软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software, [doi: 10.13328/j.cnki.jos.006093]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

高效安全向量计算及其推广*

刘旭红

(上海体育学院 经济管理学院,上海 200438)

通讯作者: 刘旭红, E-mail: liuxuhong@sus.edu.cn

摘 要: 安全多方计算是密码学的一个重要研究方向,也是目前国际密码学界的研究热点.因为许多实际问题都可以用向量来描述,研究向量的保密计算具有重要的理论与实际意义.目前关于向量保密计算问题大多是在整数集上进行研究,关于有理数向量问题的研究很少.本文主要研究有理数域上向量的安全多方计算问题,包括向量点积、向量相等、向量优势等问题,设计了安全高效的计算协议,扩大了向量保密计算的应用范围.对本文协议的安全性分析和效率分析表明本文协议在安全性和效率方面与现有协议相比具有明显优势.最后,利用所设计的协议解决了一些新的向量问题和计算几何问题.

关键词: 密码学;安全多方计算;向量点积;向量优势;推广应用

中图法分类号: TP311

中文引用格式: 刘旭红.高效安全向量计算及其应用.软件学报. http://www.jos.org.cn/1000-9825/6093.htm
英文引用格式: LIU XH. Efficient Secure Vector Computation and Its Applications. Ruan Jian Xue Bao/Journal of Software, (in Chinese). http://www.jos.org.cn/1000-9825/6093.htm

Efficient Secure Vector Computation and Its Extension

LIU Xu-Hong

(School of economics and management, Shanghai University of Sport, Shanghai 200438, China)

Abstract: Secure multiparty computation is an important research topic of cryptography and focus of the international cryptographic community. Many practical problems can be described using vectors. Therefore, it is of important theoretical and practical significance to study secure multiparty vector computation. Existing secure vector computation protocols are for integer vectors, and there are few works on rational vectors. To fill the gap, we study the secure multiparty computation for rational vectors, including computing the dot product of two vectors, determining whether two vectors are equal, and whether one vector dominates another. We propose efficient protocols for these problems and extend the application of secure vector computation. We prove these new protocols are secure. The efficiency analysis shows that our new protocols outperform existing protocols. Finally, we apply these new protocols to solve some new vector computation problems and some computational geometric problems.

Key words: cryptography; secure multiparty computation; scalar product; vector dominance; popularization and application

信息技术的迅猛发展将人类带入了信息社会.利用互联网获取信息、交换信息以及进行联合计算已成为信息社会一个非常重要的特征,与此相应,信息安全问题日益严重,也使得信息安全成为信息科学的研究热点.信息共享集信息获取、传递、加工、利用于一体,信息共享中的信息安全受到人们的特别关注.安全多方计算(Secure Multiparty Computation, SMC)是实现隐私信息共享的关键技术,是国际密码学界研究的热点.1982 年姚期智[1]提出了两个参与者的安全多方计算问题,1988 年 Ben 和 Goldwasser[2]引入了多个参与者的安全多方计算问题并对其进行了深入研究.

安全多方计算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.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.国际著名计算机科学家、密码学家 Cramer[3]指出: 如果能够

* 收稿时间: 2019-12-17; 修改时间: 2020-02-09, 2020-04-04; 采用时间: 2020-05-28; jos 在线出版时间: 2021-04-18

保密地计算任何函数,计算科学就拥有了一个新的强有力的工具.Goldreich 等[4,5]的工作奠定了安全多方计算的理论基础,他们证明了所有安全多方计算问题在理论上都是可解的,并提出了通用的解决方案,但因为效率的原因,应用通用解决方案来解决某个具体的安全多方计算问题是不实际的.根据具体问题的特点研究具体的解决方案是解决实际问题的途径.近年来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并研究其解决方案,推进了安全多方计算研究的发展.目前主要研究的问题有保密信息比较[6-7]、保密数据挖掘[8-10]、保密几何计算[11-13]、保密数据库查询[14-16]、保密竞拍[17-19]等.

向量是一个高度抽象的数学对象,许多问题可以抽象为向量计算问题[20-23],例如:(1)在信息检索、数据挖掘、机器翻译、文本复制等领域有着广泛应用的文本相似度计算,通过对涉及到的文本进行向量化,然后计算向量间的余弦相似度,进而得到文本的相似度;(2)在计算几何中,关于图形相似性的保密计算问题、空间位置关系的保密计算问题中,往往需要对原始数据进预处理,该过程不可避免存在有理数域上的数据.因此,计算几何中的诸多问题可以抽象为有理数向量计算问题进行解决;(3)统计学在金融工程、经济等领域具有重要的基础作用,其中期望、方差、加权平均值等问题的计算都可以抽象为向量计算.因此利用向量运算可以解决文本相似度问题、自动问答系统、论文抄袭识别以及几何图形检索、经济学等方面的许多问题.一个向量具有多个分量,各分量可以代表不同的含义,很多情形下对向量进行某种运算等价于对其分量分别做相应的运算.正因为向量计算的这种特殊性质,使其成为科学研究中一个基本而且重要的研究方法[24].

目前关于向量问题的保密计算主要包括保密计算向量点积[25-34],保密判定向量相等[35]及保密判定向量优势[36-38]等问题.其中,向量点积问题的研究最为广泛.文献[25]基于一个理想的预处理过程设计了向量点积协议,但实现预处理阶段需要第三方参与者;文献[26]设计了在多密钥的云环境下进行计算的向量点积协议,计算复杂性较高;文献[27]基于第三方服务器以及可逆矩阵设计了两个点积协议,在基于可逆矩阵的协议中,每个参与者会泄露其 n 维向量的 $n/2$ 个线性关系式(当向量维数或数据范围较小时该协议存在一定的安全隐患),且需要计算逆矩阵而导致其计算成本较高;文献[28]设计了一个基于多项式共享的点积协议,由于协议需要雇用半诚实的第三方,协议容易受到合谋攻击;文献[29]在文献[30,31]的基础上设计了一个偶数维点积协议,该协议的计算复杂性较低,但泄露了私有数据部分和的某些信息,仅适用于偶数维向量点积计算.文献[30,31,32]利用不同的公钥加密算法设计了一些点积协议,其中文献[30]中基于 Paillier 加密算法设计的点积协议是目前应用得较多的一个经典的点积协议;文献[31]基于 Damgard 等人[39]的加密算法设计的点积协议与[30]中协议的设计方案类似,只是将[30]中向量数据的正整数范围推广到负整数范围,这些协议的计算复杂性都较高.文献[32]基于 Goldwasser-Micali 密码体制和不经意过滤器设计了一个有关二进制向量的点积协议,此协议相比[30,31]的协议计算效率有所提高,但数据的适用范围却受到很大限制.文献[33,34]也利用公钥加密体制设计了一些点积协议,这些协议均需要借助于云服务器进行一些计算,由于应用公钥加密体制而使得计算复杂性依然较高,且只适用于正整数范围内的向量计算.

关于两个向量相等或向量优势保密判定问题,最直接的想法是将其对应分量分别进行比较,但这样做显然会泄露很多不应泄露的信息.文献[35]基于字母表与连接运算和散列函数设计了向量相等保密判定协议,该协议仅适用于向量的分量为正整数的情形,且判定结果可能出现单边错误.文献[36]通过多次调用百万富翁协议解决向量优势问题,但解决方案效率较低且有信息泄露,为了克服[36]的缺陷,文献[37]通过引进茫然的第三方,将向量优势问题扩展为向量优势统计问题,这样做增加了协议的通信轮数,也会泄露私密向量的部分信息;文献[38]也设计了一个向量优势统计问题的保密判定协议,与[37]相比提高了计算效率,但当向量 $X=(x_1,...,x_n)$ 和 $Y=(y_1,...,y_n)$ 不具有向量优势关系时会泄露两向量中具有关系 $x_i>y_i$ 的分量数目,无法获得向量优势问题的安全解决方案.

目前,已有的向量计算协议大多是在整数集上设计构造的,而实际生活中遇到的很多问题需要转化为有理数域中的向量问题解决.由于已有的有理数向量的保密计算协议还较少,且计算效率和安全性都不高,为此本文对有理数域上的向量问题进行深入研究,并提出安全高效的解决方案.

本文的贡献如下:

(1) 利用基本代数学知识设计了简单高效的向量点积协议,并以此为基础设计了向量相等以及向量优势保密判定协议.

(2) 本文所设计的协议仅需要基本的算术运算,未利用任何公钥加密方案,因而具有很高的计算效率.利用模拟范例证明了协议在半诚实模型下是安全的.理论分析和实验结果都表明本文协议与已有相关协议相比具有较高的安全性和计算效率.

(3) 本文协议适用于计算有理数向量,具有广泛适用性.通过实例说明应用本文协议的设计思想,对更广泛的向量计算或其他实际应用问题可设计构造安全高效的解决方案.

本文第 1 节主要介绍与本文有关的基本概念和记号.第 2 节、第 3 节和第 4 节分别详细描述向量点积保密计算协议、向量相等保密判定协议和向量优势保密判定协议的具体步骤,以及各协议的正确性和安全性分析.第 5 节详细分析和比较本文协议的效率.第 6 节介绍以本文协议为基础,可以解决其他安全多方计算问题.第 7 节总结全文.

1 预备知识

本节介绍一些与本文有关的基本概念和记号,本部分内容基本取自文献[5].

双方计算:一个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 $f:(x,y)\longrightarrow(f_1(x,y),f_2(x,y))$.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 $f=(f_1,f_2)$.

理想模型:假设有一个可信的第三者(Trusted Third Party-TTP),他在任何情况下都不会泄露不该泄露的信息.两个参与者 P_1,P_2 分别将各自的私密数据 x,y 告诉可信的第三者,可信的第三者自己单独计算函数 $f(x,y)=(f_1(x,y),f_2(x,y))$,然后将计算结果 $f_1(x,y),f_2(x,y)$ 分别告诉参与者 P_1,P_2.参与者 P_1,P_2 除了从协议得到可信的第三者发送给自己的计算结果外得不到任何其它信息.上面借助于可信的第三者保密计算函数 $f(x,y)$ 的协议称为理想的两方保密计算协议(简称理想协议或 TTP 协议).理想协议是最简单并且安全性最高的保密计算协议,任何一个计算 $f(x,y)$ 的实际保密计算协议的安全性都不可能超过理想协议,实际保密计算协议可以通过和理想协议进行比较来了解其安全性.理想多方保密计算方案虽然既简单又安全,但在应用中经常受到很大限制,原因是在网络环境中要找到一个可信的第三者一般不是一件容易的事,利用可信的第三者也可能需要付出经济、时间等代价.

