	6 (M)	2	有理数	/
文献[21]	ReES	$3l+8$ (M)	2	有理数	/

通过表 1 分析可知, 本文协议仅需较少的 Hash 运算即可实现有理数 (或有理向量) 相等问题以及有理集合问题的保密判定, 扩展了相关安全多方计算问题的适用范围且计算效率优势明显。

5.2 实验测试

表 1 从理论上对本文协议和以往相关协议的效率进行了全面分析, 并进行了比较。下面进一步利用 Java 编程语言在 MyEclipse 上对本文协议 4 和文献[21]的协议进行编程实现。

(1) 实验平台

计算机的配置如下: 操作系统为 Windows10 企业版, Intel Core i5-6600 CPU @3.30 GHz, 安装内存 8.00 GB, 64 位操作系统。在此, 约定本文协议所做模拟实验均在此环境下进行。

(2) 实验结果

本文协议 4 应用单向哈希函数, 文献[21]应用 Paillier 同态加密体制, 下面分别对本文协议 4 和文献[21]进行仿真实验。实验设定 Paillier 加密方案中大素数 p 和 q 的位数为 512 bit, 因此模数的位数为 1 024 bit。保密数据的范围为 $[-100, 100]$ 。通过对本文协议 4 和文献[21]的协议进行实际计算, 对每种协议进行多次实验, 实验结果随机抽取 50 次数据求取平均值, 结果如表 2 所示。

表 2 实验结果分析
Table 2 Analysis of simulation result

协议	平均耗时/ms
本文协议 4	2.269
文献[21]	189.136

由表 2 可知, 与文献[21]相比, 本文协议 4 的计算效率较高, 且优势明显。理论分析 (表 1) 和实验测试 (表 2) 两方面均表明本文协议优势明显。

为了进一步测试效率和集合基数 (添加“伪编码”数后的元素个数) 之间的关系, 对本文协议 4 和文献[21]分别进行了多维实验测试, 实验结果如图 1 所示。

在不同基数下的仿真结果 (图 1) 表明本文协议 4 的计算效率随集合基数的变化幅度较小, 文献[21]的计算效率随集合基数的变化幅度较大, 本文协议 4 具有明显优势, 且本文协议 4 更适用于集合基数大的应用场景, 适用范围更广。

6 协议的推广及实际应用

目前关于有理数域上的安全多方计算的研究

较少,但在计算几何等方面有理数的应用十分广泛,下面根据前面协议的设计思想,对更广泛的有理数域上安全多方计算实际问题设计新的高效的保密协议。

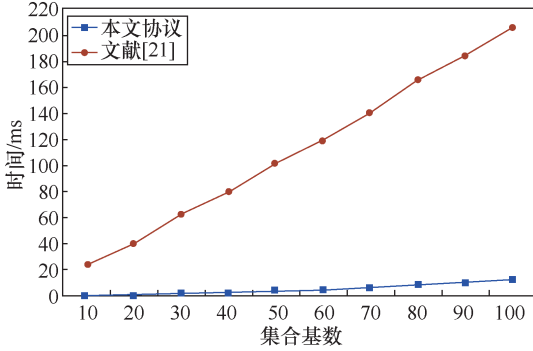


图 1 计算效率与集合基数的关系
Figure 1 The relationship between computational efficiency and set cardinality

直线位置关系问题描述

Alice 有一条直线 $l_1: y = k_1x + b_1$, Bob 有一条直线 $l_2: y = k_2x + b_2$, 双方要保密判定两条直线的位置关系(重合、平行、相交、垂直), 不泄露各自的私密信息。

