安全多方多数据排序

李顺东 杜润萌 杨颜璟 魏 琼
(陕西师范大学计算机科学学院 西安 710062)

摘 要 安全多方计算是近年来国际密码学界研究的热点.保密比较两个数据的大小是安全多方计算研究的基本问题之一,可以用它构造其他安全多方计算问题的解决方案.多方保密排序问题是两个数据比较大小的自然推广,也是安全多方计算的一个基本问题,在电子拍卖、保密竞拍、匿名投票以及安全数据挖掘等方面有广泛的应用,因此研究多方保密排序有重要的理论与实际意义.本文主要研究多方参与者数组以及联合数组中出现重复元素的排序问题,而目前已有的保密排序方案无法很好地解决这些问题.基于此,本文以新编码方法为基础,结合门限解密椭圆曲线密码系统,在半诚实模型下设计了同序位归一排序、并列同序位多重排序、增序位全排序的保密排序协议,并证明方案在半诚实模型下是安全的.设计的新编码方法不但能够用于解决本文的保密排序问题,而且可以作为解决其他很多安全多方计算问题最重要的工具.例如基于排序协议,我们可以解决最大值和最小值问题;对本文提出的协议 1 稍作修改,可以保密计算集合的并集以及并集的势,这些也是科学计算中的基本问题.本文最后分析了协议的复杂性并进行了实验测试,理论分析和实验结果都表明本文协议是高效和实用的协议.在本文的最后我们给出了恶意模型下的排序协议.

关键词 密码学;安全多方计算;排序;门限解密;模拟范例
中图法分类号 TP309 **DOI 号** 10.11897/SP.J.1016.2020.01448

Secure Multiparty Multi-Data Ranking

LI Shun-Dong DU Run-Meng YANG Yan-Jing WEI Qiong
(*School of Computer Science, Shaanxi Normal University, Xi'an* 710062)

Abstract Secure multiparty computation (SMC) has become research focus in the international cryptographic community in recent years, and a key technology of the information security in cooperative computation, cloud computing, electronic commerce, electronic voting etc. Cryptographic scholars have studied many SMC problems such as secure scientific computation, secure data mining, secure computational geometry, secure statistical analysis, and propose solutions to these problems, but there are more problems need to be studied and solved. Secure multiparty scientific computation is an important branch of SMC, which is of theoretical importance, and has many applications in the area of modern mathematics and real life. Comparing the size of two data is one of the most important problems of SMC, that is the millionaires' problem. Its solutions can construct many SMC protocols. Secure ranking is the natural generalization of the millionaires' problem and is also a basic problem in SMC. Secure ranking is widely applied to secure electronic auctions, secure bidding, anonymous voting, privacy-preserving data mining and so on. Studying secure ranking which need to know the range or the distribution of private data has important theoretical and practical significance. For example, age, salary, achievement, and parameter of product are in a relative small range in the daily life, so it is of great practical value to find the ranking of data privately in a range. This paper mainly studies the ranking problem when there

收稿日期:2019-04-02;在线发布日期:2019-09-14.本课题得到国家自然科学基金(61272435)资助.**李顺东**,博士,教授,博士生导师,主要从事密码学与信息安全研究.E-mail:shundong@snnu.edu.cn.**杜润萌**,硕士研究生,主要研究方向为密码学与信息安全研究.**杨颜璟**,硕士研究生,主要研究方向为密码学与信息安全研究.**魏 琼**,硕士研究生,主要研究方向为密码学与信息安全研究.

are repeated elements in multi-party arrays and union arrays, which cannot be solved by using the existing solutions. To solve this problem, this paper designs new encoding schemes to hide private numbers. Based on these new encoding schemes and the threshold decryption elliptic cryptosystem, we design secure ranking protocols for the following three ranking problems: the ranking in which the same numbers have the same order, and the order of the next number increases by 1; the ranking in which the same numbers have the same order, but if there are k same numbers, then the order of next number will increase by k; and the ranking in which the same numbers have different orders. Using the well-accepted simulation paradigm, we show that these protocols are secure in the semi-honest model. These new encoding schemes, in addition to being used to solve the problems of this paper, can independently be used as important tools to address other SMC problems. For example, based on our ranking protocol, we can solve the maximum and minimum problems. This problem extends the famous millionaires' problem. For example, if we want to privately compute the average of some private data, it is necessary to remove the outliers of a set of data, that is, the maximum and the minimum, to ensure the stability and accuracy of average. Our protocol can be used to do this. And Protocol 1, with a slight modification, can be used to privately compute set unions and the cardinality of set unions. These are also basic problems in secure scientific computations. We finally analyze the computational complexity and communication complexity of our protocols and test their performance. Both theoretical analysis and tests show that these protocols are efficient and practical. At the end of this paper, we design a ranking protocol for the malicious model.

Keywords　cryptography; secure multiparty computation; ranking; threshold decryption; simulation paradigm

1　引　言

安全多方计算(Secure Multiparty Computation, SMC)是指分别拥有隐私数据 $x_1,\cdots,x_n$ 的 n 个参与者联合计算函数 $f(x_1,\cdots,x_n)$ 的值,计算完成之后,每个参与者除了得到 $f(x_1,\cdots,x_n)$ 的值之外,得不到任何关于隐私数据的信息.虽然任何函数都能通过将其表示成电路形式而进行安全计算[1-2],但一般来说应用电路方法对一个复杂的函数很难设计出有效的协议.因此相关学者对于一些具体的安全多方计算提出相应的解决方案.

排序是各种信息系统中最重要的基本操作,对于一个数据序列 A,排序是指将 A 中所有数据按照从小到大的顺序排成一个序列,进而确定 A 中任意一个数据 x 在该序列中的位置.公开数据的排序已经研究得非常充分,这里不作过多的讨论.

一方面,在实际应用中,很多应用场合参与排序的数据是隐私数据,直接用上述排序算法进行排序将导致隐私泄露,带来严重的安全问题.因此当涉及到对隐私数据进行排序的时候,就需要设计相应的安全多方计算协议实现对数据的保密排序;另一方面,保密排序协议在各种数据库操作、合作入侵检测系统[3]、不经意 RAM[4] 以及私密的集合交集计算[5]等方面都有广泛应用.

本文所研究的多方保密排序问题是两方秘密数据比较问题的发展.Yao 在文献[6]中首次提出了保密比较两个数据大小的问题,即著名的百万富翁问题.对于安全排序问题,直观的想法是多次调用两方比较协议对数据进行排序,但这样会泄露很多额外的信息.因此两方协议一般都无法直接推广到多方的情形[7].

近年来,人们设计了一些保密排序协议,主要有 Ajtai 的排序网络[8]、Batcher 的归并排序[9]、随机化希尔排序[10]、不经意关键字排序[11]、快速排序[12]等.上述排序主要是对各数组中的元素进行排序,计算结束后输出一个有序且稳定的数据序列,只能保护参与者是否持有某个数据这样的隐私信息.文献[13]通过设计一个严格单调的变换使每个参与者可以得到所有各方数据的排序结果,但具体的输入数

据 x_i 是保密的.方案中所应用的单调变换思想以及秘密分享方法无法推广到每个参与方持有多个数据的情形.文献[14]先提出一种排序算法,并利用 Shamir 秘密共享实现保密排序.文献[15]借助云服务器和同态加密算法实现了保密排序.文献[14-15]提出的方案都是通过两两比较的方式实现排序,但多次执行协议会泄露哪些数据相等以及数据大小等额外信息,而且效率低下.文献[16]借助第三方把排序问题简化为多方求和问题,并结合数据压缩及编码方法实现了安全数据排序.但文献[16]提出的方案只适用于每个参与者只拥有一个数据,另外,当数据 $x_i=x_r$ 时,可能导出 $x_i\neq x_r$ 的错误.文献[17]除了与文献[16]存在同样的问题之外,还会造成一定的排名损失.即当 $x_i>x_r$($x_i<x_r$),会出现 $x_i<x_r$($x_i>x_r$)的情况.

文献[18]主要研究字符排序问题,最后作者将保密字符排序协议作为基础协议用来解决数据排序问题.文献[19-20]应用同态加密方案设计了安全多方多数据保密排序方案,文献[18-20]所研究问题与本文研究问题最接近,我们将在协议效率分析部分将本文结论与文献[18-20]进行详细的比较.

在实际中经常遇到一些私密性要求较高的保密排序问题.在这些问题中,多个参与者各自拥有一个私密的数据序列,各方希望进行合作保密计算,计算结束后,每个参与者仅可获得自己所拥有数据在所有参与者联合数据中的正确排序,而对于其他参与者的具体数据以及排序信息全部保密.这些问题在实际中具有很强的应用背景,下面举例说明.为方便起见,下文中将各参与者的数据序列表示为数组,将所有参与者的联合数据以联合数组的形式表达(对于数组 $A_1=(x_1,y_1)$,$A_2=(x_2,y_2)$,将数组 $A=(x_1,y_1,x_2,y_2)$ 称为 A_1,A_2 的联合数组,并简记为 $A=(A_1,A_2)$,以此类推).考虑下面几个问题:

(1)有若干个企业同时开发销售某类商品,他们希望保密协商该商品的大致售价,这样既能保证自己的商业秘密,又可避免恶性竞争.为此,各企业 P_i,$i\in[1,n]$ 对该产品设定一组最低定价 x_i,计划售价 y_i,最高定价 z_i,记 $A_i=(x_i,y_i,z_i)$.每个企业都希望了解自己企业设定的几个价位与其他企业价位相比是偏高还是偏低,以便进行调整,但又不泄露彼此的拟定价格.这个问题即可转化为多方保密排序问题.

(2)学校在考试结束后希望对某年级各班学生的成绩数组 A_i 进行一次排名,进而得到 A_i 中的元素在联合数组 $A=(A_1,\cdots,A_n)$ 中的序位,以确定哪个班级教学水平更好一些,并要求各个班级学生成绩以及在全年级的排名次序只有本班任课老师知道.这个问题也可转化为保密排序问题.