半诚实模型:所谓半诚实参与者是指那些在协议的执行过程中按照协议要求忠实地履行协议的参与者,但他们可能会记录下协议执行过程中收集到的所有信息,在协议执行后试图根据记录的信息推算出其他参与者的输入.如果所有的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.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.本文假设所有的参与者都是半诚实的.

模拟范例方法: 模拟范例是安全多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.当 $P_i(i=1,2)$ 的输入为 x_i 时,在执行协议 π 的过程中 P_i 所得到的信息序列记为:

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

其中 r^i 表示 P_i 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.

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

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

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

上述两式成立, 则称协议 π 保密地计算了函数 f, 其中 $\overset{c}{\equiv}$ 表示计算上不可区分.

要证明一个两方保密计算协议是安全的, 就必须构造满足上述两式的模拟器 S_1 和 S_2. 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.

2 高效的向量点积保密计算协议

定义 2:假设两个参与者 Alice 和 Bob 分别拥有私密的 n 维向量 $X=(x_1,...,x_n)$ 和 $Y=(y_1,...,y_n)$,他们希望合作保密计算两向量的点积 $X\cdot Y=x_1y_1+...+x_ny_n$,协议结束后,如果输出结果为 $X\cdot Y$,称这样的协议为向量点积协议;如果协议结束后,其中一方得到 $s\neq 0$,另一方得到 $sX\cdot Y$ 或 $X\cdot Y+s$,我们称这样的协议为共享向量点积协议.

目前已有的关于向量点积(或共享点积)问题的保密计算研究存在的主要问题是以公钥加密方案为基础设计的协议安全性较好但复杂性都很高,而避免应用公钥加密系统的协议大多都有不同程度的信息泄露.下面我们首先应用基本的代数知识和一定的设计技巧设计一个高效安全的向量点积(共享)协议.在下文中,如果一个向量的所有分量均为有理数,称这样的向量为有理数向量.

2.1 共享向量点积协议

我们首先构造一个高效的共享向量点积保密计算协议,并证明该协议的安全性.

协议 1 共享向量点积保密计算协议.

输入: Alice 输入有理数向量 $X=(x_1,...,x_n)$, Bob 输入有理数向量 $Y=(y_1,...,y_n)$.

输出: Alice 输出 s ,Bob 输出 $F(X,Y)=sX\cdot Y$.

1. Alice 将向量 $X=(x_1,...,x_n)$ 按下面方式进行分解:随机选取有理数 a_i 和有理数向量 $X_i=(x_{i1},...,x_{in})(i\in[1,t]=\{1,...,t\},2\le t\le n+1)$, 使得 $X=a_1X_1+...+a_tX_t$ 且 $a_1+...+a_t\neq 0$. Alice 将向量 $X_1,...,X_t$ 发送给 Bob.

2. (a) Bob 随机选取有理数 b_j 和有理数向量 $Y_j=(y_{j1},...,y_{jn})(j=1,2)$,使得 $Y=b_1Y_1+b_2Y_2$.

(b) Bob 选取非零随机有理数 k_1,k_2,r_1,r_2 ,计算

$$z_{11}=k_1X_1\cdot Y_1+r_1,...,z_{1t}=k_1X_t\cdot Y_1+r_1,$$
$$z_{21}=k_2X_1\cdot Y_2+r_2,...,z_{2t}=k_2X_t\cdot Y_2+r_2,$$

将 $(z_{11},...,z_{1t})$ 和 $(z_{21},...,z_{2t})$ 发送给 Alice.

3. Alice 计算

$$z_1=s(a_1z_{11}+...+a_tz_{1t}),$$
$$z_2=s(a_1z_{21}+...+a_tz_{2t}),$$

其中 $s=\frac{1}{a_1+...+a_t}$.并将 z_1,z_2 发送给 Bob.

4. Bob 计算

$$z=b_1\frac{(z_1-r_1)}{k_1}+b_2\frac{(z_2-r_2)}{k_2}.$$

5. Alice 输出 s, Bob 输出 z .

协议 1 的正确性 我们只需证明 $z=sX\cdot Y$ 成立即可.

根据内积的运算性质,对于任意的有理数向量 X 和 Y, 以及关于 X 和 Y 的任意分解方式

$$X=a_1X_1+...+a_tX_t, Y=b_1Y_1+b_2Y_2$$

均有下面等式成立:

$$z_1=s(a_1z_{11}+...+a_tz_{1t})=sk_1X\cdot Y_1+r_1$$
$$z_2=s(a_1z_{21}+...+a_tz_{2t})=sk_2X\cdot Y_2+r_2$$

则

$$z = b_1\frac{(z_1 - r_1)}{k_1} + b_2\frac{(z_2 - r_2)}{k_2} = b_1(sX \cdot Y_1) + b_2(sX \cdot Y_2)$$
$$= sX \cdot (b_1Y_1 + b_2Y_2) = sX \cdot Y.$$

因此,协议 1 是正确的.

协议 1 的安全性 为了分析协议 1 的安全性,需要考察在协议执行过程中每个参与者的私密向量的安全性,并详细分析在协议执行后可能推断出的潜在信息.

首先考虑 Alice 向量的安全性.Alice 秘密将 X 分解为 $X = a_1X_1 + ... + a_tX_t$, Bob 在整个协议的执行过程中得到 Alice 发送给他的 $X_1,...,X_t$ 以及 z_1, z_2. Bob 根据分解式 $X = a_1X_1 + ... + a_tX_t$, 可得到对应的方程组:

$$\begin{cases} x_1 = a_1x_{11} + ... + a_tx_{t1}, \\ \quad \cdots\cdots \\ x_n = a_1x_{1n} + ... + a_tx_{tn}. \end{cases} \tag{1}$$

当 $t < n$ 时,由(1)中的 $t+1$ 个方程联立可能消去 $a_1,...,a_t$ 而得到 X 向量中 $t+1$ 个分量之间的一个线性关系式.而当 $t = n$ (或 $t = n+1$)时,方程组(1)的 n 个方程含有 $2n$ (或 $2n+1$)个未知数 $x_1,...x_n$ 和 $a_1,...,a_t$, 由线性方程组的求解过程可知,需要 $n+1$ (或 $n+2$)个方程联立才可能消去私密数据 $a_1,...,a_t$ 而得到 X 向量中 $n+1$ (或 $n+2$)个分量之间的一个线性关系式.因此当 $t = n$ (或 $t = n+1$)时,根据方程组(1),Bob 得不到 Alice 向量 X 的任何信息.

Bob 根据 $z_1 = sk_1X \cdot Y_1 + r_1, z_2 = sk_2X \cdot Y_2 + r_2$, 有:

$$\begin{cases} \dfrac{z_1 - r_1}{k_1} = sX \cdot Y_1 = s(x_1y_{11} + ... + x_ny_{1n}), \\ \dfrac{z_2 - r_2}{k_2} = sX \cdot Y_2 = s(x_1y_{21} + ... + x_ny_{2n}), \end{cases} \tag{2}$$

由于(2)中两个方程含有 $n+1$ 个未知数,Bob 通过联立两个方程,可得到

$$\frac{x_1y_{11} + ... + x_ny_{1n}}{x_1y_{21} + ... + x_ny_{2n}} = l, \tag{3}$$

即 $x_1(y_{11} - ly_{21}) + ... + x_n(y_{1n} - ly_{2n}) = 0$.

进一步,当方程组(1)和(2)联立时,Bob 可得到关于 $a_1,...,a_t$ 的一个线性关系式:

$$\begin{aligned} &a_1[x_{11}(y_{11} - ly_{21}) + ... + x_{1n}(y_{1n} - ly_{2n})] + ... \\ &+a_t[x_{t1}(y_{11} - ly_{21}) + ... + x_{tn}(y_{1n} - ly_{2n})] = 0, \end{aligned} \tag{4}$$

将(4)式与方程组(1)联立得到一个以 $x_1,...,x_n$ 和 $a_1,...,a_t$ 为未知数的 $n+1$ 个方程.因此,当 $t < n$ 时,Bob 由(4)式与方程组(1)联立的 $n+1$ 个方程,最多得到 X 向量中 t 个分量之间的线性关系式;当 $t = n$ (或 $t = n+1$)时,Bob 仅能得到关于向量 X 的一个线性关系式(3).综上所述,Alice 向量 X 是安全的,且向量 X 的安全性与 t 值成正比关系.

下面考虑 Bob 向量的安全性.在协议执行中,Alice 仅获得 Bob 发送的信息 $(z_{11},...,z_{1t})$ 和 $(z_{21},...,z_{2t})$, 即有:

$$\begin{cases} z_{11} = k_1(x_{11}y_{11} + ... + x_{1n}y_{1n}) + r_1, \\ \quad \cdots\cdots \\ z_{1t} = k_1(x_{t1}y_{11} + ... + x_{tn}y_{1n}) + r_1, \end{cases} \tag{5}$$

和

$$\begin{cases} z_{21} = k_2(x_{11}y_{21} + ... + x_{1n}y_{2n}) + r_2, \\ \quad \cdots\cdots \\ z_{2t} = k_2(x_{t1}y_{21} + ... + x_{tn}y_{2n}) + r_2, \end{cases} \tag{6}$$

当 $t = n$ 时,Alice 设定矩阵 $A = (X_1,...,X_n)$ 存在可逆矩阵 A^{-1} ,通过计算

$$(z_{11},...,z_{1n})A^{-1}=k_1\cdot Y_1+r_1(1,...,1)A^{-1}$$
$$(z_{21},...,z_{2n})A^{-1}=k_2\cdot Y_2+r_2(1,...,1)A^{-1}$$

得到

$$\begin{aligned}Y&=\frac{b_1}{k_1}(k_1\cdot Y_1)+\frac{b_2}{k_2}(k_2\cdot Y_1)\\&=\frac{b_1}{k_1}[(z_{11},...,z_{1n})A^{-1}-r_1(1,...,1)A^{-1}]+\frac{b_2}{k_2}[(z_{21},...,z_{2n})A^{-1}-r_2(1,...,1)A^{-1}],\end{aligned}\tag{7}$$

其中 $r_1,r_2,\frac{b_1}{k_1},\frac{b_2}{k_2}$ 均为 Bob 的私有数据,(7)式中包含 n 个方程和 $n+4$ 个未知数,所以 Alice 根据(7)式无法确定 Bob 向量 Y. 但当 $n>4$ 时,Alice 可得到向量 Y 中 4 个分量间的线性关系式.

