软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software, 2022,33(11):4316–4333 [doi: 10.13328/j.cnki.jos.006326]

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

安全排序协议及其应用*

窦家维, 汪榆淋

(陕西师范大学 数学与统计学院, 陕西 西安 710119)

通信作者: 窦家维, E-mail: jiawei@snnu.edu.cn

摘 要: 安全多方计算(secure multi-party computation, SMC)是国际密码学界近年来的研究热点. 排序是一种基本的数据操作, 是算法研究中最基础的问题. 多方保密排序是百万富翁问题的推广, 是一个基本的 SMC 问题, 在科学决策、电子商务推荐、保密招标/拍卖、保密投票以及保密数据挖掘等方面有重要应用. 目前已有的安全多方排序解决方案大多只能适用于隐私数据范围已知而且范围较小的情况, 如果数据范围未知或者数据范围很大, 还未见到有效的解决方案. 首先, 在数据范围已知情形下, 针对同数据并列计位以及增位次计位两种不同排序方式设计保密计算协议, 进一步设计基于关键词的增位次计位方式保密排序协议; 其次, 以这些协议为基础, 在数据范围未知的情形下, 针对上述两种不同排序方式分别构造有效的保密排序方案. 应用该排序协议作为模块, 可解决许多以排序为基础的实际应用问题. 最后设计了一个安全、高效的保密 Vickrey 招投标协议, 以解决实际保密招标问题. 通过灵活运用编码技巧, 并基于 ElGamal 门限密码体制设计协议, 这些协议在半诚实模型下是安全、高效的. 应用模拟范例严格证明了协议的安全性, 并对协议的执行效率进行了实际测试. 实验结果表明, 该协议是高效的.

关键词: 安全多方计算; 保密排序; 同态加密; 门限解密; 保密招投标

中图法分类号: TP309

中文引用格式: 窦家维, 汪榆淋. 安全排序协议及其应用. 软件学报, 2022, 33(11): 4316–4333. http://www.jos.org.cn/1000-9825/6326.htm

英文引用格式: Dou JW, Wang YL. Secure Sorting Protocols and Their Applications. Ruan Jian Xue Bao/Journal of Software, 2022, 33(11): 4316–4333 (in Chinese). http://www.jos.org.cn/1000-9825/6326.htm

Secure Sorting Protocols and Their Applications

DOU Jia-Wei, WANG Yu-Lin

(School of Mathematics and Statistics, Shaanxi Normal University, Xi'an 710119, China)

Abstract: Secure multi-party computation (SMC) is a focus in the international cryptographic community in recent years. Sorting is a basic data operation and a basic problem of algorithm design and analysis. Secure multiparty sorting is the generalization of the millionaires' problem and a basic problem of SMC. It can be extensively used in scientific decision-making, e-commerce recommendation, electronic auction and bidding, anonymous voting and privacy-preserving data-mining, etc. Most existing solutions to sorting problem are applicable to the cases that the private data is known and small. If the data range is not known, they do not work. If the data range is very large, they will be very inefficient. Unfortunately, in practice, many application scenarios fall in these categories. To privately sort data in scenarios that data range is unknown or the data range is very large, two protocols are proposed first for these scenarios where the data range is small or is known to preserve the privacy of data: the scheme where the same data occupy the same order and that where the same data occupy different orders. Then, these protocols are used as building blocks to design schemes to solve the sorting problem in scenarios that data range is unknown or the data range is very large. The proposed new secure sorting protocols can be used as building blocks to solve many practical problems that inherently need sorting. Based on these protocols, a secure and efficient Vickrey auction protocol is designed. Encoding technique and threshold decryption ElGamal cryptosystem are flexibly used to design these protocols. Using the

* 基金项目: 国家自然科学基金(61272435)
收稿时间: 2020-08-31; 修改时间: 2020-12-05; 采用时间: 2021-01-19

simulation paradigm, it is proved that the protocols are secure in the semi-honest model. Finally, the efficiency of the protocols are tested. The experimental results show that the proposed protocols are efficient.

Key words: secure multi-party computation; privacy-preserving sorting; homomorrphic encryption; threshold decryption; secure bidding and auction

数据排序是一种基本的数据操作, 人们针对不同的数据情况提出了各种排序算法, 以尽可能地提高计算效率. 很多实际应用中的数据是机密数据或者隐私数据, 人们希望在保护数据隐私性的条件下对这些数据进行排序, 因此需要设计能够保护隐私的安全排序方案.