问题(2)和问题(1)的区别是每个数组 A_i 和联合数组 A 都允许有重复元素出现.排序结果要求相同元素在联合数组中具有相同的排序序位,即若数据 x 在 A 中重复出现 k 次,x 的数据序位均相同,但要在整个序列中占 k 个位次.即假设 x 在 A 中的排序位置为 r,如果在 A 中大于 x 的相邻数据为 y,那么 y 的序位应为 $r+k$.

(3)进一步考虑下面问题:允许数组 A_i 及其联合数组 A 中有重复元素出现.但在排序时,A 中相同元素占有不同的位次,即若数据 x 在 A 中重复出现 k 次,序列中首个 x 的序位为 r,那么 k 个 x 应该占有 k 个不同序位,依次为 $r,r+1,\cdots,r+k-1$,如果在 A 中大于 x 的相邻数据为 y,这时 y 的序位为 $r+k$.

这种排序方式在实际中也具有广泛应用.比如在保密拍卖中,如果将 n 个投标者所投的秘密数据记为 $x_1,\cdots,x_n$,为了确定哪些人中标,需要对 $x_1,\cdots,x_n$ 进行保密排序,由于拍卖活动最终要确定参与者的中标优先顺序,所以对 $x_1,\cdots,x_n$ 中的相同数据也要根据事先约定的规则确定其先后顺序.对于某些行业员工工资的摸底调查也可应用类似方案进行.

上述问题可分别抽象为几种不同类型的保密排序问题,并允许参与者的数组 A_i 以及联合数组 A 中出现重复元素,而目前已有的保密排序方案无法很好解决这些问题.本文主要研究解决这些保密排序问题,并针对实际应用场景中不同的排序需求,提出三种类型的序位确定方式,分别描述如下:

① 并列同序位归一排序简称归一排序.数组中相同数据共同占有一个序位,下一个较大数据的序位仅增加一位.这种排序方法即是把多重数组中的相同元素作为一个元素对待,其结果等价于对数组中所有不同的数据进行简单排序.

② 并列同序位多重排序简称多重排序.如果数据 x 在联合数组中重复出现 k 次,要求这 k 个 x 保持同一序位,但下一个较大数据的排序序位却要相应增加 k 位.

③ 增序位全排序简称增序排序.如果数据 x 在联合数组中重复出现 k 次,要求这 k 个 x 排序序位依次递增共占有 k 个不同序位.因此,如果多重数组

共有 N 个元素，那么数组元素是从序位 1 开始依次递增，其中最大元素的序位为 N.

本文主要应用椭圆曲线门限密码体制，针对上述三类排序问题在半诚实模型下设计了安全高效的保密排序协议，同时给出了恶意模型下的保密排序协议. 主要贡献如下：

1) 提出了新的排序问题. 针对多重数组排序问题定义了三种排序方案：归一排序、多重排序以及增序排序，并分别设计了相应的保密排序协议. 协议能够保证数据排序序位的隐私性.

2) 提出了新的编码方法. 对于三类排序问题，我们分别设计了新的编码方法，使得参与者的数组 A_i 对应于一个特殊的编码数组；并对编码数组运用加密选择技巧，或进行适当的同态加密运算，使问题得以解决.

3) 本文所研究问题适用于多个参与方分别具有多个数据的情形，推广了保密排序的适用范围，扩大了保密排序的实际应用价值.

本文第 2 节首先介绍后面研究所需要的相关预备知识；第 3～5 节针对前面所提出的三种排序方案具体构造保密计算协议；第 6 节对于所设计协议的效率进行分析比较；第 7 节给出了恶意模型下的保密排序协议；第 8 节是对全文的总结.

2　预备知识

2.1　安全性定义

半诚实模型[21]. 参与者 $P_1,\cdots,P_n$ 将会按照协议规则忠诚地执行协议步骤，但 P_i 有可能会记录 $P_j(i\neq j)$ 传送的中间结果和协议的最终输出结果，并根据记录的秘密信息推导出 P_j 的秘密输入，我们认为这样的参与者为半诚实参与者，称这样的模型为半诚实模型.

模拟范例[22]. 模拟范例是目前安全多方计算研究中证明安全性时广泛接受的证明方法，本文证明协议安全性时主要采用模拟范例的方法.

一些记号. 假设参与者为 $P_1,\cdots,P_n$，分别拥有秘密有序数组 $T_1,\cdots,T_n$.

(1) 定义 $T=(T_1,\cdots,T_n)$，$f(T)=(f_1(T),\cdots,f_n(T))$ 是概率多项式时间函数，定义 π 是计算 f 的协议.

(2) 在执行协议 π 期间，$P_i(i=1,\cdots,n)$ 得到的消息序列记作

$$view_i^{\pi}(T)=(T_i,r_i,m_1^i,\cdots,m_t^i,f_i(T)),$$

其中 P_i 产生的随机数用 r_i 表示，P_i 收到的第 j 个消息用 m_j^i 表示，P_i 得到的输出结果用 $f_i(T)$ 表示.

(3) 对于部分参与者集合 $I=\{P_{i1},\cdots,P_{i\alpha}\}\subset\{P_1,\cdots,P_n\}$，记

$$view_I^{\pi}(T)=(I,view_{i1}^{\pi}(T),\cdots,view_{i\alpha}^{\pi}(T)).$$

定义 1.　对于概率多项式时间函数 f 以及保密计算 f 的协议 π，如果对于每一个 $I=\{P_{i1},\cdots,P_{i\alpha}\}\subseteq\{P_1,\cdots,P_n\}$，有概率多项式时间算法 S，使得下式成立，

$$\{S(I,(T_{i1},\cdots,T_{i\alpha}),f_I(T))\}_T\overset{c}{\equiv}\{view_I^{\pi}(T)\}_T \quad (1)$$

则 π 保密计算 f，符号 $\overset{c}{\equiv}$ 表示计算不可区分[23].

2.2　恶意模型

设计恶意模型下保密计算协议时，我们需要思考协议执行时 P_i 可能做出的所有的恶意行为：(1) 参与方在协议中用其他数据替代原本规定的输入、拒绝参加协议以及在协议执行期间随时中止协议，但这三种行为在任意一个协议中都无法避免，因此不予考虑；(2) 参与方可能在协议的实际执行中应用了并非均匀分布的随机带；(3) 参与方可能试图去发送不同于半诚实模型中所规定的信息. (2)和(3)都是恶意模型下安全的协议需要防止的恶意行为.

恶意模型下安全的协议会迫使参与者像半诚实参与者一样执行协议. 关于恶意模型下的安全多方计算协议以及安全性定义可参看文献[23].

2.3　椭圆曲线密码系统

应用椭圆曲线密码系统时，经常采用的椭圆曲线方程是：

$$y^2\equiv x^3+ax+b(\mathrm{mod}p)$$
$$(a,b\in GF(p),4a^3+27b^2(\mathrm{mod}p)\neq 0) \quad (2)$$

式(2)中所有系数都属于某一有限域 $GF(p)$(p 是一个大素数).

基于椭圆曲线实现 ElGamal 密码体制[24]描述如下：构造(2)表达的椭圆曲线方程时，我们先选择一个大素数 p 以及两个整数 a,b. 设 $E_p(a,b)$ 表示椭圆曲线上的点集 $\{(x,y)\mid 0\leq x<p,0\leq y<p,$ 且 x,y 均为整数$\}$ 与无穷远点 O 的并集. 取 $E_p(a,b)$ 的一个基点 $G(x_1,y_1)$，在这里要求 G 的阶是大素数(满足 $LG=O$ 的最小正整数 L 为 G 的阶). 公开 $E_p(a,b)$ 和 G.

密钥生成. 参与者任意选择一个小于 L 的整数 k. 因此，该密码系统的私钥是 k，公钥是 $K=kG$.

加密. 参与者将明文消息 m 编码到 $E_p(a,b)$ 上一点 M，并随机选择正整数 r，计算密文：

$$(c_1,c_2)=(M+rK,rG).$$

解密. 参与者进行解密操作，继而得到明文消息 m 的编码点：

$$M=c_1-kc_2.$$

对编码点 M 进行解码即可得到明文 m.

编码方法与同态性质. 首先，直接验证可知椭圆曲线具有下面性质：

$$\begin{aligned}E(M_1)+E(M_2)&=(M_1+r_1K,r_1G)+(M_2+r_2K,r_2G)\\&=(M_1+M_2+(r_1+r_2)K,(r_1+r_2)G)\\&=E(M_1+M_2)\end{aligned}\tag{3}$$

式(3)仅表明椭圆曲线密码系统对于编码点具有加法同态性，但不能说明对明文 m 具有加法同态性.

我们利用下面方法对消息 m 进行编码[25]，即将 m 编码为 mG(这里 G 为椭圆曲线的一个基点). 这样的编码方法可使密码系统对编码点 mG 的加法同态性保持到明文消息 m 上. 这时对明文 m_1,m_2，分别将其编码后得到 $M_1=m_1G,M_2=m_2G$，根据式(3)，可得到 $E(m_1G)+E(m_2G)=E((m_1+m_2)G)$. 即不用解密，我们直接就可以从 m_1G 和 m_2G 的密文来计算 $(m_1+m_2)G$ 的密文 (c_1,c_2). 但是另一方面我们直接进行解密 (c_1,c_2)，就只能得到 $(m_1+m_2)G$ 的密文. 因此我们需要再解码才能得到 m_1+m_2. 因为椭圆曲线密码系统存在离散对数困难性的问题，所以我们无法直接由 $(m_1+m_2)G$ 解码获得 m_1+m_2，但假设存在一个确定的正整数 N，使得 $1\leq m_1,m_2\leq N$，即可通过比较集合 $\{mG,m\in[2,2N]\}$ 中哪个点与 $(m_1+m_2)G$ 相同来确定 m_1+m_2 的值.

文献[19]所设计的保密排序协议中，同样无法完全解密得到 m，也需要通过比较而获得 m.

2.4 密文重随机化

本文在应用密码系统构造协议时，需要对密文进行重随机化. 由于本文选择的密码系统是概率加密系统，与其他概率加密系统一样，椭圆曲线密码系统也有语义安全的性质，直观表述就是由公钥加密的密文计算不可区分.