当 $t<n$ (或 $t=n+1$)时,首先矩阵 $A=(X_1,...,X_t)$ 不存在可逆矩阵,则关系式(7)不存在,消去 Bob 的私有随机数 $r_1,r_2,\frac{b_1}{k_1},\frac{b_2}{k_2}$ 也无从谈起,所以 Alice 无法得到 Bob 向量 Y 的信息.其次,方程组(5)(或(6))中有 t 个方程,但包含有 $n+2$ 个未知数 $y_{11},...,y_{1n}$ 以及 k_1,t_1 (或 $y_{21},...,y_{2n}$ 以及 k_2,t_2),由于 $n+2>t$, 且所有未知数为有理数,因此即使 Alice 有无限的计算能力,也不可能通过求解(5)(或(6))直接得到向量 Y_1 (或 Y_2),从而无法得到向量 Y. 但是 Alice 根据方程组(5)可以得到以下关系式:

$$\begin{cases}\dfrac{z_{13}-z_{11}}{z_{12}-z_{11}}=\dfrac{(x_{31}-x_{11})y_{11}+...+(x_{3n}-x_{1n})y_{1n}}{(x_{21}-x_{11})y_{11}+...+(x_{2n}-x_{1n})y_{1n}},\\ \cdots\cdots\\ \dfrac{z_{1t}-z_{11}}{z_{12}-z_{11}}=\dfrac{(x_{t1}-x_{11})y_{11}+...+(x_{tn}-x_{1n})y_{1n}}{(x_{21}-x_{11})y_{11}+...+(x_{2n}-x_{1n})y_{1n}},\end{cases}\tag{8}$$

方程组(8)是以 $y_{11},...,y_{1n}$ 为未知数,含有 $t-2$ 个方程的线性方程组.由于 $t-2<n$, Alice 最多可求得关于向量 Y_1 的任意 $n-t+3$ 个分量的线性关系式(同理,Alice 最多可求得关于向量 Y_2 的任意 $n-t+3$ 个分量的线性关系式). 进一步,由 $Y=b_1Y_1+b_2Y_2$ 可知,Alice 需要将向量 Y_1 和 Y_2 联立才能求解向量 Y, 但 Alice 仅知道关于向量 Y_1 和 Y_2 各自分量独立的线性关系式,且不知道随机数 b_1,b_2 的值,所以无法联立求解,故无法得到向量 Y 的任何信息.综上所述,当 $t<n$ (或 $t=n+1$)时,Alice 得不到 Bob 向量 Y 的信息,向量 Y 是安全的.

根据上面的分析,当 $t<n$ 时,Bob 向量 Y 是安全的,Bob 最多得到向量 X 中 t 个分量之间的关系式;当 $t=n$ 时,Alice 最多得到向量 Y 中 4 个分量之间的关系式,Bob 仅能得到关于向量 X 的一个关系式(3);当 $t=n+1$ 时,Bob 向量 Y 是安全的,Bob 仅能得到关于向量 X 的一个关系式(3).因此 $t=n+1$ 时,协议 1 的安全性最好,且与理想模型下协议相比仅有微小差别:Bob 能得到 Alice 向量 X 的一个线性关系式.

关于协议 1 的安全性有下面的定理 1.

定理 1 共享向量点积保密计算协议 1 是安全的.

证明 下面应用模拟范例严格证明定理 1, 即需要构造模拟器 S_1 (或 S_2), 使两式成立.

首先构造 S_1. 接收到输入 (X,s) 后, S_1 按以下方式运行:

(i) S_1 首先任意选择有理数向量 $Y'=(y'_1,...,y'_n)$, 并随机选取有理数 b'_j 和有理数向量 $Y'_j=(y'_{j1},...,y'_{jn})(j=1,2)$, 使得 $Y'=b'_1Y'_1+b'_2Y'_2$.

(ii) S_1 任意选取非零随机有理数 k'_1,k'_2,r'_1,r'_2, 计算

$$z'_{11}=k'_1X_1\cdot Y'_1+r'_1,...,z'_{1t}=k'_1X_t\cdot Y'_1+r'_1,$$
$$z'_{21}=k'_2X_1\cdot Y'_2+r'_2,...,z'_{2t}=k'_2X_t\cdot Y'_2+r'_2,$$

由于在协议执行中,

$$view_1^{\pi}(X,Y)=(X,(z_{11},...,z_{1t}),(z_{21},...,z_{2t}),s),$$

而 S_1 在模拟过程中产生的信息序列为

$$S_1(X,f_1(X,Y))=(X,(z'_{11},...,z'_{1t}),(z'_{21},...,z'_{2t}),s).$$

由于有理数 b_j,k_1,k_2,r_1,r_2 和有理数向量 $Y_j=(y_{j1},...,y_{jn})(j=1,2)$ 是 Bob 随机选取的,对 Alice 来说,有

$$b_j\overset{c}{\equiv}b_j',\quad Y_j\overset{c}{\equiv}Y_j',$$

$$k_1\overset{c}{\equiv}k_1',\quad k_2\overset{c}{\equiv}k_2',\quad r_1\overset{c}{\equiv}r_1',\quad r_2\overset{c}{\equiv}r_2'$$

因此 $(z_{11},...,z_{1t})\overset{c}{\equiv}(z_{11}',...,z_{1t}'),(z_{21},...,z_{2t})\overset{c}{\equiv}(z_{21}',...,z_{2t}')$, 故

$$\{S_1(X,f_1(X,Y))\}_{x_i,y_i\in Q}\overset{c}{\equiv}\{view_1^{\pi}(X,Y)\}_{x_i,y_i\in Q}.$$

接收到输入 $(Y,f_2(X,Y)=F(X,Y))$ 后, S_2 按以下方式运行:

(i) S_2 首先任意选择有理数向量 $X'=(x_1',...,x_n')$, 并随机选择有理数 a_i' 和有理数向量 $X_i'=(x_{i1}',...,x_{in}')(i\in[1,t]=\{1,...,t\},,2\le t\le n+1)$, 使得 $F(X',Y)=F(X,Y),X'=a_1'X_1'+...+a_t'X_t'$ 且 $a_1'+...+a_t'\neq 0$.

(ii) S_2 计算

$$z_{11}'=k_1X_1'\cdot Y_1+r_1,...,z_{1t}'=k_1X_t'\cdot Y_1+r_1,$$
$$z_{21}'=k_2X_1'\cdot Y_2+r_2,...,z_{2t}'=k_2X_t'\cdot Y_2+r_2,$$

以及

$$z_1'=s'(a_1'z_{11}'+...+a_t'z_{1t}'),$$
$$z_2'=s'(a_1'z_{21}'+...+a_t'z_{2t}'),$$

其中 $s'=\frac{1}{a_1'+...+a_t'}$.

(iii) S_2 计算

$$z'=b_1\frac{(z_1'-r_1)}{k_1}+b_2\frac{(z_2'-r_2)}{k_2}.$$

由于在协议执行中,

$$view_2^{\pi}(X,Y)=(Y,(X_1,...,X_t),(z_1,z_2),F(X,Y)),$$

而 S_2 在模拟过程中产生的信息序列为

$$S_2(Y,f_2(X,Y))=(Y,(X_1',...,X_t'),(z_1',z_2'),F(X',Y)).$$

首先, 由于 $a_i(i\in[1,t])$ 为 Alice 任意选取的随机数,因此 $a_i'\overset{c}{\equiv}a_i,s'\overset{c}{\equiv}s$, 故有 $(X_1',...,X_t')\overset{c}{\equiv}(X_1,...,X_t),(z_1',z_2')\overset{c}{\equiv}(z_1,z_2)$. 又因为 $F(X',Y)=F(X,Y)$, 因此

$$\{S_2(Y,f_2(X,Y))\}_{x_i,y_i\in Q}\overset{c}{\equiv}\{view_2^{\pi}(X,Y)\}_{x_i,y_i\in Q}.$$

证毕.

注解 1 设定 $t=n+1$, 如果在协议 1 中 Alice 取 $s=1$, 即 $a_1+...+a_t=1$, 这时协议输出结果为 $F(X,Y)=X\cdot Y$, 协议 1 成为一个点积协议.此时 Bob 收到 Alice 发送的数据 z_1,z_2 中不再包含未知数 s, Bob 最多可得到两个关系式:

$$\frac{z_1-r_1}{k_1}=X\cdot Y_1=x_1y_{11}+...+x_ny_{1n},$$

$$\frac{z_2-r_2}{k_2}=X\cdot Y_2=x_1y_{21}+...+x_ny_{2n}.$$

且方程组(1)与 $a_1+...+a_t=1$ 联立仍无法得到关于向量 X 的任何信息,所以 Alice 向量 X 仍是安全的.在此过程中 Alice 收到 Bob 发送的数据 $(z_{11},...,z_{1t})$ 和 $(z_{21},...,z_{2t})$ 没有发生改变,因此向量 Y 的安全性没有受到影响.

注解 2 由于利用向量分解的思想,协议 1 中参与方的向量至少需分解为两个随机向量,由此知对于维数 $n\ge 3$ 的向量点积问题可直接利用本文协议 1 进行解决.对于维数 $n=2$ 的向量点积问题,用本文协议 1,Bob 仍仅得到关于 Alice 向量 X 的一个线性关系式;Alice 根据 (z_{11},z_{12}) (或 (z_{21},z_{22}))无法消去随机数 k_1,r_1 (或 k_2,r_2),所以 Alice 得不到向量 Y_1,Y_2 的任何信息,即得不到向量 Y 的任何信息.因此,本文协议 1 适用于 $n\ge 2$ 的向量点积问题.

如果参与计算的向量维数较小且限制各分量仅能取值 0 或 1 时,本文协议 1 不能安全的解决向量点积问题.

例如,当 $n=3$ 且 $x_i,y_i\in\{0,1\},i=1,2,3$ 时,向量 $X=(x_1,x_2,x_3)$ 的可能取值仅有 8 种情形.Bob 得到的(3)中含有三个未知数 x_1,x_2,x_3,通过直接代入 0 和 1 进行测试,可知 (x_1,x_2,x_3) 的有些取值满足(3),有些不满足,因此就泄露了向量 X 的部分信息.文献[27]基于可逆矩阵设计的共享点积协议也存在类似安全性问题,当向量维数 n 较小且限制各分量取值为整数时,任一方参与者可通过其获得的 $n/2$ 个等式推导获取对方足够多的私有信息.文献[29]设计了偶数维共享向量点积协议也存在安全性问题,参与者分别泄露其私密向量相邻分量之间的关系式 $x_{2k-1}+x_{2k}$ 和 $y_{2k-1}+y_{2k}$ 的值.文献[28,30-34]中设计的共享点积协议,当向量维数较小且各分量限制取值为 0 或 1 时,每个参与者都可在一定程度上推导获取对方向量的某些信息,只是信息泄露的程度有所不同.针对这种向量维数较小且各分量限制取值为 0 或 1 的情形,可以利用公钥加密方案进行解决,且计算效率较高.因此,本文协议 1 适用于向量维数 $n\geq 2$ 且向量的分量取值无限制范围的情形.