在安全多方计算(secure multi-party computation, SMC)中, 若干个参与者利用各自的保密数据进行联合计算. 计算完成后, 每个参与者仅能获得规定的输出结果, 而无法获得其他任何额外信息. Yao[1]首先提出并研究了一个安全两方计算问题. 随后, Goldreich 等人对其进行了深入的研究[2,3]. 目前, 关于 SMC 的研究已经形成了较完整的理论体系, 从理论上证明了所有的 SMC 问题都是可解的, 并提出了通用的解决方案. 由于通用方案对于大多数具体问题并不实用, 人们即针对各种实际问题研究设计高效的具体解决方案, 目前已经取得了很多较好的研究成果[4−12].

保密排序是 SMC 的一个基本问题, 在科学决策、电子商务推荐、保密拍卖、保密排名以及数据库操作等方面具有广泛的应用, 因此对于保密排序问题的研究也得到了广泛关注. 近年来, 人们提出了一些保密排序协议[13−19]. 文献[14,15]主要应用秘密分享的方法设计协议, 文献[14]应用多项式分享方法设计了一个严格单调变换 T, 使得参与者 P_i 能将自己的秘密数据 x_i 保密变换为 $T(x_i)=x_i'$, 所有参与者最后得到结果 (P_i,x_i'), i=1,…,n. 单调变换保证了由 x_i' 无法求逆得到相应的 x_i, 因此参与者能够在保护输入数据 x_i 隐私性的条件下获得所有数据的排序结果, 但该协议会泄露其他参与者数据的排序. 文献[15]以已有的比较协议、位分解以及相等测试等协议为基础设计了几类保密排序协议, 解决了数据范围有全集限制及无全集限制情形下的几类排序问题. 由于所设计的排序协议需要多次调用已有的基础协议, 协议复杂性很高. 文献[14,15]中的协议是信息论安全的, 应用秘密分享方法设计协议需要参与者之间有安全的通信信道.

文献[16–19]在有全集限制条件下, 以同态加密算法为基础研究了各种排序问题, 设计具有计算安全性的排序协议, 这些协议不需要参与者之间有安全通信信道. 文献[16]中的协议 1 和协议 2 主要以 Paillier 加密算法为基础设计构造, 由于只有 P_1 拥有私钥可以独立解密, 协议 1 设计简单但无法抵抗有 P_1 参与的合谋攻击; 协议 2 在协议 1 的基础上应用秘密分割的方法以增强抵抗合谋攻击能力, 但这又增加了协议的通信复杂性. 文献[17]利用 RSA 加密算法并借助半可信的第三方和不经意传输设计协议, 文献[18]应用具有加法同态性的加密算法研究了与文献[17]同样的问题, 避免借助第三方和不经意传输, 但这两个协议的通信复杂性都比较高. 文献[16]中的协议 3 和文献[19]的协议主要以椭圆曲线加密算法为基础进行设计, 两者研究的问题有部分类似, 后者的计算效率较高.

目前, 关于排序问题主要是在数据范围有全集限制条件下进行研究的. 关于在无全集限制情形下的排序问题, 文献[15]以秘密分享方法为基础设计了一个信息论安全的解决方案, 本文主要研究解决计算安全相关问题的解决方案.. 我们首先在数据范围有全集限制条件下设计了几类高效排序协议, 并以此为基础进一步构造了数据范围无全集限制情形下的解决方案. 应用模拟范例严格证明了所设计协议在半诚实模型下的安全性. 本文协议 1 和协议 2 所研究的问题与文献[16–19]的工作相关, 在效率分析部分, 将与这些工作进行详细的分析比较. 下面我们首先描述本文所研究的排序问题.

- 问题描述

假设 n 个参与者 $P_1,\ldots,P_n$ 分别拥有私密数据 $x_1,\ldots,x_n$, 他们希望对联合序列 X=[$x_1,\ldots,x_n$]中的元素排序得到一个由小到大排序的序列 T, 最后, P_i 仅知道数据 x_i 在序列 T 中的位置, 得不到其他参与方数据与其排序位置的任何信息. 本文将 x 在序列 T 中的位置序号称为 x 的排序位次.

当 X 中有相同数据时, 相同数据排序位次可以有不同的定义方式. 本文针对实际中典型的排序应用问题, 给出两种不同类型的位次计算方式, 并对所定义的排序方式设计 SMC 协议. 关于两种不同类型的排序位次计

算方式, 分别描述如下.