如果需要对椭圆曲线密码系统的密文 (c_1,c_2) 进行重随机化，则可选取随机正整数 r_1，计算

$$\begin{aligned}(c_1',c_2')&=(c_1+r_1K,c_2+r_1G)\\&=(M+(r+r_1)K,(r+r_1)G).\end{aligned}$$

显然 (c_1',c_2') 仍然是 M 的密文，但与 (c_1,c_2) 比较，它已是一个全新的密文形式. 如果没有解密密钥，无法判断它们是否为同一明文对应的密文. 在下文中所应用的重随机化过程即由上述运算过程实现，并令 $R(C)$ 表示重随机化，其中 C 表示一个密文.

2.5 门限椭圆曲线密码系统

门限解密[26-27]的主要功能是在协议执行期间对抗参与者的合谋攻击. 本文使用的是能够抵抗最大合谋攻击者的 (n,n) 门限密码系统. 在该体系中，公钥由 n 个参与者联合生成，而密钥分布在这 n 个参与者之中. 任何参与者都可以直接用公钥加密明文消息，然而解密一个密文就需要 n 个参与者共同合作. 本文构造椭圆曲线 (n,n) 门限系统如下：

密钥生成. 选定一条 $E_p(a,b)$ 及其上的一个生成元 G. $E_p(a,b)$ 和 G 作为公开参数，每个参与者 $P_i,i\in[1,n]$ 任意选择一个小于 L 的整数 k_i 作为私钥，计算 $K_i=k_iG$，共同生成公钥：

$$K=\sum_{i=1}^{n}K_i=\sum_{i=1}^{n}k_iG.$$

加密. 首先将明文消息编码到 $E_p(a,b)$ 上一点 M，并选择一个随机数 r：$1\leq r\leq L-1$，计算密文：

$$(c_1,c_2)=(M+rK,rG).$$

合作解密. 对密文 (c_1,c_2)，每个参与者 P_i 计算并公布 k_ic_2，进一步计算下面公式得到编码点 M：

$$M=c_1-\sum_{i=1}^{n}k_ic_2.$$

3 并列同序位归一法排序

下文中将有重复元素出现的数组称为多重数组，无重复元素出现的数组称为标准数组. 如果没有特殊说明，下文中均默认数组为多重数组.

3.1 基本原理

问题描述. 假设有 n 个参与者 $P_1,\cdots,P_n$，对于每个 $i\in[1,n]$，P_i 拥有一个私密的有序数组 $T_i=(t_{i1},\cdots,t_{ie_i})$，其中数组 T_i 是标准数组. 根据归一排序的定义，只需考虑各自数组为标准数组即可. n 个参与者希望进行合作保密计算，计算结束后 P_i 仅可获知自己的每个元素 $t_{is},s\in[1,e_i]$ 在联合数组 $T=(T_1,\cdots,T_n)$ 中归一排序的序位 r_{is}.

计算原理. (1) 参与者 $P_1,\cdots,P_n$ 商定一个全集 $J=[1,N]$，满足 $T_i\subseteq J$. 在全集 J 之下，每个参与者 P_i 根据数组 T_i 构造一个 N 维向量：

$$\boldsymbol{Y}_i=(y_{i1},\cdots,y_{iN})\tag{4}$$

其中对每一个 $j\in J$，定义：

$$y_{ij}=\begin{cases}1, & j\in T_i\\0, & j\notin T_i\end{cases}\tag{5}$$

(2) 令 $\boldsymbol{Y}_1^*=\boldsymbol{Y}_1$，$P_i$ ($i=1,\cdots,n-1$) 依次将向量 $\boldsymbol{Y}_i^*$ 发给 P_{i+1}，P_{i+1} 根据 $\boldsymbol{Y}_i^*$ 和 Y_{i+1} 构造新向量：

$$\boldsymbol{Y}_{i+1}^*=(y_{(i+1)1}^*,\cdots,y_{(i+1)N}^*),$$

其中对每一个 $j\in J$，定义：

$$y_{(i+1)j}^*=\begin{cases}y_{(i+1)j}, & y_{(i+1)j}=1\\ y_{ij}^*, & y_{(i+1)j}=0\end{cases}\tag{6}$$

(3) P_i 按下面方式计算 r_{is}：

$$r_{is}=\sum_{j=1}^{t_{is}}y_{nj}^*\tag{7}$$

命题 1.　参与者 P_i 具有的数组 T_i 中元素 t_{is} 在联合数组 T 中的归一排序序位 r_{is} 可由式(7)计算获得.

证明.　根据归一排序的定义可知，同一元素不管重复多少次，在最后排序时它们只占一个序位，因此在考虑参与者各自的秘密数组或他们的联合数组时，只需删去重复元素，把所有数组简化为标准数组处理即可.

(1) 根据向量 $\boldsymbol{Y}_i$ 的构成方式，若数据 $x\in[1,N]$ 在某个 T_i 中出现，则 $y_{ix}=1$. 进一步根据 $\boldsymbol{Y}_n^*$ 的构造方式可知

$$x\in T\Leftrightarrow y_{ix}^*=1.$$

并且可知 x 在联合数组 T 中的排位即为 $\sum_{j=1}^{x}y_{nj}^*$.

(2) 对于任意 $i\in[1,n]$ 以及 $s\in[1,e_i]$，由于 $t_{is}\in T$，因此 $x=t_{is}$ 在 T 中的序位是 $r_{is}=\sum_{j=1}^{t_{is}}y_{nj}^*$. 这即为式(7). 命题得证.　　证毕.

由此可知，无论 $x\in T$ 属于哪个参与者的私密数组(也可能多人同时拥有 x)，元素 x 在 T 中的排序位次都是 $\sum_{j=1}^{x}y_{nj}^*$，即向量 $\boldsymbol{Y}_n^*$ 中前 x 个分量之和.

实例 1.　假设参与者 P_1,P_2,P_3 分别拥有数组 $T_1=(t_{11},t_{12})=(1,3)$，$T_2=(t_{21},t_{22})=(2,3)$ 以及 $T_3=(t_{31})=(6)$. 选取 $N=7$，3 个参与者按照方式(4)分别构造 7 维向量如下：

$$\boldsymbol{Y}_1=(1,0,1,0,0,0,0),$$
$$\boldsymbol{Y}_2=(0,1,1,0,0,0,0),$$
$$\boldsymbol{Y}_3=(0,0,0,0,0,1,0).$$

P_1 发送向量 $\boldsymbol{Y}_1^*=\boldsymbol{Y}_1$ 给 P_2，P_2 按照方式(6)得：

$$\boldsymbol{Y}_2^*=(1,1,1,0,0,0,0),$$

P_2 发送 $\boldsymbol{Y}_2^*$ 给 P_3. P_3 操作与 P_2 一致，得到：

$$\boldsymbol{Y}_3^*=(1,1,1,0,0,1,0).$$

根据 $\boldsymbol{Y}_3^*$，各参与者按照式(7)确定其各自数组的元素在联合数组中的排序位次，排序结果如表 1 所示.

表 1　并列同序位归一排序结果

元素	序位	算法
1	1	1
2	2	1+1
3	3	1+1+1
6	4	1+1+1+0+0+1

由表 1 可知，元素 $t_{11}=1$ 的序位是 1，其值等于 $\boldsymbol{Y}_3^*$ 的第一个元素 1；元素 $t_{21}=2$ 的序位是 2，其值等于 $\boldsymbol{Y}_3^*$ 的前两个元素求和得到 1+1=2；元素 $t_{12}=t_{22}=3$ 的序位是 3，其值等于 $\boldsymbol{Y}_3^*$ 的前 3 个元素求和得到 1+1+1=3；元素 $t_{31}=6$ 的序位是 4，其值等于 $\boldsymbol{Y}_3^*$ 的前六个元素求和得到 1+1+1+0+0+1=4.

3.2　归一排序协议

命题 1 是计算归一排序序位的基本原理，下面应用椭圆曲线门限密码体制，设计构造保密计算协议. 具体设计如下.

协议 1.　保密归一排序协议.

输入：$P_1,\cdots,P_n$ 各自的秘密有序数组 $T_1=(t_{11},\cdots,t_{1e_1}),\cdots,T_n=(t_{n1},\cdots,t_{ne_n})$.

输出：对于任意 $i\in[1,n]$，P_i 获得输出结果 $f_i(T_1,\cdots,T_n)$：元素 t_{is}，$s\in[1,e_i]$ 在联合数组 $T=(T_1,\cdots,T_n)$ 中按照归一排序方法获得的序位 r_{is}.

准备： n 个参与者 P_i，$i\in[1,n]$ 首先商定一条椭圆曲线 $E_p(a,b)$ 以及椭圆曲线上的一个基点 G. 参与者 P_i 任意选择小于 L 的整数 k_i 作为私钥，计算 $K_i=k_iG$，共同生成公钥：

$$K=\sum_{i=1}^{n}K_i=\sum_{i=1}^{n}k_iG.$$

1. 对于每个 $i\in[1,n]$，参与者 P_i 操作如下：

(a) P_i 将自己的有序数组 T_i 按照式(4)构造对应的 N 维向量：

$$\boldsymbol{Y}_i=(y_{i1},\cdots,y_{iN}).$$

(b) P_i 对 $\boldsymbol{Y}_i$ 中的分量 y_{ij} 进行编码，计算 $M_{ij}=y_{ij}G$，$j\in[1,N]$，得到：

$$M_i=(M_{i1},\cdots,M_{iN}).$$

2. P_1 应用公钥 K 加密 M_1(即加密 M_1 的每个分量)，得到：

$$E(M_1)=(E(M_{11}),\cdots,E(M_{1N}))\tag{8}$$

P_1 将 $Z_1=E(M_1)$ 发送给 P_2.

3. 对于 $i=2,\cdots,n-1$，P_i 收到 P_{i-1} 发送来的 $Z_{i-1}=(z_{(i-1)1},\cdots,z_{(i-1)N})$ 后，依次进行下面运算：

(a) P_i 根据 $\boldsymbol{Y}_i$ 的值对 Z_{i-1} 进行重随机化并进行替换，得到向量：

$$\boldsymbol{Z}_i=(z_{i1},\cdots,z_{iN}),$$

其中对 $j\in[1,N]$，

$$z_{ij}=\begin{cases}E(M_{ij}), & y_{ij}=1\\ R(z_{(i-1)j}), & y_{ij}=0\end{cases}\tag{9}$$

(b) P_i 将 $\boldsymbol{Z}_i$ 发送给 P_{i+1}.