3 向量相等保密判定问题

问题描述 Alice 和 Bob 分别具有有理数向量 $X=(x_1,...,x_n)$ 和 $Y=(y_1,...,y_n)$, 双方想要保密判定两个向量 X 和 Y 是否相等.如果不相等,不应向对方泄露向量 X 或 Y 的任何信息.

计算原理 首先证明下面结论.

命题 1 对于任意两个有理数向量 $X=(x_1,...,x_n)$ 和 $Y=(y_1,...,y_n)$, $X=Y$ 的充要条件是下面等式成立:

$$|X|^2+|Y|^2=2X\cdot Y, \tag{9}$$

其中 $|X|^2=x_1^2+...+x_n^2$, $|Y|^2=y_1^2+...+y_n^2$.

证明 由于

$$\begin{aligned} X=Y &\Leftrightarrow |X-Y|=0\\ &\Leftrightarrow (x_1^2+...+x_n^2)+(y_1^2+...+y_n^2)-2(x_1y_1+...+x_ny_n)=0\\ &\Leftrightarrow |X|^2+|Y|^2=2X\cdot Y, \end{aligned}$$

因此命题 1 得证.

命题 1 是判定两向量是否相等的基本原理,即将判定两向量 X,Y 是否相等的问题转化为判定条件(9)是否成立的问题.下面我们以协议 1 为基础构造向量相等保密判定协议,Alice 和 Bob 调用协议 1(约定 $t=n+1$),Alice 得到随机数 $s>2$, Bob 得到 $z=sX\cdot Y$. Bob 计算 $u=z-|Y|^2=sX\cdot Y-|Y|^2$, 并将 u 发送给 Alice;Alice 计算

$$\begin{aligned} w&=\frac{s}{s-2}(u-|X|^2)=\frac{s}{s-2}(sX\cdot Y-|X|^2-|Y|^2)\\ &=\frac{s}{s-2}((s-2)X\cdot Y+(2X\cdot Y-|X|^2-|Y|^2))\\ &=sX\cdot Y+\frac{s}{s-2}(2X\cdot Y-|X|^2-|Y|^2), \end{aligned}$$

并将 w 发送给 Bob.Bob 判断 w 是否等于 $z=sX\cdot Y$. 若 $w=z$, 则 $2X\cdot Y-|X|^2-|Y|^2=0$, 即 $X=Y$; 否则 $X\neq Y$. 下面为叙述方便,定义二元谓词:

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

协议 2 向量相等保密判定协议.

输入: Alice 和 Bob 分别输入有理数向量 $X=(x_1,...,x_n)$ 和 $Y=(y_1,...,y_n)$.

输出: Bob 输出 $P(X,Y)$.

1. 将 X,Y 作为协议 1 的输入向量,Alice 和 Bob 调用协议 1,Alice 得到一个随机数 $s>2$, Bob 得到 $z=sX\cdot Y$.
2. Bob 计算 $u=z-|Y|^2$, 并将 u 发送给 Alice.
3. Alice 计算 $w=\frac{s}{s-2}(u-|X|^2)$, 并将 w 发送给 Bob.
4. Bob 输出 $y=P(w,z)$.

协议 2 的正确性 根据二元谓词 $P(x,y)$ 的定义,需要证明 $w=z \Leftrightarrow X=Y$. 因此只需证明下式成立即可:

$$w=\frac{s}{s-2}(z-|X|^2-|Y|^2).$$

根据向量的运算性质,对于任意有理数向量 X,Y, 以及关于 X 和 Y 的任意关系式 $z=sX\cdot Y$ 均有下面等式成立:

$$w=\frac{s}{s-2}(u-|X|^2)=\frac{s}{s-2}(z-|X|^2-|Y|^2).$$

因此,协议 2 是正确的.

协议 2 的安全性 在协议 2 执行中首先调用协议 1,Alice 得到一个保密随机数 s, Bob 得到保密值 $z=sX\cdot Y$.

首先考虑 Alice 向量的安全性.由协议 1 的安全性证明,在调用协议 1 的过程中 Bob 最多可得到向量 X 的一个线性关系,在协议 2 的后续执行中又收到 Alice 发送的 $w=\frac{s}{s-2}(u-|X|^2)$.当向量 $X\neq Y$ 时,Bob 根据得到的关系式 $w=\frac{s}{s-2}(u-|X|^2)$ 得不到向量 X 的任何信息(其中 s 是 Alice 的私有数据).因此,在协议 2 中 Alice 向量 X 的安全性与协议 1 一致.

下面再考虑 Bob 向量的安全性.在调用协议 1 的过程中 Alice 得到 Bob 向量 Y_1 (和 Y_2)的任意 2 个分量的线性关系式,得不到向量 Y 的信息.在协议 2 的后续执行中,Alice 又获得 Bob 发送的 $u=z-|Y|^2$, 对于 Alice 而言, $y_1,...,y_n$ 作为未知数,由关系式 $u=z-|Y|^2$ 仅能获得关于 Y 的一个二次关系式.

根据上面的分析,在协议 2 的执行过程中,Bob 最多能得到 Alice 向量 X 的一个线性关系式,Alice 最多得到关于 Bob 向量 Y 的一个二次关系式.根据有理数的稠密性,即使参与者有无限的计算能力也无法得到对方的私有向量.协议 2 的安全性与协议 1 类似,安全性很高.

关于协议 2 的安全性仅叙述下面定理 2, 定理 2 的证明类似于定理 1, 故从略.

定理 2 向量相等保密判定协议 2 是安全的.

4 向量优势保密判定问题

向量优势保密判定问题 假设有两个有理数向量 $X=(x_1,...,x_n)$ 和 $Y=(y_1,...,y_n)$, 如果对于所有 $i\in[1,n]$, 关系 $x_i>y_i$ 成立,则称向量 X 关于 Y 具有向量优势,并记为 $X>Y$; 否则,称向量 X 关于 Y 不具有向量优势.向量 X 和 Y 的优势保密判定问题即是要保密判定是否有关系 $X>Y$ 成立.

基本原理 我们知道判断向量 $X>Y$, 即判断

$$Z=X-Y=(x_1-y_1,...,x_n-y_n):=(z_1,...,z_n)$$

中每个分量均大于 0,那么只需判断向量 Z 中最小分量 $\min\{z_1,...,z_n\}>0$ 即可.

为下面叙述方便,定义二元谓词如下:

$$P(X,Y)=\begin{cases}1, & \text{如果} X>Y;\\ 0, & \text{否则.}\end{cases}$$

协议 3 向量优势保密判定协议.

输入: Alice 输入有理数向量 $X=(x_1,...,x_n)$, Bob 输入有理数向量 $Y=(y_1,...,y_n)$.

输出: Alice 和 Bob 输出 $P(X,Y)$.

1. Alice 随机选取有理数向量 $R=(r_1,...,r_n),r_i>0$ 计算

$$Z_1=X+R:=(z_{11},...,z_{1n}),$$

将向量 Z_1 发送给 Bob.

2. Bob 计算

$$Z_2=Z_1-Y:=(z_{21},...,z_{2n}),$$

并随机选取有理数向量 $K=(k_1,...,k_n),k_i>0$ 计算

$$Z_3=(k_1z_{21},...,k_nz_{2n}):=(z_{31},...,z_{3n}),$$

将向量 Z_3 发送给 Alice.

3. Alice 计算

$$Z_4=(\frac{z_{31}}{r_1},...,\frac{z_{3n}}{r_n}):=(z_{41},...,z_{4n}),$$

并随机选取有理数 s 计算

$$Z_5=Z_4+sE:=(z_{51},...,z_{5n}),$$

将 Z_5 发送给 Bob.

4. Bob 计算

$$Z_6=Z_5-K:=(z_{61},...,z_{6n}),$$

并选取向量 Z_6 中的最小分量 z_{min} 发送给 Alice.

5. Alice 比较 z_{min} 和 s, 当 $z_{min}>s$ 时,输出 $P(X,Y)=1$;否则,输出 $P(X,Y)=0$.

协议 3 的正确性 由 Z_1 到 Z_5 的计算过程可知,

$$Z_6=(z_{61},...,z_{6n})=(\frac{k_1(x_1-y_1)}{r_1}+s,...,\frac{k_n(x_n-y_n)}{r_n}+s).$$

进一步由 z_{min} 的定义可知,

$$\begin{aligned}z_{min}>s &\Leftrightarrow \forall i\in[1,n],z_{6i}>s\\ &\Leftrightarrow \forall i\in[1,n],\frac{k_i(x_i-y_i)}{r_i}>0\\ &\Leftrightarrow \forall i\in[1,n],x_i-y_i>0\\ &\Leftrightarrow X>Y,\end{aligned}$$

因此协议 3 是正确的.

协议 3 的安全性 为了分析协议 3 的安全性,需要考察在协议执行过程中每个参与者的私密向量的安全性,并详细分析在协议执行后可能推断出的潜在信息.

首先考虑 Alice 向量的安全性.Bob 得到 Alice 发送的向量 Z_1 和 Z_5, 由 Z_1,Z_5 的定义可知

$$Z_1=(z_{11},...,z_{1n})=(x_1+r_1,...,x_n+r_n),$$

$$Z_5=(z_{51},...,z_{5n})=(\frac{k_1(x_1-y_1)}{r_1}+s+k_1,...,\frac{k_n(x_n-y_n)}{r_n}+s+k_n),$$

分别根据 Z_1 和 Z_5 无法得到向量 X 的任何信息.结合 Z_1,Z_5 可以得到以下等式:

$$\frac{k_i(x_i-y_i)}{z_{1i}-x_i}-\frac{k_j(x_j-y_j)}{z_{1j}-x_j}=(z_{5i}-k_i)-(z_{5j}-k_j)=z_{6i}-z_{6j},\tag{10}$$

其中 x_i,x_j 是未知量,一个方程两个未知数,Bob 根据上述等式无法解得 x_i,x_j, 仅可以得到向量 X 中两个分量 x_i,x_j 间的二次关系式.因此,Alice 向量 X 是安全的.