(i) 同数据并列计位法: 对于序列中出现的若干个相同数据(比如某数据 x 出现 k 次), 要求这 k 个 x 保持同一位次, 但下一个较大数据的排序位次要相应地增加 k 位.

(ii) 同数据增位次计位法: 对于序列中出现的若干个相同数据(比如某数据 x 出现 k 次), 要求这 k 个 x 排序位次严格递增, 共占有 k 个不同位次. 如此, 如果序列 X 中共有 n 个元素, 那么排序后序列 T 中所有数据排序位次从 1 开始严格递增, 最后一个数据的位次为 n.

下文中将应用计数排序法的基本思想, 结合适当的编码方法和修改后的 ElGamal 门限密码体制设计保密排序协议, 并应用所设计的保密排序思想设计一个安全公平的招投标协议. 本文的贡献如下:

(1) 首先, 应用计数排序的思想, 并灵活运用编码技巧设计了数据范围在有全集限制条件下的保密排序协议(协议 1 和协议 2), 其中, 增位次排序方案具有稳定性. 进一步设计了基于关键词的保密排序协议(协议 3), 解决了在初始位次为参与者隐私数据情形下的增位次计位排序问题, 基于关键词的排序算法具有稳定性. 严格论证了协议的安全性以及抵抗合谋攻击的能力.

(2) 以协议 1(或协议 2、协议 3)为基础, 根据基数排序的基本思想, 设计了数据范围未知情形下的并列位次(增位次)计位方式下的保密排序方案(方案 1、方案 2), 方案能够完全保护数据的隐私. 根据 SMC 组合定理, 证明了排序方案在半诚实模型下是安全的.

(3) 上述排序协议的设计思想能够应用于解决各种实际问题. 本文以上述协议为基础设计了 Vickrey 保密招投标协议(协议 4). 由于所提的问题本身对于招标者及投标者的数据隐私性保护要求很高, 所设计的协议具有较高的安全性及公平性.

(4) 由于计数排序是在数据范围较小时比较好的排序算法, 本文所设计的协议简单、高效. 本文协议保证每个参与者只知道自己的排序位次, 对于其他参与者的数据和位次完全保密. 和已有成果相比, 本文的解决方案能够更全面地保护数据隐私. 本文协议可以推广应用于更复杂的数据情形, 比如每个参与者具有一个私密数据序列的情形.

1 预备知识

1.1 半诚实模型及安全性

- 半诚实模型[3]

在半诚实模型中, 要求所有参与者都是半诚实的, 即参与者在执行协议时能够忠实地履行协议, 但他们会完整记录协议执行中收到的信息, 在完成协议后希望能从这些信息中推断出其他参与者的一些私密信息. 此模型下通常由下面方法证明 SMC 协议的安全性(称为模拟范例).

n 个参与者 $P_1,\ldots,P_n$ 利用协议 Π 保密地计算函数 $f(\bar{x})=(f_1(\bar{x}),\ldots,f_n(\bar{x}))$, 其中, $\bar{x}=(x_1,\ldots,x_n)$, x_i 是 P_i 的保密输入数据, $f_i(\bar{x})$ 为其输出结果. 将 P_i 在协议执行中记录的信息序列表示为

$$view_i^{\Pi}(\bar{x})=(x_i,r_i,M_i^1,\ldots,M_i^t,f_i(\bar{x})).$$

其中, $M_i^1,\ldots,M_i^t$ 为 P_i 收到的所有信息. 对于给定的参与者子集 $I=\{P_{i_1},\ldots,P_{i_s}\}\subseteq\{P_1,\ldots,P_n\}$, 记:

$$f_I(\bar{x})=(f_{i_1}(\bar{x}),\ldots,f_{i_s}(\bar{x}))\text{ 以及 }view_I^{\Pi}(\bar{x})=(I,view_{i_1}^{\Pi}(\bar{x}),\ldots,view_{i_s}^{\Pi}(\bar{x})).$$

半诚实模型下协议的安全性定义如下[3].

定义 1. 在半诚实模型下, 如果对于任意子集 $I=\{P_{i_1},\ldots,P_{i_s}\}$ 都存在概率多项式时间算法 S, 使得下式成立:

$$\{S(I,(x_{i_1},\ldots,x_{i_s}),f_I(\bar{x}))\}_{\bar{x}\in(\{0,1\}^*)^n}\overset{c}{\equiv}\{view_I^{\Pi}(\bar{x})\}_{\bar{x}\in(\{0,1\}^*)^n}\tag{1}$$

其中, $\overset{c}{\equiv}$ 表示计算不可区分. 则称 Π 是 n 元函数 f 的一个保密计算协议.