(c) P_n 最后得到密文向量 $\boldsymbol{Z}_n=(z_{n1},\cdots,z_{nN})$，并公布.

4. 参与者 P_i，$i\in[1,n]$ 计算如下公式

$$C_{is}=R\Big(\sum_{j=1}^{t_{is}}z_{nj}\Big)=(c_{is1},c_{is2}),$$

并公开.

5. 参与者 P_l，$l\in[1,n]$ 计算 k_lc_{is2}，并公布.

6. P_i 计算 $R_{is}=c_{is1}-(k_1+\cdots+k_n)c_{is2}$，并将 R_{is} 与集合 $\{jG,j\in[1,N]\}$ 中元素作比较，如果 $R_{is}=j_0G$，则令 $r_{is}=j_0$. P_i 输出 j_0.

3.3 协议的正确性

定理 1. 协议 1 能正确计算归一排序序位.

证明. 根据命题 1，只需证明 $r_{is}=\sum_{j=1}^{t_{is}}z_{nj}$ 即可.

首先，在协议 1 第 1 步中，P_i 将自己的有序数组 T_i 按照式(4)构造对应的 N 维向量 $\boldsymbol{Y}_i$，再将 $\boldsymbol{Y}_i$ 中分量 y_{ij}，$j\in[1,N]$ 应用编码方式 $M_{ij}=y_{ij}G$ 编码到椭圆曲线上，所应用的编码方式可以保证参与者在对 $\boldsymbol{Y}_i$ 数组中的分量进行求和计算时也具有加法同态性.

协议 1 第 2 步和第 3 步是 P_i 对 Z_{i-1} 进行保密替换，最后在第 3(c)步获得 $\boldsymbol{Y}_n^*$ 对应编码向量的密文向量 $\boldsymbol{Z}_n$. 并且在协议第 4 步，P_i 为了确定 t_{is} 对应的序位 r_{is}，对加密向量的前 t_{is} 项求和，根据加密算法的加法同态性，第 4 步所获得的 C_{is} 确为 r_{is} 编码数据的密文.

协议 1 第 5 步和第 6 步是所有参与者进行联合解密得到 R_{is}，其中 $R_{is}=r_{is}G$. 由于椭圆曲线上离散对数的困难性，一般来说由 $r_{is}G$ 反求 r_{is} 是困难的，但可以通过将 R_{is} 与 $\{jG,j\in[1,N]\}$ 中元素逐个比较. 若 $R_{is}=j_0G$，则可知 j_0 即为我们所求的 t_{is} 在联合数组 T 中的排序序位 r_{is}. 证毕.

3.4 协议的安全性

本节我们构造式(1)模拟器，证明协议 1 的安全性. 半诚实模型下的恶意行为只能发生在协议执行之后，不能影响协议的输出[23]，因此有下面结论.

定理 2. 在半诚实模型下，以命题 1 为基础，协议 1 能安全地计算归一排序，而且可以抵抗最大数量的合谋攻击者，也就是 $n-1$ 个参与者的合谋攻击.

证明. 因为 P_i 的地位均是平等的，所以我们仅证明 P_1 数组中的数据以及排序信息是安全的即可. 为此只需要证明协议 1 对于合谋者集合 $I=\{P_2,\cdots,P_n\}$ 是安全的即可(因为 I 的任何子集能够得到的信息都不超过 I 能够得到的信息). 对于 I 中 $n-1$ 个参与者构成的合谋者集合，需要构造相应的模拟器 S，使得式(1)成立. S 首先构造椭圆曲线公钥系统，设其公钥为 K'，对应的私钥 $k'=k'_1+\cdots+k'_n$. S 按如下方式运行：

(1) 接受输入 $(I,T_I,f_I(T_1,\cdots,T_n))$ 以后，S 随机选取 $T'_1=(t'_{11},\cdots,t'_{1e_1})$，使得 $f_I(T_1,\cdots,T_n)=f_I(T'_1,T_2,\cdots,T_n)$.

(2) S 将 T'_1 按照方式(4)转化为向量 $\boldsymbol{Y}'_1=(y'_{11},\cdots,y'_{1N})$，将所有分量编码到椭圆曲线上得到 $M'_1=(M'_{11},\cdots,M'_{1N})$，加密 M'_1，得到：

$$E(M'_1)=(E(M'_{11}),\cdots,E(M'_{1N})).$$

(3) S 根据向量 $\boldsymbol{Y}_2,\cdots,\boldsymbol{Y}_n$ 的取值情况类似于协议第 3 步对 $E(M'_1)$ 的分量进行替换或重随机化，最终得到：

$$Z'_n=(z'_{n1},\cdots,z'_{nN}).$$

(4) 对于 $i=1$，$s\in[1,e_1]$，S 计算

$$C'_{1s}=R\Big(\sum_{j=1}^{t'_{1s}}z'_{nj}\Big)=(c'_{1s1},c'_{1s2}).$$

对于 $i\in[2,n]$，$s\in[1,e_i]$，S 计算

$$C'_{is}=R\Big(\sum_{j=1}^{t_{is}}z'_{nj}\Big)=(c'_{is1},c'_{is2}).$$

(5) 对于 $i\in[1,n]$，$s\in[1,e_i]$，S 计算 $k'_1c'_{is2}$，$k'_2c'_{is2},\cdots,k'_nc'_{is2}$，对 C'_{is} 解密并进一步解码得到 r'_{is}.

在协议中，

$$view_1^{\pi}(T_1,T_2,\cdots,T_n)=$$
$$\{T_2,\cdots,T_n,(E(M_{11}),\cdots,E(M_{1N})),(c_{1s1},c_{1s2})$$
$$(s\in[1,e_1]),k_1c_{is2}(i\in[2,n],s\in[1,e_i]),$$
$$f_I(T_1,T_2,\cdots,T_n)\}.$$

令

$$S(T_2,\cdots,T_n,f_I(T_1,T_2,\cdots,T_n))=$$
$$\{T_2,\cdots,T_n,(E(M'_{11}),\cdots,E(M'_{1N})),(c'_{1s1},c'_{1s2})$$
$$(s\in[1,e'_1]),k'_1c_{is2}(i\in[2,n],s\in[1,e_i]),$$
$$f_I(T'_1,T_2,\cdots,T_n)\}.$$

一方面，本文应用 (n,n) 门限椭圆密码系统，即 I 中所有参与者合谋操作也无法对任何一个密文解密. 另一方面，椭圆密码系统有语义安全性，因此对于 $(E(M_{11}),\cdots,E(M_{1N}))$ 与 $(E(M'_{11}),\cdots,E(M'_{1N}))$，$(c_{1s1},c_{1s2})$ 与 (c'_{1s1},c'_{1s2}) 都是计算不可区分的.

由于椭圆曲线密码系统存在离散对数困难性的问题，根据 $k_1c_{is2}(i\in[2,n],s\in[1,e_i])$，$I$ 中任意一个参与者无法分离 k_1c_{is2}，更不用说得到 P_1 的私钥 k_1，从而认为 k_1c_{is2} 与 $k'_1c'_{is2}$ 计算不可区分.

最后，由于 $f_I(T_1,T_2,\cdots,T_n)=f_I(T'_1,T_2,\cdots,T_n)$，因此

$$\{view_I^{\pi}(T_1,\cdots,T_n)\}_{T_i\subseteq[1,N]^{e_i},i\in[1,n]} \overset{c}{\equiv}$$
$$\{S(T_2,\cdots,T_{n-1},f_I(T_1,\cdots,T_n))\}_{T_i\subseteq[1,N]^{e_i},i\in[1,n]}.$$

证毕.

4　并列同序位多重排序

4.1　协议的基本原理

问题描述. n 个参与者 $P_1,\cdots,P_n$ 分别拥有有序数组 $T_1,\cdots,T_n$，他们希望进行合作保密计算，计算结束后参与者 P_i 仅可获知自己的每个元素 $t_{is}\in T_i$，$i\in[1,n]$，$s\in[1,m_i]$ 在联合数组 $T=(T_1,\cdots,T_n)$ 中的多重排序序位 r_{is}.

计算原理. (1) 参与者 $P_1,\cdots,P_n$ 商定一个全集 $J=[1,N]$，使得 T_i，$i\in[1,n]$ 的所有元素都属于 J. 为了方便起见，将 T_i 以重数方式形式表达：$\hat{T}_i=[(t_i^1,d_i^1),\cdots,(t_i^{e_i},d_i^{e_i})]$，其中 t_i^k，$k\in[1,e_i]$ 为 T_i 中所有互异的元素，满足 $t_i^1<t_i^2<\cdots<t_i^{e_i}$，这些元素在 T_i 中的重复次数分别为 $d_i^1,d_i^2,\cdots,d_i^{e_i}$（元素仅出现一次时约定重复次数为 1）.

(2) 参与者 P_i 首先构造一个 N 维向量 $\boldsymbol{U}_i$ 与 $\hat{T}_i$ 相对应.

$$\boldsymbol{U}_i=(u_{i1},u_{i2},\cdots,u_{iN}) \tag{10}$$

其中 $1\le s<t_i^1$ 时，$u_{is}=0$；当 $t_i^1\le s<t_i^2$ 时，$u_{is}=d_i^1$；当 $t_i^2\le s<t_i^3$ 时，$u_{is}=d_i^1+d_i^2$；依次类推，当 $t_i^{e_i}\le s\le N$ 时，$u_{is}=d_i^1+d_i^2+\cdots+d_i^{e_i}$. 以此方式，$P_i$ 拥有的数组 T_i 与向量 $\boldsymbol{U}_i$ 相对应.