下面考虑 Bob 向量的安全性.在协议执行中,Alice 仅获得 Bob 发送的向量 Z_3 和最小分量 z_{min}. 根据协议的执行过程可知,

$$Z_3=(k_1(z_{11}-y_1),...,k_n(z_{1n}-y_n)),$$

其中 y_i,k_i 均为未知量,因此 Alice 根据 Z_3 无法得到向量 Y 的任何信息.根据最小分量 z_{min} 的定义可知,Alice 无法获知对于哪个 $i\in[1,n]$ 有 $z_{min}=z_{6i}$, Alice 只能对每一个 $i\in[1,n]$, 求解下式中的 y_i :

$$\begin{cases}z_{min}=z_{6i}=\dfrac{k_i(x_i-y_i)}{r_i}+s\\ z_{3i}=k_i(x_i-y_i+r_i),\end{cases}\tag{11}$$

若将所得的解记为 y_i', 显然 $y_i'=y_i$ 当且仅当 $z_{min}=z_{6i}$, 如果仅有一个 $i\in[1,n]$ 满足 $z_{min}=z_{6i}$, 此时 Alice 仅有 $1/n$ 的概率猜对满足 $y_i'=y_i$ 的 i 值.除此之外,在协议的整个执行过程中,Alice 得不到关于向量 Y 的任何信息.因此,Bob 向量 Y 是安全的.

根据上面的分析及有理数的稠密性可知,协议 3 的安全性和理想模型相比仅有很小的差别.关于协议 3 的安全性仅叙述下面定理 3,定理 3 的证明类似于定理 1,故从略.

定理 3 向量优势保密判定协议 3 是安全的.

5 效率分析与比较

本部分将本文的主要结果与近期较好的工作进行比较,在进行效率分析时,一般地基本算术运算与模乘运算的计算复杂性相比较可忽略不计.文献[25,27,28]均研究了向量点积的保密计算问题,但文献中提出的解决方案需要借助第三方参与者或服务器;文献[25,26,30-34]也研究了向量点积的保密计算问题,但上述文献中提出的解决方案均需要借助不同的密码体制进行保密计算,仅适用于整数范围内的向量计算且计算复杂性较高.文献[29]利用基本算数运算设计了偶数维点积协议,适用于有理数向量的保密计算且计算复杂性较低.因此,本文协议 1 主要与文献[29]进行详细的分析和比较.而关于向量相等保密计算问题和向量优势保密计算问题,现有的相关文献较少,其中文献[35]利用散列函数设计了向量相等保密判定协议,文献[36-38]利用不同的公钥加密算法设计了向量优势协议,这些协议均只适用于整数集上的向量计算.目前还没有关于有理数向量相等问题和有理数向量优势问题的研究文献,因此本文协议 2 仅与文献[35]进行效率分析和比较,本文协议 3 主要与计算效率相对较高的文献[38]进行效率分析与比较.

为了便于分析比较,假设两方参与者的向量都是 n 维的.由于我们所设计的协议仅应用了基本算术运算(普通有理数的加减乘除运算),而已有文献大多是应用公钥加密系统设计协议,所以分析计算复杂性时同时考虑了基本算术运算次数和模乘运算次数.并且我们以通信数据量衡量协议的通信复杂性.

点积协议的分析比较 本文协议 1 中仅应用了有理数的加法及乘法等基本算术运算.具体地,接收到 Alice 的分解向量后,Bob 执行了 $2t(n+1)$ 次乘法以及 $2tn$ 次加法运算,且计算 z 执行了 4 次乘法以及 3 次加法运算. Alice 执行了 $2(t+1)$ 次乘法运算和 $2(t-1)$ 次加法运算,因此协议 1 共需要 $4tn+6t+7$ 次基本算术运算.文献[29]设计了偶数维点积共享协议,协议中 Alice 执行了 $3n/2$ 次乘法运算以及 $9n/2-1$ 次加法运算,Bob 执行了 $2n$ 次乘法运算以及 $9n/2-1$ 次加法运算,因此文献[29]的点积协议共需要 $25n/2-2$ 次基本算术运算.为了具体地将本文协议 1 与文献[29]进行对比,我们通过以下两种情形 1 和情形 2 分别进行效率分析.

情形 1: $n<5, t=2$ 时,协议 1 比文献[29]多执行 $21-9n/2$ 次基本算术运算. $n=2$ 时,协议 1 比文献[29]多执行 12 次基本算术运算; $n=3$ 时,协议 1 比文献[29]多执行 7 次基本算术运算; $n=4$ 时,协议 1 比文献[29]多执行 3 次基本算术运算.因此,在 $n<5$ 的情形下,两种方案的计算效率相近,均可以高效解决维数较小的向量保密计算问题,例如:信息检索、机器翻译等输入数据量较小的场景.

情形 2: $n\geq 5, t=2$ 时,协议 1 比文献[29]少执行 $9n/2-21$ 次基本算术运算. $n=5$ 时,协议 1 比文献[29]少执行 1 次基本算术运算; $n=6$ 时,协议 1 比文献[29]少执行 6 次基本算术运算; $n=7$ 时,协议 1 比文献[29]少执行 11 次基本算术运算.进一步 $n=10$ 时,协议 1 比文献[29]少执行 24 次基本算术运算; $n=15$ 时,协议 1 比文献[29]少执行 46 次基本算术运算; $n=20$ 时,协议 1 比文献[29]少执行 69 次基本算术运算.因此,在 $n\geq 5$ 的情形下,随着向量维数 n 的增加,本文协议 1 的计算效率优势更加明显.故本文协议 1 可以高效解决向量维数大的安全多方计算问题,例如:文本相似度问题、文本挖掘、论文抄袭识别等输入数据量较大的场景.

本文的点积协议 1 需要的通信数据量为 $t(n+2)+2$; 文献[29]的点积协议需要的通信数据量为 $3n$.当 $t=2$ 时,本文协议 1 的通信数据量为 $2n+6$,且随着向量维数 n 的增加,本文协议 1 的通信效率优势更加明显.

当 $t=2$ 时协议 1 的安全性,根据前面的分析,Bob 最多得到 Alice 向量 X 中任意 $t+1=3$ 个分量之间的一个线性关系式,Alice 得不到 Bob 向量 Y 的任何信息.而文献[29]中 Alice 和 Bob 分别能够得到对方私密向量相邻分量之间的关系式 $y_{2k-1}+y_{2k}$ 和 $x_{2k-1}+x_{2k}$ 的值.

因此,在选取 $t=2$ 时,本文协议 1 的安全性和效率与文献[29]相比有所提高,更适用于解决大数据情形下的向量保密计算问题.且本文可以根据具体的向量点积保密计算问题的安全性要求灵活选择分解个数 t. 本文协议 1 具有更高的灵活性和广泛适用性.

向量相等协议的分析比较 本文协议 2 研究了两向量相等的判定问题,就我们所知,目前直接研究向量相等问题的文献较少,文献[35]中的协议 3 可用于比较两向量相等.由于文献[35]的方案仅适用于正整数向量,适用范围有限,且会出现单边错误,而本文协议 2 是基于共享点积协议 1 设计的,效率更高,适用范围也更广泛.

向量优势协议的分析比较 我们将本文协议 3 与主要研究向量优势保密判定问题的文献[38]协议 2 进行比较.本文协议 3 主要应用基本算术运算,其中 Alice(或 Bob)进行的加减法为 $2n$ 次,乘除法运算为 n 次,整个协议需要的基本算术运算各为 $3n$ 次.文献[38]应用 Paillier 公钥加密算法设计的协议 2,基本运算是模乘运算,整个协议共需要 $2(mn+1)\log N+n$ 次模乘运算,其中 N 是加密系统中两个大素数的乘积.一般来说,基本算术运算与模乘运算的计算复杂性相比较可忽略不计.因此本文协议 3 的计算复杂性相比[38]是非常小的.

本文协议 3 需要的通信数据量为 4;文献[38]的协议 2 需要的通信数据量为 $mn+2$.本文协议 3 的通信负载较低.

详细的协议效率比较结果如表 1 所示.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,M(或 B)表示模乘(或基本算术)运算次数,通信复杂性为通信数据量.

Table 1 Efficiency analysis and range analysis of the protocols

表 1 协议的效率分析与适用范围的比较

文献	计算功能	计算复杂性(M 或 B)	通信数据量	适用范围
文献[29]	点积	$25n/2-2$ (B)	$3n$	有理数
文献[38]	优势	$2(mn+1)\log N+n$ (M)	$mn+2$	正整数
本文协议 1	点积	$8n+19$ (B)	$2n+6$	有理数
本文协议 3	优势	$6n$ (B)	4	有理数

根据 Paillier 公钥加密系统, N 是两个大素数的乘积,一般长度为 1024 比特.我们注意到实际生活中有关向量计算问题,向量维数 n 一般都满足 $n \ll \log N$, 在此情形下本文协议 1 的计算复杂性低于已有文献,通信复杂性也不比已有协议的通信复杂性高;协议 3 的计算复杂性远低于已有文献,而我们的协议适用于有理数范围内的向量计算,适用性更广,并且本文协议 1 和协议 3 在有理数向量情形下安全性更高,泄露的信息对安全性影响极小.

协议效率实验测试 前面已从理论上对本文所设计协议的效率进行了全面分析,并与已有相关结果进行了比较.下面进一步进行实验测试,并将本文协议的执行结果与已有的效率较高的协议的执行结果进行比较,在此我们取定协议 1 中的向量分解个数 $t=2$.

(I) **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 @3.30GHz,安装内存 8.00GB,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.

(II) **实验结果** 由于本文协议 1、3 和文献[29]执行的是算术(指数)基本运算,文献[30,37,38]应用的是 Paillier 加密方案,下面分别在两种不同环境下进行仿真实验.

实验设定 Paillier 加密算法中使用的大素数 p,q 的位数为 256 比特,并且统一限定保密数据的范围为[-100, 100].下面分别对本文协议 1 和文献[29,30]的协议,以及本文协议 3 和文献[37,38]的协议进行实际计算,对每种协议在不同维向量下进行多次实验,实验结果随机抽取 50 组数据求取平均值.由于本文协议 1 和协议 3 在进行一次实验时耗时太少无法显示,因此通过对每组数据分别循环运行 10000 次和 100 次求平均值得到结果.结果如表 2 和表 3 所示.

Table 2 Analysis of simulation result of Protocol 1

表 2 协议 1 实验结果分析

	文献[29]	文献[30]	本文协议 1
10000 次实验平均耗时(ms)	90.308	18926.5	63.092

Table 3 Analysis of simulation result of Protocol 3

表 3 协议 3 实验结果分析

	文献[37]	文献[38]	本文协议 3
100 次实验平均耗时(ms)	5798.4	4323.8	11.9033

由表 2 和 3 可知,本文协议 1 和协议 3 的效率较高,优势明显.