直线位置关系计算原理

首先, 根据两条直线位置关系的判定定理可知, ① 当两条直线的斜率 $k_1 = k_2$ 且截距 $b_1 = b_2$ 时, 两直线重合, 记为 $l_1 = l_2$; ② 当两条直线的斜率 $k_1 = k_2$ 但截距 $b_1 \neq b_2$ 时, 两直线平行, 记为 $l_1 \parallel l_2$; ③ 当两条直线的斜率 $k_1 \neq k_2$ 时, 两直线相交, 记为 $l_1 \times l_2$; ④ 当两条直线的斜率满足 $k_1 \times k_2 = -1$ 时, 两直线垂直, 记为 $l_1 \perp l_2$ 。特别地, 当直线与 X 轴平行时, 记直线斜率为 $-0/1$; 当直线与 Y 轴平行时, 记直线斜率为 $1/0$ 。那么, 两直线位置关系描述如下。

$$\begin{aligned} (k_1, b_1) = (k_2, b_2) &\Leftrightarrow l_1 = l_2 \\ (k_1, b_1) \neq (k_2, b_2), k_1 = k_2 &\Leftrightarrow l_1 \parallel l_2 \\ k_1 \neq k_2 &\Leftrightarrow l_1 \times l_2 \\ -\frac{1}{k_1} = k_2 &\Leftrightarrow l_1 \perp l_2 \end{aligned}$$

其次, Alice 和 Bob 利用有理向量的编码方法 VBMF, 将向量 $(k_1, b_1), (k_2, b_2)$ 分别编码为整数 u_1, v_1 ; Alice 利用有理数编码方法 BMF, 将斜率 $k_1, -1/k_1$ 分别编码为整数 u_2, u_3 ; Bob 利用有理数编

码方法 BMF, 将斜率 k_2 编码为整数 v_2 , 则

$$\begin{aligned} u_1 = v_1 &\Leftrightarrow l_1 = l_2 \\ u_1 \neq v_1, u_2 = v_2 &\Leftrightarrow l_1 \parallel l_2 \\ u_2 \neq v_2 &\Leftrightarrow l_1 \times l_2 \\ u_3 = v_2 &\Leftrightarrow l_1 \perp l_2 \end{aligned}$$

根据上述计算原理设计协议, 为叙述方便, 定义二元谓词 $P(l_1, l_2)$ 如下。

$$P(l_1, l_2) = \begin{cases} 0, & l_1 = l_2 \\ 1, & l_1 \parallel l_2 \\ 2, & l_1 \times l_2 \\ 3, & l_1 \perp l_2 \end{cases}$$

具体协议如下。

协议 6 直线位置关系的保密判定协议

输入 Alice 输入直线 $l_1: y = k_1x + b_1$, Bob 输入直线 $l_2: y = k_2x + b_2$

输出 $P(l_1, l_2)$

准备 Alice、Bob 商定单向哈希函数 Hash()

1) 按照计算原理所述, Alice 利用 VBMF 编码方案和 BMF 编码方案, 将向量 (k_1, b_1) 以及 $k_1, -\frac{1}{k_1}$ 分别编码为整数 u_1, u_2, u_3 ; 类似地, Bob 利用 VBMF 编码方案和 BMF 编码方案, 将向量 (k_2, b_2) 以及 k_2 分别编码为整数 v_1, v_2 。

2) Bob 选取充分大的随机数 t_1, t_2 (均为有理数), 分别计算 $\hat{v}_1 = v_1 + t_1$ 和 $\hat{v}_2 = v_2 + t_2$, 并将结果发送给 Alice。

3) Alice 选取充分大的随机数 r_1, r_2, r_3 (均为有理数), 计算 $W = (u_1 + r_1, u_2 + r_2, u_3 + r_3) := (w_1, w_2, w_3)$, $V = (\text{BM}(\hat{v}_1 + r_1), \text{BM}(\hat{v}_2 + r_2), \text{BM}(\hat{v}_2 + r_3)) := (\bar{v}_1, \bar{v}_2, \bar{v}_3)$ 以及 $Z = (\text{Hash}(\bar{v}_1), \text{Hash}(\bar{v}_2), \text{Hash}(\bar{v}_3)) := (z_1, z_2, z_3)$, 并将 W, Z 发送给 Bob。