(3) 以参与者 P_i 的向量 $\boldsymbol{U}_i$ 作为矩阵 $\boldsymbol{U}$ 的第 i 行，可以构成下面的 $n\times N$ 阶矩阵 $\boldsymbol{U}$：

$$\boldsymbol{U}=\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1N} \\ u_{21} & u_{22} & \cdots & u_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nN} \end{pmatrix}.$$

(4) P_i 按下面方式计算 t_i^k，$k\in[1,e_i]$ 在联合数组 $T=(T,\cdots,T_n)$ 中的排序位次 r_{ik}（由于 T 中重复元素的排序位次相同，故只需确定 T_i 中所有不同元素的位次即可）：

$$r_{ik}=\begin{cases} 1, & t_i^k=1 \\ \sum_{s=1}^{n} u_{s(t_i^k-1)}+1, & t_i^k\ne 1 \end{cases} \tag{11}$$

命题 2.　对于多重排序，由式(11)给出的序位计算方法是正确的.

证明.　首先将联合数组 T 表示为重数表示形式：$\hat{T}=[(x_1,k_1),\cdots,(x_m,k_m)]$.

(1) 根据 T 的构造，T 中各元素的排序序位为：

x_1 的序位应为 $r_1=1$，x_2 的序位应为 $r_2=k_1+1$，x_j 的序位应为 $r_j=k_1+k_2+\cdots+k_{j-1}+1,\cdots,x_m$ 的序位应为 $r_m=k_1+\cdots+k_{m-1}+1$.

(2) 根据向量 $\boldsymbol{U}_i$ 和矩阵 $\boldsymbol{U}$ 的构成方式，矩阵 $\boldsymbol{U}$ 中各列元素和分别为：

x_1 列之前各列元素之和均为 $s_1=0$；从 x_1 列直到 x_2-1 列，各列元素之和均为 $s_2=k_1$；从 x_j 列直到 $x_{j+1}-1$ 列，各列元素之和均为 $s_{j+1}=k_1+k_2+\cdots+k_j$；最后，我们可以发现从第 x_{m-1} 列到第 x_m-1 列，对每列元素求和都是 $s_m=k_1+\cdots+k_{m-1}$.

对于每一个 $j\in[1,m]$，比较(1)中的 r_j 和(2)中的 s_j 可知 $r_j=s_j+1$，这说明 T 中的元素 x_j 的序位刚好等于 $\boldsymbol{U}$ 中第 x_j-1 列之和，再加 1. 因此对于 $i\in[1,n]$，$s\in[1,m_i]$，T_i 中的元素 t_{is}，总存在 x_j 使得 $t_{is}=x_j$，因此 t_{is} 的排序序位满足式(11).　　证毕.

实例 2.　参与者 P_1,P_2,P_3 分别拥有秘密有序数组 $T_1=(2,2,2,3)$，$T_2=(2,3,3,5,7)$ 和 $T_3=(4,4,5,6)$. 令 $N=9$，P_1,P_2,P_3 各自按照式(10)构造一个 9 维向量，并将其合成为以下 3×9 矩阵：

$$\boldsymbol{Z}=\begin{pmatrix} 0 & 3 & 4 & 4 & 4 & 4 & 4 & 4 & 4 \\ 0 & 1 & 3 & 3 & 4 & 4 & 5 & 5 & 5 \\ 0 & 0 & 0 & 2 & 3 & 4 & 4 & 4 & 4 \end{pmatrix}.$$

根据矩阵 $\boldsymbol{Z}$，各参与者按照方式(11)确定其所有元素在联合数组中的多重排序序位如表 2 所示.

表 2　按照并列同序位多重法得到的排序结果

元素	序位	算法
2	1	0+1
3	5	4+1
4	8	7+1
5	10	9+1
6	12	11+1
7	13	12+1

根据表 2，元素 2 在联合数组中的排序位置是 1，为 0+0+0+1；元素 3 在联合数组中的排序位置是 5，为 3+1+0+1；其他类似，每个元素的排序位置都满足式(11). 由于此例中有序的联合数组 T 为 (2,2,2,2,3,3,3,4,4,5,5,6,7)，直接检验可知表 2 的排序序位是正确的.

4.2　多重排序协议

下面应用椭圆曲线门限密码体制，设计构造保密计算协议 2.

协议 2. 多重排序序位保密计算.

输入:$P_i(i\in[1,n])$各自的秘密有序数组 $\hat{T}_i=[(t_i^1,d_i^1),\cdots,(t_i^{e_i},d_i^{e_i})]$.

输出:对于任意给定的 $i\in[1,n],k\in[1,e_i]$,P_i输出 T_i 的元素 t_i^k在联合数组 $T=(T_1,\cdots,T_n)$中的多重排序序位 r_{ik}(下面假设 $t_i^k\neq1$,否则直接输出 $r_{ik}=1$ 即可).

准备:n 个参与者 P_i,$i\in[1,n]$首先商定一条椭圆曲线 $E_p(a,b)$以及椭圆曲线上的一个基点 G.每个参与者 P_i选择小于 L 的整数 k_i作为私钥,计算 $K_i=k_iG$,共同生成公钥:

$$K=\sum_{i=1}^{n}K_i=\sum_{i=1}^{n}k_iG.$$

1. 对于每个 $i\in[1,n]$,参与者 P_i操作如下:

(a) P_i将自己的有序数组 $\hat{T}_i$按照方式(10)构造 N 维向量:

$$\boldsymbol{U}_i=(u_{i1},\cdots,u_{iN}).$$

(b) P_i对 $\boldsymbol{U}_i$中的分量编码,得到:

$$M_i=(M_{i1},\cdots,M_{iN}).$$

(c) P_i应用公钥 K 加密 M_i(即加密 M_i的每个分量),得到 $\boldsymbol{C}_i$,并公布:

$$\boldsymbol{C}_i=(C_{i1},\cdots,C_{iN})=(E(M_{i1}),\cdots,E(M_{iN})) \tag{12}$$

2. 所有参与者 $P_i(i\in[1,n])$构造下面矩阵:

$$\boldsymbol{M}=\begin{pmatrix}C_{11}&C_{12}&\cdots&C_{1N}\\C_{21}&C_{22}&\cdots&C_{2N}\\\vdots&\vdots&\ddots&\vdots\\C_{n1}&C_{n2}&\cdots&C_{nN}\end{pmatrix}.$$

3. 对任意给定的 $i\in[1,n]$以及 $k\in[1,e_i]$,操作:

(a) 参与者 P_i计算:

$$R_{ik}=\sum_{s=1}^{n}C_{s(t_i^k-1)}+E(G)=(A_{ik},B_{ik}) \tag{13}$$

(b) 所有参与者 P_l,$l\in[1,n]$分别计算 k_lB_{ik},并公布.

(c) 参与者 P_i计算 $G_{ik}=A_{ik}-\sum_{l=1}^{n}k_lB_{ik}$,$P_i$进一步解码 G_{ik}得到 h_{ik}.

(d) P_i输出 $h_{ik}(r_{ik})$.

4.3 协议 2 的正确性

定理 3. 协议 2 能正确计算多重排序序位.

证明. 根据命题 2,当 $t_i^k\neq1$ 时,只需证明 $h_{ik}=r_{ik}=\sum_{s=1}^{n}u_{s(t_i^k-1)}+1$ 即可.

首先,在协议 2 第 1 步中,每个 P_i将自己的有序数组 $\hat{T}_i$按照式(10)构造对应的 N 维向量 $\boldsymbol{U}_i$,再将 $\boldsymbol{U}_i$中分量应用编码方式 $M_{ij}=u_{ij}G$ 编码到椭圆曲线上并进行加密,参与者 P_i最终得到密文向量 $\boldsymbol{C}_i$.所用编码方式保证了在应用椭圆曲线加密方案进行计算时对 $\boldsymbol{U}_i$中的分量数据也具有加法同态性.

协议 2 的第 2 步是参与者根据计算原理构造矩阵 $\boldsymbol{U}$ 对应编码矩阵的密文矩阵 $\boldsymbol{M}$.

协议 2 第 3 步针对任意给定的 $i\in[1,n]$以及 $k\in[1,e_i]$,计算 t_i^k在联合数组 T 中的排序序位.

在第 3(a)步中,P_i计算了 $\boldsymbol{M}$ 中第 t_i^k-1 列所有元素的和再加 $E(G)$,得到 R_{ik}.由于 G 是数据 1 的编码点,根据加密算法的加法同态性,R_{ik}等价于 $\boldsymbol{U}$ 中第 t_i^k-1 列元素和再加 1.第 3(b)和 3(c)步是所有参与者合作解密 R_{ik},P_i最终得到解密结果 G_{ik},并对 G_{ik}进行解码得到 h_{ik}.根据计算原理,h_{ik}为 t_i^k在 T 中的排序序位,即 $h_{ik}=r_{ik}$.因此协议 2 是正确的.

证毕.

4.4 协议 2 的安全性

分析协议 2 的安全性,需要分析协议每一步的安全性.在协议前两步中所有的参与者对自己的数组编码并加密,其地位是平等的.第 3 步的目的是计算 P_i的元素 t_i^k的排序序位.在这个过程中,P_i首先计算 $\boldsymbol{M}$ 的第 t_i^k-1 列元素的和,求和后 P_i又对其加了一个"1"的编码点 G 的密文 $E(G)$,最后得到 R_{ik}.由于 $E(G)$是 P_i加密的密文,根据应用的门限密码系统以及密文的不可区分性,可以了解即使 $n-1$ 个参与者进行合谋也无法分离出密钥以及得到 t_i^k的任何信息.协议最后一步参与者 P_i合作解密 R_{ik},由于解密结果只有 P_i知道,因此 P_i元素的排序序位信息也是安全的.根据加密算法的安全性,P_i在协助其他参与者进行解密的过程中也无任何信息泄露.这就证明了任意一个参与者 P_i的数据和排序信息都是安全的.根据定理 2,我们省略定理 4 的证明过程.

定理 4. 在半诚实模型下,以命题 2 为基础,协议 2 能安全计算多重排序,而且可以抵抗最大数量的合谋攻击者,也就是 $n-1$ 个参与者的合谋攻击.