6 协议的推广举例

对于前面所设计的协议进行适当修改或者组合,或直接应用前面各协议的设计思想,对于更广泛的科学计算问题或实际应用问题,能够设计构造安全高效的解决方案.下面举例进行说明.

6.1 多方向量相等保密判定问题

问题描述 考虑 m 个参与者 $P_1,...P_m$, $P_i(i\in[1,m])$ 分别具有私密的 n 维有理数向量 X_i. 他们想合作保密判定所有向量是否相等,而不泄露各自的 X_i.

计算原理 类似于命题 1,对于任意 m 个 n 维有理数向量 $X_1,...,X_m$, 这些向量 $X_1=...=X_m$ 的充要条件是下面等式成立:

$$(X_1-X_2)^2+...+(X_1-X_m)^2=0, \tag{12}$$

即:

$$2(X_1\cdot X_2+...+X_1\cdot X_m)-(m-1)|X_1|^2=|X_2|^2+...+|X_m|^2, \tag{13}$$

其中 $|X_i|^2$ 与命题 1 中 $|X|^2$ 的定义相同.

因此,可将判定多方向量 $X_1,...,X_m$ 是否相等的问题转化为判定条件(13)是否成立的问题.下面我们以协议 1 为基础构造多方向量相等保密判定协议.为叙述方便,定义谓词:

$$P(X_1,...,X_m)=\begin{cases}1, & \text{如果} X_1=\cdots=X_m;\\ 0, & \text{否则.}\end{cases}$$

协议 4 多方向量相等保密判定协议.

输入: m 个参与者 $P_i(i\in[1,m])$ 分别输入有理数向量 X_i.

输出: $P(X_1,...,X_m)$.

准备 (a) 对于每一个 $P_i(i\in[2,m])$, P_i 分别随机选取非零有理数 S_i 构造向量 $X'_i=(2S_i^2X_i,-S_i^2)$; 并随机选取有理数 $a_{i1},...,a_{in}$ 和有理数向量 $X_{i1},...,X_{in}$, 使得 $X'_i=a_{i1}X_{i1}+...+a_{in}X_{in}$;

(b) P_1 构造向量 $X'_1=(X_1,|X_1|^2)$, 并随机选取有理数 b_{i1},b_{i2} 和有理数向量 $U_{i1},U_{i2}(i\in[2,m])$, 使得 $X'_1=b_{i1}U_{i1}+b_{i2}U_{i2}$.

1. $P_i(i\in[2,m])$ 和 P_1 分别将向量 X'_i 和 X'_1 作为协议 1 的输入向量,按照准备阶段的(a)和(b)合作调用协议 1(注解 1 的点积协议), P_1 得到 $z_2,...,z_m$, 并计算 $Z=z_2+...+z_m$, 其中 $z_i=X'_i\cdot X'_1$.

2. 每一个 $P_i,i=2,...,m$,

(a) 从 P_{i-1} 处接收到 $w_{i-1}=S_i^2|X_i|^2$;

(b) 将 w_i 与 w_{i-1} 相加,最后得到 $W=w_2+...+w_m$ 并发送给 P_1.

3. P_1 比较 Z 和 W: 如果 $Z=W$, 输出 $P(Z,W)=1$; 否则,输出 $P(Z,W)=0$.

协议 4 的正确性 需要证明 $W=Z\Leftrightarrow X_1=\cdots=X_m$. 根据

$$\begin{aligned}Z&=z_2+...+z_m=X'_2\cdot X'_1+...+X'_m\cdot X'_1\\&=(2S_2^2X_2\cdot X_1-S_2^2|X_1|^2)+...+(2S_m^2X_m\cdot X_1-S_m^2|X_1|^2),\end{aligned}$$

那么, $W-Z=(S_2X_2-S_2X_1)^2+...+(S_mX_m-S_mX_1)^2$. 因此,

$$W = Z \Longleftrightarrow \forall i \in [2,m], S_i X_i - S_i X_1 = 0$$
$$\Longleftrightarrow \forall i \in [2,m], X_i - X_1 = 0$$
$$\Longleftrightarrow X_1 = \cdots = X_m.$$

因此,协议 4 是正确的.

协议 4 的安全性 协议 4 是基于向量点积协议 1 设计的,当不考虑合谋攻击时,由协议 1 的安全性分析可知,参与者 $P_i(i \in [2,m])$ 向量的安全性完全类似于协议 1 中 Alice 向量的安全性.参与者 P_1 向量的安全性完全类似于协议 1 中 Bob 向量的安全性.因此,在协议 4 中各参与者的向量是安全的.

下面我们主要分析协议 4 中的合谋攻击.由于在协议 4 中,参与者 P_1 拥有向量点积 $X_1' \cdot X_2', ..., X_1' \cdot X_m'$, 若 P_1 与其他参与者合谋,很容易判断出哪些向量 $X_i(i \in [2,m])$ 与向量 X_1 相等.因此,我们将仅考虑参与者 P_1 不参与合谋的情况.假设参与者 $P_3, ..., P_m$ 合谋,他们最多可得到参与者 P_2 的 $S_2^2 X_2^2$ 值,由于 S_2 是 P_2 的保密数据,因此无法得到向量 X_2 的任何信息.

根据协议 4 的设计过程可知,协议仅应用基本的算术运算解决多方向量相等保密判定问题,计算效率较高.但是,协议 4 需要限制参与者 P_1 不参与合谋攻击,具有一定的局限性.据我们所知,现有的抵抗合谋攻击的保密计算协议均需要借助公钥加密方案,例如,利用 ElGamal 同态加密方案可以构造门限密码体制,进而设计可以抵抗合谋攻击的多方保密计算协议.现有信息论安全的协议很难有效抵抗合谋攻击,均需要设定一些限制条件.

6.2 向量和矩阵的保密计算问题

假设 Alice 有一个 m 维有理数向量 $X = (x_1, ..., x_m)$, Bob 有一个 $m \times n$ 阶的有理数矩阵 $A = (a_{ij})$, Alice 和 Bob 想要保密计算 X 和 A 的乘积.

记矩阵 A 的第 i 个列向量为 $A_i(i = 1, ..., n)$, 则有下面关系:

$$Y = XA = (X \cdot A_1, ..., X \cdot A_n),$$

因此,计算 X 和 A 的乘积问题即转化为计算 X 与 A 的列向量的内积问题.利用这一原理,构造向量与矩阵乘积保密计算协议如下:

协议 5 向量和矩阵乘积保密计算协议.

输入: Alice 和 Bob 分别输入有理数向量 $X = (x_1, ..., x_m)$ 和有理数矩阵 $A = (a_{ij})_{m \times n}$.

输出: Alice 输出 $Y = XA$.

1. Alice 将向量 $X = (x_1, ..., x_m)$ 按以下方式进行分解:随机选取有理数 $a_1, ..., a_t$ 和有理数向量 $X_i = (x_{i1}, ..., x_{im})$ $(i \in [1,t], 2 \le t < n)$, 使得 $X = a_1 X_1 + ... + a_t X_t$. Alice 将向量 $(X_1, ..., X_t)$ 发送给 Bob.

2. Bob 计算 $z_1 = X_1 A, ..., z_t = X_t A$, 将 $z_1, ..., z_t$ 发送给 Alice.

3. Alice 计算 $Z = a_1 z_1 + ... + a_t z_t$.

4. Alice 输出 Z.

协议 5 的正确性 我们只需要证明 $Z = XA$ 成立即可.

根据矩阵的运算性质,对于任意向量 $X = (x_1, ..., x_m)$, 以及关于 X 的任意分解方式 $X = a_1 X_1 + ... + a_t X_t$, 均有下面等式成立:

$$Z = a_1 z_1 + ... + a_t z_t = a_1 X_1 A + ... + a_t X_t A = (a_1 X_1 + ... + a_t X_t) A = XA.$$

因此,协议 5 是正确的.

协议 5 的安全性 协议 5 的设计思想本质上是协议 1 设计思想的推广,向量 X 的安全性与协议 1 类似,矩阵 A 的各列向量的安全性与协议 1 中 Bob 向量 Y_1 (或 Y_2)的安全性类似.由于矩阵 A 由 n 个列向量联合构成,根据协议设计可知 A 的各列向量之间没有任何关系,这也保证了矩阵的各元素数据具有较高的安全性.

6.3 点与多边形的位置关系

问题描述 假设 Alice 有一个私密点 $P_0(x_0, y_0)$, Bob 有一个私密凸多边形 S, 其顶点按逆时针顺序排列为

$P_1,...,P_m(m\geq 3)$, 顶点坐标为: $P_i(x_i,y_i)(i\in[1,m])$. Alice 和 Bob 想要保密判定点 $P_0(x_0,y_0)$ 是否在多边形 S 的内部.

计算基本原理 根据点与直线的位置关系可知,两个点 Q_1,Q_2 在一条直线 f 的同一侧的充要条件是:

$$f(Q_1)f(Q_2)>0.$$

那么对于点与多边形的位置关系,我们可以知道点 P_0 位于多边形 S 内部的充要条件是:选取多边形内任意一点 Q_i (包括多边形端点 $P_j,j\notin\{i,i+1\}$),将点 P_0 和点 Q_i 代入多边形 S 的每一条边 $P_iP_{i+1}(i\in[1,m])$ 所在的直线 f_i 都满足 $f_i(P_0)f_i(Q_i)>0$. 即当点 P_0 和点 Q_i 在多边形的每条边 P_iP_{i+1} 的同一侧时,点 P_0 位于多边形内部.如图 1 所示:当点 P_0 和 Q_1 在直线 $f_1,f_2,...,f_7$ 的同一侧时,点 P_0 在多边形内部.当然对于每一条直线 f_i, Bob 可以选择不同的点 Q_i 进行计算,在此约定 Bob 对直线 f_i 取点 $Q_i=P_{i+2}$.