4) Bob 计算 $\bar{W} = (\text{BM}(w_1 + t_1), \text{BM}(w_2 + t_2), \text{BM}(w_3 + t_2)) := (\bar{w}_1, \bar{w}_2, \bar{w}_3)$ 以及 $Z' = (\text{Hash}(\bar{w}_1), \text{Hash}(\bar{w}_2), \text{Hash}(\bar{w}_3)) := (z'_1, z'_2, z'_3)$ 。

Bob 比较 Z 与 Z' 的各个分量, 如果 $z_1 = z'_1$, 则输出 $P(l_1, l_2) = 0$; 如果 $z_1 \neq z'_1, z_2 = z'_2$, 则输出 $P(l_1, l_2) = 1$; 如果 $z_2 \neq z'_2$, 则输出 $P(l_1, l_2) = 2$; 如果 $z_3 = z'_3$, 则输出 $P(l_1, l_2) = 3$ 。

协议 6 的正确性和安全性 协议 6 的设计思

想本质上是协议 1 和协议 3 设计思想的推广。根据协议 1 和协议 3 的正确性和安全性可知, 协议 6 是正确且安全的。

7 结束语

本文提出了有理数域上一些基本的安全多方计算问题, 设计了有理数(数组)的编码方法, 并以此为基础解决了有理数域上相等问题和集合问题的保密计算。通过对本文协议进行适当的应用, 可以进一步解决有理数域上其他的安全多方计算问题, 对后续有理数域上安全多方计算问题的研究和应用具有重要意义。目前, 本文的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有理数域上的安全多方计算问题。

参考文献:

- [1] YAO A C. Protocols for secure computations[C]//Proceedings of the 23th Annual Symposium on Foundations of Computer Science. 1982: 160-164.
- [2] GOLDWASSER S. Multi-party computations: past and present[C]//The 16th Annual ACM Symposium on Principles of Distributed Computing. 1997: 1-6.
- [3] CRAMER R, DAMGARD IB, NIELSEN J B. Secure multiparty computation[M]. London: Cambridge University Press, 2015.
- [4] GOLDREICH O, MICALI S, WIGDERSON A. How to play any mental game[C]//Proceedings of the 19th Annual ACM Conference on Theory of Computing. 1987: 218-229.
- [5] GOLDREICH O. The fundamental of cryptography: basic applications[M]. London: Cambridge University Press, 2004.
- [6] 张兴兰, 郑炜. 基于博弈论的安全多方计算的研究[J]. 网络与信息安全学报, 2018, 4(1): 52-56.
- [7] ZHANG X L, ZHENG W. Research on security multi-party computing based on game theory[J]. Chinese Journal of Network and Information Security, 2018, 4(1): 52-56.
- [8] HIROFUMI M, NORITAKA S, HIROMI M. A proposal of profit sharing method for secure multiparty computation[J]. International Journal of Innovative Computing Information and Control, 2018, 14(2): 727-735.
- [9] 邢欢, 张琳. 基于 paillier 的隐私保护下关联规则挖掘方法[J]. 网络与信息安全学报, 2016, 2(1): 53-59.
- [10] XING H, ZHANG L. Privacy-preserving mining of association rules based on paillier encryption algorithm[J]. Chinese Journal of Network and Information Security, 2016, 2(1): 53-59.
- [11] SHAO Y, HONG W J, LI Z J. A new method to compute ratio of secure summations and its application in privacy preserving distributed data mining[J]. IEEE Access, 2019, 7: 20756-20766.
- [12] 窦家维, 王颖囡, 葛雪. 区间关系保密计算若干问题研究[J]. 电子学报, 2021, 49(1): 50-57.
- [13] DOU J W, WANG Y N, GE X. Some research on secure interval relation computation[J]. Acta Electronica Sinica, 2021, 49(1): 50-57.
- [14] 张明武, 冷文韬, 沈华. 隐私保护的点与任意多边形位置关系判定[J]. 密码学报, 2019, 6(4): 443-454.
- [15] ZHANG M W, LENG W T, SHEN H. Privacy preservation protocol to determine position relation between points and arbitrary polygons[J]. Journal of Cryptologic Research, 2019, 6(4): 443-454.
- [16] 何贤芒. 基于差分隐私保护技术的多方求和查询方法[J]. 网络与信息安全学报, 2020, 6(3): 14-18.
- [17] HE X M. Multi-party summation query method based on differential privacy[J]. Chinese Journal of Network and Information Security, 2020, 6(3): 14-18.
- [18] QAOSAR M, ZAMAN A, SIDDIQUE M A, et al. Privacy-preserving secure computation of skyline query in distributed multi-party databases[J]. Information, 2019, 10(3): 119.
- [19] DU W L. A study of several specific secure two-party computation problems[D]. West Lafayette: Purdue University, 2001.
- [20] CHEN X B, XU G, NIU X X, et al. An efficient protocol for the private comparison of equal information based on the triplet entangled state and single-particle measurement[J]. Optics Communications, 2010, 283(7): 1561-1565.
- [21] GONG L M, YANG B, TA O X, et al. Secure rational numbers equivalence test based on threshold cryptosystem with rational numbers[J]. Information Sciences, 2018, 466: 44-54.
- [22] 李顺东, 杨晓莉, 左祥建, 等. 保护私有信息的图形相似判定[J]. 电子学报, 2017, 45(9): 2184-2189.
- [23] LI S D, YANG X L, ZUO X J, et al. Privacy protecting similitude determination for graphics similarity[J]. Chinese Journal of Electronics, 2017, 45(9): 2184-2189.
- [24] 李顺东, 杜润萌, 杨颜璟, 等. 有理数相等的保密判定[J]. 电子学报, 2020, 48(10): 1933-1937.
- [25] LI S D, DU R M, YANG Y J, et al. Privately determining equality of rational numbers[J]. Chinese Journal of Electronics, 2020, 48(10): 1933-1937.
- [26] FREEDMAN M J, NISSIM K, PINKAS B. Efficient private matching and set intersection[C]//Advances in Cryptology- Eurocrypt. 2004: 1-19.
- [27] LI S D, WANG D S, DAI Y Q. Symmetric cryptographic protocols for extended millionaires' problem[J]. Science in China(Series F: Information Sciences), 2009, 52(6): 974-982.
- [28] 窦家维, 刘旭红, 王文丽. 有理数域上两方集合的高效保密计算[J]. 计算机学报, 2020, 43(8): 1397-1413.
- [29] DOU J W, LIU X H, WANG W L. Privacy preserving two-party rational set computation[J]. Chinese Journal of Computers, 2020, 43(8): 1397-1413.
- [30] HIDAKA A, SASAZUKI S, GOTO A, et al. Plasma insulin, C-peptide and blood glucose and the risk of gastric cancer: the Ja-

- pan public health center-based prospective study[J]. International Journal of Cancer Journal International Du Cancer, 2015, 136(6): 1402-1410.
- [23] LIU X, CHOO K K R, DENG R H, et al. Efficient and privacy-preserving outsourced calculation of rational numbers[J]. IEEE Transactions on Dependable and Secure Computing, 2018, 15(1):27-39.
- [24] 巩林明, 李顺东, 邵连合, 等. 数轴上保密关系测定协议[J]. 软件学报, 2020, 31(12):3950-3967.
- GONG L M, LI S D, SHAO L H, et al. Protocols for secure test on relationship on number axis[J]. Journal of Software, 2020, 31(12): 3950-3967.
- [25] 窦家维, 王文丽, 刘旭红, 等. 有理区间的安全多方计算与应用[J]. 电子学报, 2018, 46(9): 2057-2062.
- DOU J W, WANG W L, LIU X H, et al. Secure multiparty computation of rational interval and its applications[J]. Chinese Journal of Electronics, 2018, 46(9): 2057-2062.

[作者简介]



刘旭红（1992-），女，山西吕梁人，上海体育学院助教，主要研究方向为应用数学、密码学理论、安全多方计算、密码学原理与算法。



孙晨（1985-），女，山东菏泽人，上海体育学院副教授，主要研究方向为控制科学与工程、动作识别、健康数据挖掘。