5 增序位全排序序位保密计算

5.1 协议的基本原理

问题描述. n 个参与者 P_i,$i\in[1,n]$各自拥有秘密有序数组 T_i,将其表达为重数形式 $\hat{T}_i=[(t_i^1,d_i^1),\cdots,(t_i^{e_i},d_i^{e_i})]$,他们希望进行合作保密计算,最后参与者 P_i可获知自己每个元素在联合数组 $T=(T_1,\cdots,T_n)$中按照增序排序的序位.

计算原理. 与多重排序方式类似,刚开始每个参与者需要将自己的保密有序数组 $\hat{T}_i$按照式(10)构造向量 $\boldsymbol{U}_i$,然后将这些向量作为行向量构成矩阵 $\boldsymbol{U}$.不同于多重排序方式,在联合数组中除了要求不同元素按先后顺序排列,所有相同元素的序位也要

依次递增. 因此，在增序位全排法方式下 P_i 按下面方式计算其元素 t_i^k 在联合数组 T 中的首位排序序号(完整序位为从 r_{ik} 到 $r_{ik}+d_i^k-1$).

$$r_{ik}=\begin{cases}\sum_{s=1}^{n}u_{s(t_i^k-1)}+1, & i=1\\ \sum_{s=1}^{i-1}u_{st_i^k}+\sum_{s=i}^{n}u_{s(t_i^k-1)}+1, & 1<i\le n\end{cases}\tag{14}$$

命题 3.　对于增序排序，由式(14)给出的排序计算方法是正确的.

证明.　在下面证明中，把 $\boldsymbol{U}$ 的第 j 列记为向量 $\boldsymbol{v}_j$，并借用记号 $x\in X$ 表示数组 X 包含元素 x，以及 ST 表示有序的联合数组.

将联合数组 T 改写成扩充重数表示形式：$[(1,k_1),(2,k_2),\cdots,(N,k_N)]$，其中 $k_j, j\in[1,N]$ 为元素 $j\in T$ 的重数(对于某个 $j_0\in[1,N]$，如果 j_0 不属于 T，则记 $k_{j_0}=0$).

根据矩阵 $\boldsymbol{U}$ 的构成方式，向量 $\boldsymbol{v}_j$ 各分量的和是有序联合数组 ST 中不大于 j(包括 j)的所有元素的和，即为 $k_1+k_2+\cdots+k_j$.

对于 $i\in[1,N]$，考虑某个数据 $m\in ST$ 对于参与者 P_i 的排位问题，假设 P_i 具有数据 m 的重数是 s_i.

(1) 如果 $m\in T_1$，对于 T_1 中的 m，在 ST 中的最小序位是 $r_{1m}=k_1+\cdots+k_{m-1}+1$，即 v_{m-1} 个分量的和再加 1，这表明式(14)的第一式是正确的.

(2) 我们使用数学归纳法来证明式(14)的第二式.

如果 $m\in T_2$，对于 T_2 中的 m，在 ST 中的最小序位 r_{2m} 应是在 r_{1m} 的基础上再加上 m 在 T_1 中的重数 s_1，这个数据恰是把 v_{m-1} 的第一个分量换为 v_m 的第一个分量后再求和，并加 1 而得到. 因此式(14)第二式对 $i=2$ 成立.

假设对于 $i\ge 3, m\in T_i$，并且 T_i 中的 m 在 ST 中的最小序位 r_{im} 可由式(14)的第二式进行计算，即把 v_{m-1} 的前 $i-1$ 个分量都换为 v_m 的对应分量后再求和，并加 1 而得到. 进一步，如果 $m\in T_{i+1}$，对于 T_{i+1} 中的 m，在 ST 中的最小序位 $r_{(i+1)m}$ 应是在 r_{im} 的基础上再加上 m 在 T_i 中的重数 s_i，这个数据恰是把 v_{m-1} 的前 i 个分量都换为 v_m 的对应分量后再求和，并加 1 而得到. 因此式(14)的第二式也是正确的.　　证毕.

根据实例 2 中的矩阵 $\boldsymbol{Z}$，各参与者按照方式(14)确定其所有元素在联合数组中的序位如表 3 所示.

表 3　按照增序位全排法得到的排序结果

ST 中互异元素	T_1 元素序位	T_2 元素序位	T_3 元素序位
2	1～3	4	—
3	5	6～7	—
4	—	—	8～9
5	—	10	11
6	—	—	12
7	—	13	—

根据表 3，P_1 的元素 2 在联合数组中的排序位置是 1～3，首位序号 1 为 0+0+0+1；P_2 的元素 2 在联合数组中的排序位置是 4，为 0+0+3+1；P_1 的元素 3 的排序位置是 5，为 3+1+0+1，P_2 的元素 3 的排序位置是 6～7，首位序号 6 为 1+0+4+1；其他类似. 由于此例中 3 个数组构成的有序联合数组为(2,2,2,2,3,3,3,4,4,5,5,6,7)直接检验可知表 3 的序位排法是正确的.

5.2　增序排序序位保密计算

命题 3 是本文计算增序排序的基本原理，下面将应用椭圆曲线密码体制设计保密计算协议. 增序排序序位的保密计算协议只需将协议 2 的第 3 步进行修改，其余保持不变，具体操作参看协议 3(修改见协议 3 的 3′).

协议 3.　增序位全排法序位保密计算.

输入：$P_i(i\in[1,n])$ 各自的秘密有序数组 $\hat{T}_i=[(t_i^1,d_i^1),\cdots,(t_i^{e_i},d_i^{e_i})]$.

输出：参与者 P_i 仅可获得自己的任意元素(记为 m)在联合数组 T 中按照增序排序的首位排序序位 $r_i(m)$.

3′. 对于任意给定的 $i\in[1,n]$ 以及 $m\in T_i$，操作：

(a) 参与者 P_i 计算

$$R_i(m)=(A_i(m),B_i(m))=\begin{cases}\sum_{s=1}^{n}C_{s(m-1)}+E(G), & i=1\\ \sum_{s=1}^{i-1}C_{sm}+\sum_{s=i}^{n}C_{s(m-1)}+E(G), & 1<i\le n\end{cases}\tag{15}$$

(b) 所有参与者 $P_l, l\in[1,n]$ 分别计算 $k_l B_i(m)$，并公布.

(c) 参与者 P_i 计算 $G_i(m)=A_i(m)-\sum_{l=1}^{n}k_l B_i(m)$，$P_i$ 进一步解码 $G_i(m)$ 得到 $r_i(m)$.

(d) P_i 输出 $r_i(m)$.

5.3　协议 3 的正确性和安全性

协议 3 和协议 2 非常类似，在数据编码以及加密和解密等方面完全相同，只有根据密文矩阵 $\boldsymbol{M}$ 进行计算时所用的计算公式不同，协议 2 中式(13)是实现式(11)的保密计算，而协议 3 中式(15)是实现式(14)的保密计算. 因此两个协议的区别主要是在

第 3(a)步对于密文的计算方面,但过程极其相似.根据协议 2 和命题 3,类似定理 3 和定理 4,我们省略定理 5 的证明过程.

定理 5. 协议 3 在半诚实模型下可正确和安全地计算增序排序,能抵抗 $n-1$ 个参与者的合谋攻击.

6 效率分析与比较

本文所设计的协议与文献[18-20]所研究问题最为接近.因此,本节中首先对协议 1～3 的计算效率进行详细分析,然后将我们的结果与文献[18-20]的结果进行全面比较.

本文协议 1～3 均应用椭圆曲线密码系统保密实现安全多方多数据排序,该加密方案进行的是模加运算,模加运算的复杂性主要跟加密明文消息时选择随机数 r 的位数有关联($0\leq r\leq p-1$).其他协议主要应用的是 Paillier 加密系统、椭圆曲线加密系统、RSA 加密系统以及 Lifted ElGamal 加密系统.为了方便分析,统一以模加运算和模指数运算的次数作为测评方案计算复杂性的基准,用 M_a 表示模加运算,M_e 表示模指数运算;另外,测评方案通信复杂性的基准统一为通信次数.

6.1 本文协议复杂性分析

计算复杂性. 在协议 1 中,参与者 P_1 对数组 $\boldsymbol{Y}_1$ 中的分量全部加密,$P_i, i\in[2,n]$ 需加密数组 $\boldsymbol{Y}_i$ 中 e_i 个分量,参与者 $P_i, i\in[1,n]$ 合作解密 $(e_1+\cdots+e_n)$ 次.因此协议 1 计算复杂性为 $(N+e_1+2(e_2+\cdots+e_n))\log r[M_a]$.在协议 2 中,每个参与者 $P_i, i\in[1,n]$ 需要加密 $N+e_i$ 个元素.参与者 $P_i, i\in[1,n]$ 利用自己的私钥合作解密 $(e_1+\cdots+e_n)$ 次.因此协议 2 计算复杂性为 $(nN+2(e_1+\cdots+e_n))\log r[M_a]$.由于协议 3 和协议 2 计算原理基本相同,只是计算方式不同,因此计算复杂性与协议 2 一致.

通信复杂性. 在协议 1 中构造公钥需要 $n-1$ 次通信.参与者 P_i 传递 $E(M_i)$,需要 $n-1$ 次通信.P_n 公布 $\boldsymbol{Z}$ 需要 1 次通信.协议 1 最后三步共需要 $2n$ 次通信.因此执行协议 1 总共需要 $4n-1$ 次通信,通信复杂性是 $O(n)$.在协议 2 中构造公钥需要 $n-1$ 次通信.参与者 P_i 公布 $\boldsymbol{C}_i$,需要 n 次通信.参与者公开 R_{ik} 和 k_lB_{ik} 共需要 $2n$ 次通信.所以协议 2 共需要 $4n-1$ 次通信,通信复杂性是 $O(n)$.由于协议 3 和协议 2 计算原理基本相同,只是计算方式不同,因此通信复杂性与协议 2 一致.

6.2 与已有结果比较

本文协议 1～3 既可以解决参与者持有一个数据的排序问题,也可以解决参与者持有多个数据的排序问题.文献[18]提出的方案适用于每个参与者持有一个数据,文献[19-20]提出的方案适用于每个参与者持有多个数据.为方便比较,我们分两种情形讨论,并统一用 N 表示每个参与者拥有数组的维数,n 表示参与者的个数,CF 表示计算功能,C_1 表示计算复杂性,C_2 表示通信复杂性.