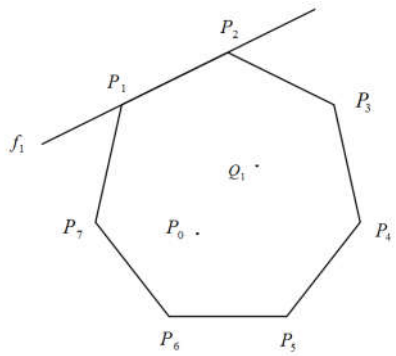

Fig.1 Position relationship between point and polygon

图 1 点与多边形的位置关系

Bob 首先写出多边形的边 $P_iP_{i+1}(i\in[1,m],P_{m+1}=P_1)$ 所在直线方程:

$$f_i(x,y)=a_ix+b_iy+c_i=0,$$

对于每个 $i\in[1,m]$, 为使上面表达式中的函数 $f_i(x,y)$ 完全确定,如果多边形的边 P_iP_{i+1} 不平行于 x 轴,则约定取 $a_i=1$, 否则,取 $b_i=1$. 根据计算基本原理可知,点 P_0 位于多边形内部的充要条件是下式成立:

$$f_1(x_0,y_0)f_1(P_3)>0,f_2(x_0,y_0)f_2(P_4)>0,...,f_m(x_0,y_0)f_m(P_2)>0.$$

Bob 进一步计算 $s_i=f_i(P_{i+2})$, 则上式等价于 $s_if_i(x_0,y_0)>0,i\in[1,m]$, 或

$$\begin{cases}(s_1a_1)x_0+(s_1b_1)y_0+s_1c_1>0,\\(s_2a_2)x_0+(s_2b_2)y_0+s_2c_2>0,\\\quad\cdots\cdots\\(s_ma_m)x_0+(s_mb_m)y_0+s_mc_m>0.\end{cases}\tag{14}$$

对于 $i\in[1,m]$, 记 $u_i=s_ia_i,v_i=s_ib_i,w_i=s_ic_i$, 并记 $X=(x_0,y_0,1),\ A_i=(u_i,v_i,w_i)$ 以及

$$A=(A_1,...,A_m)=\begin{pmatrix}u_1&u_2&...&u_m\\v_1&v_2&...&v_m\\w_1&w_2&...&w_m\end{pmatrix}\tag{15}$$

则条件(14)可表示为 $Y=XA=(X\cdot A_1,...,X\cdot A_n)>0$, 如此,可以通过计算 $Y=XA$,根据其每个分量是否大于 0 来判断点 P_0 是否在多边形 S 内部.

为下面叙述方便,定义二元谓词 $P(P_0,S)$ 如下:如果点 P_0 在多边形 S 内部,定义 $P(P_0,S)=1$; 否则, $P(P_0,S)=0$. 根据上面的计算原理容易设计出点与多边形位置关系的保密判定协议,具体如下:

协议 6 点与凸多边形的位置关系判定协议.

输入: Alice 输入点 $P_0(x_0,y_0)$, Bob 输入凸多边形 $S:P_1P_2...P_m$, 顶点坐标为 $P_i:(x_i,y_i)(i\in[1,m])$.

输出: Bob 输出 $P(P_0,S)$.

1. Bob 按照(15)式计算矩阵 A.

2. Alice 将向量 $X=(x_0,y_0,1)$ 按以下方式进行分解:随机选取有理数 t_1,t_2,t_3 和有理数向量 $X_i=(x_{i1},x_{i2},x_{i3})$ $(i=1,2,3)$, 使得 $X=t_1X_1+t_2X_2+t_3X_3$. 其中 $t=t_1+t_2+t_3>0$ 并将 (X_1,X_2,X_3) 发送给 Bob.

3. Bob 随机选取有理数 $r_i>0, i\in[1,m]$ 和有理数 s, 计算

$$\overline{A}=(r_1A_1,...,r_mA_m)$$

和

$$Z_1=X_1\overline{A}+sE, Z_2=X_2\overline{A}+sE, Z_3=X_3\overline{A}+sE,$$

其中 $E=(1,...,1)$ 并将 (Z_1,Z_2,Z_3) 发送给 Alice.

4. Alice 计算 $Z=\frac{1}{t}(t_1Z_1+t_2Z_2+t_3Z_3)$, 选取向量 Z 中的最小分量 z_{min} 发送给 Bob.

5. Bob 比较 z_{min} 与 s. 如果 $z_{min}>s$, Bob 输出 $P(P_0,S)=1$; 否则,输出 $P(P_0,S)=0$.

协议 6 的正确性 我们首先证明 $Z=\frac{1}{t}X\overline{A}+sE$ 成立.

根据矩阵运算性质,对于向量 $X=(x_0,y_0,1)$, 及 X 的任意分解方式 $X=t_1X_1+t_2X_2+t_3X_3$, 均有下面等式成立:

$$\begin{aligned}Z&=\frac{1}{t}(t_1Z_1+t_2Z_2+t_3Z_3)\\&=\frac{1}{t}(t_1(X_1\overline{A}+sE)+t_2(X_2\overline{A}+sE)+t_3(X_3\overline{A}+sE))\\&=\frac{1}{t}X\overline{A}+sE.\end{aligned}$$

进一步,根据有理数 $r_i>0, i\in[1,m]$ 和 $t>0$ 可知,

$$\begin{aligned}&Y=XA=(X\cdot A_1,...,X\cdot A_m)>0\\&\Leftrightarrow X\overline{A}=(r_1XA_1,...,r_mXA_m)>0\\&\Leftrightarrow z_{min}>s.\end{aligned}$$

因此,协议 6 是正确的.

协议 6 的安全性 协议 6 是结合协议 1 和协议 3 的设计思想进行设计的,类似于协议 1 和协议 3 的安全性证明思想进行证明.首先考虑 Alice 向量的安全性.Alice 秘密选择有理数 t_1,t_2,t_3 和有理数向量 X_1,X_2,X_3, 将向量 $X=(x_0,y_0,1)$ 进行分解.Bob 在协议执行中得到 Alice 发送的 X_1,X_2,X_3, Bob 由关系式 $X=t_1X_1+t_2X_2+t_3X_3$, 即

$$\begin{cases}x_0=t_1x_{11}+t_2x_{21}+t_3x_{31},\\y_0=t_1x_{12}+t_2x_{22}+t_3x_{32},\\1=t_1x_{13}+t_2x_{23}+t_3x_{33}.\end{cases}\tag{16}$$

推算 X 的相关信息.方程组(16)中有三个方程,但含有 5 个未知数 x_0,y_0,t_1,t_2,t_3, 由于 t_1,t_2,t_3 均由 Alice 随机选择, Bob 即使有无限的计算能力,也得不到向量 X 的任何信息.在第 4 步中,Bob 还收到 Alice 发送的向量 Z 中最小分量 z_{min}. 如果 Bob 猜测对于某一个 $i_0\in[1,m], z_{min}=z_{i_0}$. 这时 Bob 可联立求解(16)以及 $z_{min}=(r_{i_0}/t)X\cdot A_{i_0}+s$, 得到关于 x_0,y_0 的一个关系式,由于 Bob 无法获知对于哪个分量有 $z_{min}=z_i$, Bob 能猜对 x_0,y_0 关系式的概率仅有 $1/m$.

下面再考虑 Bob 向量的安全性.在协议执行中,Alice 仅获得 Bob 发送的信息:

$$\begin{cases}Z_1=X_1\overline{A}+sE=(r_1X_1\cdot A_1+s,...,r_mX_1\cdot A_m+s),\\Z_2=X_2\overline{A}+sE=(r_1X_2\cdot A_1+s,...,r_mX_2\cdot A_m+s),\\Z_3=X_3\overline{A}+sE=(r_1X_3\cdot A_1+s,...,r_mX_3\cdot A_m+s).\end{cases}\tag{17}$$

若记 $Z_j=(z_{j1},...,z_{jm})(j=1,2,3)$, 则(17)等价于下面方程组:

$$\begin{cases}z_{1i}=r_iX_1\cdot A_i+s,\\z_{2i}=r_iX_2\cdot A_i+s,\\z_{3i}=r_iX_3\cdot A_i+s,\\(i=1,...,m).\end{cases}\tag{18}$$

对 Alice 来说,方程组(18)有 $3m$ 个方程,却具有 $4m+1$ 个独立的未知数 $u_i,v_i,w_i,r_i(i=1,...,m)$ 以及 s, 且由于所有未知数均可取任意有理数,因此即使 Alice 有无限的计算能力也不可能通过求解(18)而获得 Bob 的矩阵 A.但对于每一个 $i\in[1,m]$. Alice 由(18)可以得到下面关系式:

$$k_i=\frac{z_{1i}-z_{2i}}{z_{1i}-z_{3i}}=\frac{(X_1-X_2)\cdot A_i}{(X_1-X_3)\cdot A_i},$$

从而得到向量 A_i 所有分量的一个线性关系式.

根据上面的分析可知,在协议执行过程中 Bob 仅可能以 $1/m$ 的概率猜对 x_0,y_0 的一个关系式,而对于每一个 $i\in[1,m]$, Alice 仅能获得 A_i 所有分量 u_i,v_i,w_i 间的一个关系式,进一步根据(15)可知,Bob 的多边形各顶点坐标是完全安全的.因此,协议 6 是安全的.

协议 6 的效率分析 为了分析协议 6 的实际效率,需要详细分析在协议执行过程中的计算复杂性和通信复杂性.由于本文协议 6 仅应用了基本算术运算,所以分析计算复杂性时仅考虑基本算术运算次数,以通信轮数和通信数据量(传输的数据个数)衡量协议的通信复杂性,同时利用 Java 编程语言对本协议进行实验测试.

计算复杂性 在协议 6 中,Bob 计算矩阵 A 执行了 $3m$ 次乘法运算,接收到 Alice 的分解向量后执行了 $10m$ 次乘法运算和 $9m$ 次加法运算.Alice 执行了 $3m+9$ 次乘法运算和 6 次加法运算.因此协议 6 共需要执行 $25m+15$ 次基本算术运算.

通信复杂性 在协议 6 中共需要 2 轮通信,在协议过程中的通信数据量为 $4m+10$.

实验测试 实验设定保密数据的范围为[-100, 100].下面分别对本文协议 6 在不同维向量下进行多次实验,实验结果随机抽取 50 组数据求取平均值.结果如表 4 所示.

Table 4 Analysis of simulation result of Protocol 6

表 4 协议 6 实验结果分析

协议	计算复杂性(B)	通信数据量	50 组实验数据平均值(ms)
本文协议 6	$25m+15$ (B)	$4m+10$	112.4925

在表 4 中,计算复杂性一栏,B 表示基本算术运算次数.由表 4 可知,协议 6 的计算复杂性和通信复杂性较低.

7 结论

向量问题的保密计算是安全多方计算的一项基本内容,实际生活中遇到的许多问题均可转化为向量问题得到解决,其中保密数据挖掘、保密统计分析、保密计算几何等问题与向量问题密不可分.特别地,目前关于有理数域上向量问题的保密计算研究较少且不成熟.

本文利用代数学基本知识,设计了简单高效的向量点积保密计算协议,向量相等保密判定协议及向量优势保密判定协议.严格的理论分析和实验结果表明我们的协议是安全高效的.应用这些协议作为基本模块不仅可以解决更多的安全多方计算问题,而且计算效率更高.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的安全多方向量计算问题.