情形 1. 参与者持有一个数据.文献[18]协议 2 应用 Paillier 密码体制,协议 3 应用门限椭圆曲线密码体制解决了数据排序问题.

本文协议 1～3 与文献[18]详细比较如表 4 所示,其中 A_1 表示是否可以推广到多方多数据出现重复元素的排序问题.

表 4 与文献[18]效率分析和应用场景的比较

	CF	C_1	C_2	A_1
协议 1	归一排序	$(N+2n-1)\log r[M_a]$	$O(n)$	√
协议 2	多重排序	$n(N+2)\log r[M_a]$	$O(n)$	√
协议 3	增序排序	$n(N+2)\log r[M_a]$	$O(n)$	√
文献[18]协议 2	普通排序	$2n(N+2)[M_e]$	$O(n)$	×
文献[18]协议 3	普通排序	$n(N+2)\log r[M_a]$	$O(n^2)$	×

根据表 4,协议 1 的计算复杂性比文献[18]协议 2～3 低,通信复杂性和文献[18]协议 2 一样,比文献[18]协议 3 低;协议 2～3 的计算复杂性比文献[18]协议 2 低,通信复杂性和文献[18]协议 2 一样.协议 2～3 计算复杂性和文献[18]协议 3 一样,通信复杂性比文献[18]协议 3 低.

文献[18]方案的思想无法推广到多方多数据出现重复元素的排序问题.本文提出的协议 1～3 研究三种不同的排序,可以推广到参与者数组和联合数组为多重数组的排序问题.

情形 2. 参与者持有多个数据.文献[19]研究安全多方多数据排序,利用 RSA 加密体制并借助半可信的第三方和不经意传输协议.

文献[20]所研究问题与文献[19]类似.假设文献[20]应用门限 Lifted ElGamal 加密方案,避免半可信第三者和不经意传输,但也不能进行完全解密.

本文协议 1～3 与文献[19-20]详细比较如表 5 所示,其中 e_i 表示每个参与者拥有私密数据的个数,A_2 表示是否适用于参与者数组以及联合数组为多重数组.

表 5　与文献[19-20]效率分析和应用场景的比较

	CF	C_1	C_2	A_2
协议 1	归一排序	$(N+e_1+2(e_2+\cdots+e_n))\log r\ [M_a]$	$O(n)$	√
协议 2	多重排序	$(nN+2(e_1+\cdots+e_n))\log r\ [M_a]$	$O(n)$	√
协议 3	增序排序	$(nN+2(e_1+\cdots+e_n))\log r\ [M_a]$	$O(n)$	√
文献[19]	普通排序	$3nN+2(n+1)(e_1+\cdots+e_n)[M_e]$	$O(n^2)$	×
文献[20]	普通排序	$2nN+(e_1+\cdots+e_n)\ [M_e]$	$O(n^2)$	×

根据表 5，本文协议 1～3 的计算复杂性和通信复杂性都比文献[19-20]低. 文献[18-19]的方案仅仅适合参与者拥有的数组以及联合数组为标准数组的排序问题，本文提出的排序方案适用于多重数组.

综上所述，无论是情形 1 还是情形 2，本文提出的排序方案不仅在计算复杂性和通信复杂性占据优势，而且在解决文献[18-20]的问题方面更具有普遍性.

6.3　本文结果的实际可行性

为了展现方案的实际可行性，本文使用 java 语言编程测试执行协议 1 所需要的时间.

实验测试环境. Windows XP 专业版 32 位操作系统，CPU 为 Intel(R) Core(TM) i3-2100 CPU@3.10 GHz，内存是 2.99 GB. 语言环境：MyEclipse. 程序语言：Java.

实验算法. 协议 1 中椭圆曲线密码系统选择素数 p,a,b 的长度为 256 比特；文献[18]协议 2 中 Paillier 加密系统选择素数的长度为 512 比特；文献[20]中 Lifted ElGamal 加密系统选择素数的长度为 512 比特.

实验方法. 实验中我们使用现成的开源包(https://sourceforge.net/projects/jecc/). 实验数据范围 1～100，设定 $N=20$，参与者个数分别设定为 $P_i, i\in\{3,\cdots,25\}$. 为了保证数据的准确性，进行 100 次模拟实验测试，然后对测试结果求平均值. 同时为了保证数据的科学性，与情形 1 和情形 2 相对应，我们的实验也分两种情况.

实验 1.　P_i有一个数据，由图 1 展示实验结果.

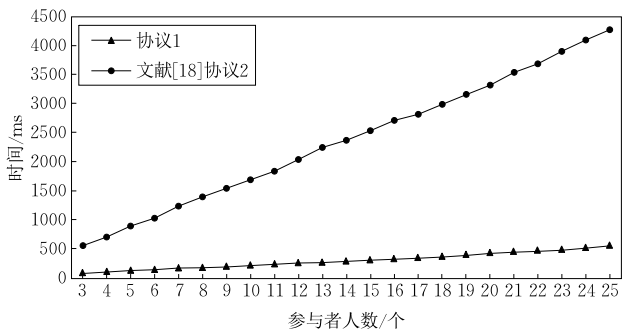

图 1　协议执行时间随参与者人数增长的变化规律

实验 2.　P_i有 10 个数据，由图 2 展示实验结果.

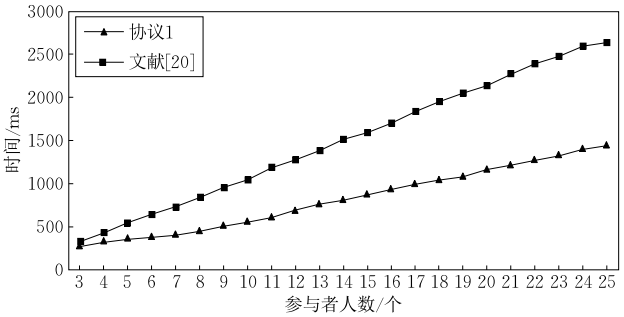

图 2　协议执行时间随参与者人数增长的变化规律

由图 1、图 2 可知，协议 1 和文献[18,20]在持有数据不变的情况下，协议执行时间随着参与者人数的增加大致呈线性增长，但协议 1 的增长率明显低于文献[18]协议 2 与文献[20]的协议增长率. 从图 1、图 2 中的数据可以看出无论解决的是参与者持有一个数据还是多个数据的排序问题，协议 1 都是高效的.

7　恶意模型下的排序协议

在本节中，我们将应用文献[23]中协议编译器的基本思想，利用输入承诺函数、认证计算函数、增强掷币函数以及零知识证明系统[28]，下面我们根据协议 2 的原理设计协议 4.

设计协议 4 需要解决 3 个问题：(1) 协议执行期间，保证后面所有的计算都在参与者的输入数据上面进行；(2) 防止任意一个参与者 P_i不按照协议步骤对数组加密以及对矩阵中密文元素操作乘积；(3) 合作解密时，保证所有参与者公布正确的解密信息. 具体我们有协议 4.

协议 4.　恶意模型下多重排序安全计算协议.

输入：参与者 $P_i(i\in[1,n])$各自的秘密有序数组 $\hat{T}_i=[(t_i^1,d_i^1),\cdots,(t_i^{e_i},d_i^{e_i})]$. 门限椭圆曲线密码系统：$P_i$私钥为 k_i，公钥为 K.

输出：对于任意给定的 $i\in[1,n],k\in[1,e_i]$，P_i输出 T_i 的元素 t_i^k在联合数组 $T=(T_1,\cdots,T_n)$中的多重排序序位 r_{ik} (假设 $t_i^k\neq1$，否则直接输出 $r_{ik}=1$ 即可).

在协议执行期间，所有参与者需要验证他们收到的所有数据的正确性，如有证明未通过验证，则中止协议.

1. P_i使用的随机数 r_i通过所有参与者联合调用增强掷币函数为其生成.

2. 对于每个 $i\in[1,n]$，P_i进行如下操作：

(a) P_i对数组 $\hat{T}_i$每一对元素 $(t_i^k,d_i^k),k\in[1,e_i]$进行承诺，并将自己的有序数组 $\hat{T}_i$按照方式(10)构造 N 维向量

$\boldsymbol{U}_i=(U_{i1},\cdots,U_{iN})$.

(b) P_i 对 $\boldsymbol{U}_i$ 中的每个分量进行编码得到 $M_i=(M_{i1},\cdots,M_{iN})$.

(c) P_i 加密 M_i 的每个分量,从而得到向量 $\boldsymbol{C}_i=(C_{i1},\cdots,C_{iN})=(E(M_{i1}),\cdots,E(M_{iN}))$. 并将密文 $\boldsymbol{C}_i$ 随同关于其对应明文知识的证明公布.

3. 所有参与者构造下面矩阵:

$$\boldsymbol{M}=\begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1N} \\ C_{21} & C_{22} & \cdots & C_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nN} \end{pmatrix}.$$

4. 对任意给定的 $i\in[1,n]$ 以及 $k\in[1,e_i]$,操作:

(a) P_i 应用认证计算函数计算下面密文:

$$R_{ik}=\sum_{s=1}^{n}C_{s(t_i^k-1)}+E(G)=(A_{ik},B_{ik}),$$

P_i 公布 R_{ik}.

(b) 参与者 $P_l,l\in[1,n]$ 计算 k_lB_{ik},并公布计算结果. 所有参与者应用文献[28]中的零知识证明系统验证 k_lB_{ik} 和 k_lG 中的 k_l 是否相同.

(c) P_i 计算

$$G_{ik}=A_{ik}-\sum_{l=1}^{n}k_lB_{ik},$$

P_i 进一步解码 G_{ik} 得到 h_{ik},并输出 $h_{ik}(r_{ik})$.

协议的安全性分析. 如果在协议执行中,所有参与者对于其收到的所有数据都能通过验证,协议 4 没有中止执行,则协议是安全的. 简要论证如下:

(1) 首先协议第 1 步中,要求所有参与者应用增强掷币函数为每个参与者生成执行协议所需要的随机带去构造随机数,这样即可保证协议执行期间参与者需要用到的数据确实是随机数.