References:

[1] Yao AC. Protocols for secure computations. In: Proc. of the 23th Annual Symposium on Foundations of Computer Science, IEEE, 1982. 160-164. DOI: 10.1109/SFCS.1982.38

[2] Ben-Or M, Goldwasser S, Wigderson A. Completeness theorems for non-cryptographic fault-tolerant distributed computation. In: Proc. of the 20th Annual ACM Symposium on Theory of Computing. New York: ACM Press, 1988. 1-10.DOI：10.1145/62212.62213

[3] Cramer R. Introduction to Secure Computation. Berlin: Springer-Verlag, 1999. 16-62. DOI：10.1007/3-540-48969-X_2

[4] Goldreich O, Micali S, Wigderson A. How to play any mental game. In: Proc. of the 19th Annual ACM Conference on Theory of Computing. New York: ACM Press, 1987. 218-229. DOI：10.1145/28395.28420

[5] Goldreich O. The Fundamental of Cryptography: Volume 2, Basic Applications. London: Cambridge University Press, 2004. 599-729. DOI：10.1017/CBO9780511721656

[6] Tang CM, Shi GH, Yao ZA. Secure multi-party computation protocol for sequencing problem. Science China Information Sciences, 2011, 54(8): 1654-1662. DOI：10.1007/s11432-011-4272-1

[7] Toft T. Sub-linear, secure comparison with two non-colluding parties. In: Proc. of International Conference on Practice and Theory in Public Key Cryptography Conference on Public Key Cryptography. Berlin: Springer-Verlag, 2011. 174-191. DOI：10.1007/978-3-642-19379-8_11

[8] Yi X, Rao FY, Bertino E, Bouguettaya A. Privacy-preserving association rule mining in cloud computing. In: Proc. of the 10th ACM Symposium on Information, Computer and Communications Security. Singapore, 2015: 439-450. DOI：10.1145/2714576.2714603

[9] Li YP, Chen MH, Li QW, Zhang W. Enabling multilevel trust in privacy preserving data mining. IEEE Transactions on Knowledge and Data Engineering, 2012, 24(9): 1598-1612. DOI：10.1109/TKDE.2011.124

[10] Kantardzic, Mehmed. Data mining: concepts, models, methods, and algorithms. Hoboken, USA: John Wiley & Sons, 2011.1-25. DOI：10.1002/9781118029145.ch1

[11] Du WL, Atallah MJ. Secure multi-party computation problems and their applications: A review and open problems. In: Proc. of the 2001 Workshop on New Security Paradigms. ACM, 2001. 13–22. DOI：10.1145/508171.508174

[12] Li SD, Wu CY, Wang DS, Dai YQ. Secure multiparty computation of solid geometric problems and their applications. Information Sciences, 2014, 282:401-413. DOI：10.1016/j.ins.2014.04.004

[13] Li SD, Wang DS, Dai YQ. Efficient secure multiparty computational geometry. Chinese Journal of Electronics, 2010,19(2):324–328.

[14] Liu YJ, Luo X, Joneja A, Ma CX, Fu XL, Song DW. User-adaptive sketch-based 3D CAD model retrieval. IEEE Transactions on Automation Science and Engineering, 2013, 10(3): 783-795. DOI: 10.1109/TASE.2012.2228481

[15] Fong PK, Weber-Jahnke JH. Privacy preserving decision tree learning using unrealized data sets. IEEE Transactions on Knowledge and Data Engineering, 2012, 24(2): 353-364. DOI: 10.1109/TKDE.2010.226

[16] Liu LG, Sun H, Jia HL, Zhang Y. CGIM: classificatory group index method for efficient ranked search of encrypted cloud data. Chinese Journal of Electronics, 2019, 47(2): 331-336. DOI: 10.3969/j.issn.0372-2112.2019.02.011

[17] Huang H, Li XY, Sun Y, Huang LS. PPS: Privacy-preserving strategyproof social-efficient spectrum auction mechanisms[J]. IEEE Transactions on Parallel and Distributed Systems, 2015, 26(5): 1393-1404. DOI：10.1109/TPDS.2014.2315200

[18] Li MJ, Juan JST, Tsai JHC. Practical electronic auction scheme with strong anonymity and bidding privacy. Information Sciences, 2011, 181(12): 2576-2586. DOI：10.1016/j.ins.2011.02.005

[19] Yin X, Tian YL, Wang HL. Delegation auction scheme for big data pricing. Chinese Journal of Electronics, 2018, 46(5): 1113-1120. DOI: 10.3969/j.issn.0372-2112.2018.05.014

[20] Chrétien S, Zhen WOH. Incoherent submatrix selection via approximate independence sets in scalar product graphs. In: Nicosia G, Pardalos P, Umeton R, Giuffrida G, Sciacca V, eds. Machine Learning, Optimization, and Data Science. Springer Cham, 2019. 95-105. https://doi.org/10.1007/978-3-030-37599-7_9

[21] Hofmann J, Fey D, Riedmann M, Eitzinger J, Hager G, Wellein G. Performance analysis of the Kahan-enhanced scalar product on current multi-core and many-core processors. Concurrency and Computation: Practice and Experience, 2017, 29(9). DOI：10.1002/cpe.3921

[22] Lin WP, Wang K, Zhang ZL, Chen H. Revisiting security risks of asymmetric scalar product preserving encryption and its variants. In: Proc. of the 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS), Atlanta USA: IEEE, 2017. 1116-1125. DOI：10.1109/ICDCS.2017.20

[23] You YP, Li XH. Ordering scalar products with applications in financial engineering and actuarial science. Journal of Applied Probability, 2016, 53(1): 47-56. https://doi.org/10.1017/jpr.2015.7

[24] Zhou SF, Dou JW, Guo YM, Mao Q, Li SD. Secure multiparty vector computation. Chinese Journal of Computers, 2017, 40(5): 1134-1150. DOI：10.11897/SP.J.1016.2017.01134

[25] Siabi B, Berenjkoub M, Susilo W. Optimally efficient secure scalar product with applications in Cloud computing. IEEE Access. 2019, 7: 42798-42815. DOI：10.1109/ACCESS.2019.2908230

[26] Rong H, Wang HM, Huang K, Liu J, Xian M. Privacy-preserving scalar product computation in Cloud environments under multiple keys. In: Yin H, ed. Intelligent Data Engineering and Automated Learning. Springer, Cham, 2016: 248-258. https://doi.org/10.1007/978-3-319-46257-8_27

[27] Du WL, Zhan ZJ. A practical approach to solve secure multi-party computation problems. In: Proc. of the 2002 Workshop on New Security Paradigms, Virginia Beach, 2002. 127-135. https://doi.org/10.1145/844102.844125

[28] Shaneck M, Kim Y. Efficient cryptographic primitives for private data mining. In: Proc. of the 43rd Hawaii International Conference on System Sciences. Hawaii: IEEE, 2010. 1-9. DOI：10.1109/HICSS.2010.172

[29] Zhu Y, Takagi T. Efficient scalar product protocol and its privacy-preserving application. International Journal of Electronic Security Digital Forensics, 2015, 7(1): 1-19. DOI：10.1504/IJESDF.2015.067985

[30] Goethals B, Laur S, Lipmaa H, Mielikäinen T. On private scalar product computation for privacy-preserving data mining. In: Proc. of International Conference on Information Security and Cryptology, Seoul, Korea, 2004. 104-120. DOI: 10.1007/11496618_9

[31] Yang B, Yang CH, Yu Y, Xie D. A secure scalar product protocol and its applications to computational geometry. Journal of Computers, 2013, 8(8): 2018-2026. DOI：10.4304/jcp.8.8.2018-2026

[32] Dong CY, Chen LQ. A fast secure dot product protocol with application to privacy preserving association rule mining. In: Proc. of the Advances in Knowledge Discovery and Data Mining. Berlin: Springer International Publishing, 2014. 606-617. DOI: 10.1016/0022-4804(81)90076-7

[33] Sheng G, Wen T, Guo Q, Yin Y. Privacy preserving inner product of vectors in cloud computing. International Journal of Distributed Sensor Networks, 2014, 2014(1): 1-6. https://doi.org/10.1155/2014/537252

[34] Liu F, Ng WK, Zhang W. Secure scalar product for big-data in mapreduce. In: Proc. of the 2015 IEEE First International Conference on Big Data Computing Service and Applications, San Francisco, CA, 2015. 120-129. https://doi.org/10.1109/BigDataService.2015.9

[35] Li SD, Yang XL, Zuo XJ, Zhou SH, Kang J, Liu X. Privacy-Preserving Graphical Similarity Determination. Chinese Journal of Electronics, 2017,45(9):2184-2189. DOI: 10.3969/j.issn.0372-2112.2017.09.019

[36] Atallah MJ, Du WL. Secure Multi-party computational geometry. In: Proc. of the 7th International Workshop on Algorithms and Data Structures. Berlin: Springer Berlin Heidelberg, 2001. 165-179. https://doi.org/10.1007/3-540-44634-6_16

[37] Liu W, Luo SS, Wang YB. Secure two-party vector dominance statistic protocol and its applications. Chinese Journal of Electronics, 2010, 38(11): 2573-2577

[38] Li SD, Zuo XJ, Yang XL, Gong LM. Secure vector dominance protocol and its applications. Chinese Journal of Electronics, 2017, 45(5): 1117-1123. DOI:10.3969/j.issn.0372-2112.2017.05.014

[39] Damgard I, Jurik M. A length-flexible threshold cryptosystem with applications. In: Proc. of the 2003 Australasian Conference on Information Security and Privacy. Berlin: Springer Berlin Heidelberg, 2003. 350-364. https://doi.org/10.1007/3-540-45067-X_30

附中文参考文献:

[16] 刘良桂,孙辉,贾会玲,张宇,面向高效加密云数据排序搜索的类别分组索引方法.电子学报,2019,47(2):331-336

[19] 尹鑫,田有亮,王海龙.面向大数据定价的委托拍卖方案.电子学报,2018,46(5):1113-1120

[24] 周素芳,窦家维,郭奕旻,毛庆,李顺东.安全多方向量计算.计算机学报,2017,40(5):1134-1150

[35] 李顺东,杨晓莉,左祥建,周素芳,亢佳,刘新.保护私有信息的图形相似判定.电子学报, 2017,45(9):2184-2189

[37] 刘文,罗守山,王永滨.安全两方向量优势统计协议及其应用.电子学报,2010,38(11):2573-2577

[38] 李顺东,左祥建,杨晓莉,巩林明.安全向量优势协议及其应用.电子学报,2017,45(5):1117-1123