(2) 在第 2(a)步中,首先要求每个参与者 $P_i,i\in[1,n]$ 将 $\hat{T}_i$ 按照式(10)构造对应的 N 维向量 $\boldsymbol{U}_i=(U_{i1},\cdots,U_{iN})$. 这时如果有一个恶意参与者没有按照协议 2 正确执行协议步骤,这相当于 P_i 修改原始集合 $\hat{T}_i$,我们在 2.2 节曾说明这种行为不予考虑,第 2(b)步同理.

(3) 在第 2(c)步中,P_i 在公布密文 $\boldsymbol{C}_i$ 时,同时公布了关于这些密文对应明文的证明,这可保证这些加密运算的正确性.

(4) 在第 4(a)步中,P_i 根据认证计算函数计算 R_{ik},能保证所公布的 R_{ik} 是正确的.

(5) 在第 4(b)步中,P_i 合作解密时,需要所有参与者公布各自的解密信息. 同时应用文献[28]提出的关于零知识证明的协议验证 k_lB_{ik} 的有效性,这可保证解密结果 G_{ik} 是正确的.

由此可知,在协议 4 中参与者都必须按要求正确地执行协议 2,任何偏离协议规定的行为都会引起协议中止. 如果协议正常结束,则结果一定是正确的. 又由于协议 4 是在协议 2 的基础上设计构造的,因此各参与者隐私数据的保密性仍然保持. 给出下面定理,限于篇幅,证明过程省略.

定理 6. 协议 4 在恶意模型下,安全计算多重排序.

8 结 论

本文研究了百万富翁的推广问题:多方保密排序问题,并定义了三种针对数组中有重复元素出现的排序:归一排序、多重排序,增序排序. 基于所设计的新的编码方法与具有加法同态性的椭圆曲线密码体制,对于这三种排序方案设计了安全高效的计算协议,协议适用于任意数组的保密排序. 在本文的最后给出了恶意模型下的排序协议. 理论分析和实验结果都表明本文协议是高效和实用的.

参 考 文 献

[1] Ben-Or M, Goldwasser S, Wigderson A. Completeness theorems for non-cryptographic fault-tolerant distributed computation//Proceedings of the 20th Annual ACM Symposium on Theory of Computing. Chicago, USA, 1988: 1-10

[2] 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

[3] Jonsson K V, Kreitz G, Uddin M. Secure multi-party sorting and applications//Proceedings of the Applied Cryptography and Network Security. Nerja, Spain, 2011: 122

[4] Damgard I, Meldgaard S, Nielsen J B. Perfectly secure oblivious RAM without random oracles//Proceedings of the 8th Conference on Theory of Cryptography. Berlin: Springer-Verlag, 2010: 108

[5] Huang Y, Evans D, Katz J. Private set intersection: Are garbled circuits better than custom protocols//Proceedings of the NDSS Symposium 2012. San Diego, Chile, 2012: 1-2

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

[7] Du W, Atallah M J. Privacy-preserving cooperative scientific computations//Proceedings of the 14th IEEE Workshop on Computer Security Foundations. Nova Scotia, Canada, 2001: 273-294

[8] Ajtai M, Komlos J, Szemeredi E. An $O(n\log n)$ sorting network//Proceedings of the 15th Annual ACM Symposium on Theory of Computing. New York, USA, 1983: 1-9

[9] Batcher K E. Sorting networks and their applications// Proceedings of the Spring Joint Computer Conference. New York, USA, 1968: 307-314

[10] Vitanyi P. On the Average-Case Complexity of Shellsort. Random Structures and Algorithms. Germany: Wiley Periodicals, Inc, 2018: 905-911

[11] Zhang B. Generic constant-round oblivious sorting algorithm for MPC//Proceedings of the International Conference on Provable Security. Xi'an, China, 2011: 240-256

[12] Hamada K, Kikuchi R, Dai I, et al. Practically efficient multi-party sorting protocols from comparison sort algorithms //Proceedings of the International Conference on Information Security and Cryptology. Seoul, Korea, 2012: 202-216

[13] Tang Chun-Ming, Shi Gui-Hua, Yao Zheng-An. Secure multi-party computation protocol for sequencing problem. Science China, 2011, 41(7): 789-797(in Chinese)
(唐春明，石桂花，姚正安. 排序问题的安全多方计算协议. 中国科学，2011，41(7)：789-797)

[14] Marszałek Z. Parallel fast sort algorithm for secure multiparty computation. Journal of Universal Computer Science, 2018, 24(4): 488-514

[15] Dai H, Ren H, Chen Z Y, et al. Privacy-preserving sorting algorithms based on logistic map for clouds. Security and Communication Networks. 2018: 2373545:1-2373545:10

[16] Wang Ning, Gu Hao-Min, Zheng Tong. A practical and efficient secure multi-party sorting protocol. Computer Applications and Software, 2018, 35(10): 311-317(in Chinese)
(王宁，顾昊旻，郑彤. 一种实用高效的安全多方排序协议. 计算机应用与软件，2018，35(10)：311-317)

[17] Yang D Q, Qu B Q, Philippe C M. Privacy-preserving social media data publishing for personalized ranking-based recommendation. IEEE Transactions on Knowledge and Data Engineering, 2019, 31(3): 507-520

[18] Li Shun-Dong, Kang Jia, Yang Xiao-Yi, et al. Secure multi-party characters sorting. Chinese Journal of Computers, 2018, 41(5): 206-222(in Chinese)
(李顺东，亢佳，杨晓艺等. 多个字符排序的安全多方计算. 计算机学报，2018，41(5)：206-222)

[19] Qiu Mei, Luo Shou-Shan, Liu Wen, et al. A solution secure multi-party multi-data ranking problem based on RSA encryption scheme. Journal of Electronics, 2009, 37(5): 1119-1123(in Chinese)
(邱梅，罗守山，刘文等. 利用 RSA 密码体制解决安全多方多数据排序问题. 电子学报，2009，37(5)：1119-1123)

[20] Liu W, Luo S S, Wang Y B, et al. A protocol of secure multi-party multi-data ranking and its application in privacy preserving sequential pattern mining//Proceedings of the 4th International Joint Conference on Computational Sciences and Optimization. 2011: 272-275

[21] Li S D, Wang D S, Dai Y Q. Efficient secure multiparty computational. Chinese Journal of Electronics, 2010, 19(2): 324-328

[22] Reimer B, Fried R, Mehler B, et al. Brief report: Examining driving behavior in young adults with high functioning autism spectrum disorders: A pilot study using a driving simulation paradigm. Journal of Autism and Developmental Disorders, 2013, 43(9): 2211-2217

[23] Goldreich O. Foundations of Cryptography: Volume 2, Basic Applications. London: Cambridge University Press, 2004: 599-764

[24] Yang Bo. Foundations of Modern Cryptography. Beijing: Tsinghua University Press, 2015: 106-112(in Chinese)
(杨波. 现代密码学基础. 北京：清华大学出版社，2015：106-112)

[25] Li L, El-Latif A A A, Niu X. Elliptic curve ElGamal based homomorphic image encryption scheme for sharing secret images. Signal Processing, 2012, 92(4): 1069-1078

[26] Desmedt Y, Frankel Y. Threshold cryptosystems//Proceedings of the on Advances in Cryptology. Santa Barbara, USA, 1989: 307-315

[27] Long Y, Chen K F. New constructions of dynamic threshold cryptosystem. Journal of Shanghai Jiaotong University(Science), 2014, 19(4): 431-435

[28] Chatzigiannakis I, Pyrgelis A, Spirakis P G, et al. Elliptic curve based zero knowledge proofs and their applicability on resource constrained devices//Proceedings of the IEEE 8th International Conference on Mobile Adhoc and Sensor Systems. Valencia, Spain, 2011: 715-720

LI Shun-Dong, Ph. D., professor, Ph. D. supervisor. His main research interests include modern cryptography and information security.

DU Run-Meng, M. S. candidate. Her main research interests include modern cryptography and information security.

YANG Yan-Jing, M. S. candidate. Her main research interests include modern cryptography and information security.

WEI Qiong, M. S. candidate. Her main research interests include modern cryptography and information security.

Background

Secure multiparty computation (SMC) plays an important role in information security and is a pivotal privacy preserving technology both in cyberspace and in cooperative computation. It has become a research focus in the international cryptographic community in recent years. Since SMC was introduced, cryptographic scholars have studied many SMC problems arising in various fields, such as scientific computation, computational geometry, data mining, statistical analysis and social management; there are also many new problems that must be studied, and many previously addressed problems also require further study in order to develop more efficient solutions.

In this study, we evaluated a new SMC problem, which consisted of ranking the union of arrays of participants in a multiple array; this work has important practical significance and broad application prospects in information security. This approach is of great significance to solve the ranking in real life. For example, the age, salary, achievement, and parameters of product span relatively small ranges in daily life. At the same time, these problems are extensively used in fields such as private biding and auction, anonymous voting and secure data mining. In addition, it is also of great mathematical significance to solve ranking. Based on ranking protocol, we can solve the maximum and minimum problems. This problem extends the famous millionaires' problem which is a basic problem in scientific computation. For example, if we want to privately compute the average of some private data, it is necessary to remove the outliers of a set of data, that is, the maximum and the minimum, to ensure the stability and accuracy of average. Our protocol can be used to do this.

The existing solutions mainly study the ranking problem when there is no repeated element in multi-party arrays and union arrays. This paper solves ranking problems in different scenarios and reduces the computational complexity. We design new encoding schemes to hide private numbers and to map plaintexts to points on an elliptic curve. Based on these new encoding schemes and the threshold decryption elliptic cryptosystem, we design secure ranking protocols for the following three ranking problems: the ranking in which the same numbers have the same order, and the order of the next number increases by 1; the ranking in which the same numbers have the same order, but if there are k same numbers, then the order of next number will increase by k; and the ranking in which the same numbers have different orders.

We have been studying SMC for more than ten years, and have done much work on this topic. Our work is supported by the National Natural Science Foundation of China (Grant No. 61272435).