1.2 计数排序法与基数排序法

计数排序的基本思想是: 对待排序列中的每个元素 x, 确定该序列中小于 x 的元素个数, 由此可确定 x 的

排序位次. 本文首先以计数排序思想为基础, 设计数据范围较小情况下稳定的保密排序协议(协议 1–协议 3).

基数排序的主要思路是: 首先, 将所有待排序整数(均为非负整数)统一为位数相同的整数, 位数较少的前面补 0; 其次, 从最低位开始, 每一位进行一次稳定排序. 所有数位排序完成以后, 整个序列就变成了一个有序序列. 在基数排序中, 同一数位的排序子程序要用稳定排序算法, 如此即可将上一轮的排序成果保留下来.

1.3 ElGamal加密系统

ElGamal 加密系统是一种公钥加密系统, 并具有乘法同态性[20]. 简述如下.

- 密钥生成

根据安全参数τ, 密钥生成算法生成一个素数 p(长度为τ比特), 选择 Z_p^*的一个生成元 g, 并选择随机数 $k\in Z_p^*$为私钥, 与其对应的公钥则为 $h=g^k \bmod p$.

- 加密

为加密消息 $m(m\in Z_p^*)$选择一个随机数 r, 计算:

$$(c_1,c_2)=E(m)=(g^r \bmod p, mh^r \bmod p).$$

- 解密

对于密文 $C=(c_1,c_2)$, 解密得明文为: $m=D(C)=c_2\cdot c_1^{-k} \bmod p$.

将 ElGamal 加密系统修改为 Lifted ElGamal, 可使其对原始数据具有加法同态性. Lifted ElGamal 密钥生成算法与 ElGamal 加密系统相同, 如果将其加密算法和解密算法分别记为 $\hat{E}$ 和 $\hat{D}$, 则有:

- 加密

为加密消息 $m(2^m<p)$选择一个随机数 r, 密文为

$$(c_1,c_2)=\hat{E}(m)=(g^r \bmod p, 2^m h^r \bmod p).$$

- 解密

对于密文(c_1,c_2), 解密得到:

$$m=\hat{D}(c_1,c_2)=\log_2[c_2\cdot c_1^{-k} \bmod p].$$

- 同态性质

假设

$$\hat{E}(m_1)=(g^{r_1} \bmod p, 2^{m_1}h^{r_1} \bmod p),$$
$$\hat{E}(m_2)=(g^{r_2} \bmod p, 2^{m_2}h^{r_2} \bmod p).$$

有下面加法同态性:

$$\hat{E}(m_1)\times\hat{E}(m_2)=(g^{r_1+r_2} \bmod p, 2^{m_1+m_2}h^{r_1+r_2} \bmod p)=\hat{E}(m_1+m_2).$$

ElGamal 密码系统及 Lifted ElGamal 都是语义安全的[21], 即通过在加密过程中选择不同的随机数, 同一明文可以加密成多个密文形式, 并且所有密文都是计算不可区分的.

注解 1. 在 Lifted ElGamal 系统中, 加密明文 m 实际是在 ElGamal 系统中加密 2^m (要求 $2^m<p$), 这样做的目的是将 ElGamal 加密系统的乘法同态性转化为加法同态性. 应用 ElGamal 系统直接解密时仅得到 2^m, 如要进行完全解密, 需要再做一次对数运算. 在加密中, 明文 m 要满足条件 $2^m<p$. 对于具体的实际问题, 根据明文 m 的取值范围, 通过适当选择充分大的素数 p, 这个条件能够满足.

1.4 门限密码体制

门限密码体制[22,23]是 SMC 中对抗合谋攻击的一个重要手段. 在(t,n)门限密码体制中, 任何人都可以用公钥加密消息, 但至少需要 t 个人合作才能解密, 少于 t 个人时无法得到明文的任何信息. 本文将应用 Lifted ElGamal 构造(n,n)门限密码体制, 具体如下.

- 联合生成公钥

参与者 $P_1,\ldots,P_n$首先生成 ElGamal 密码系统的公共参数 p, g. 每个 P_i, $i\in[1,n]=\{1,\ldots,n\}$选取随机正整数 $sk_i<$

p, 计算并公布 $h_i=g^{sk_i} \bmod p$. 参与者联合生成公钥 pk: $h=g^{sk_1+\ldots+sk_n} \bmod p$, 联合持有私钥 $sk=(sk_1+\ldots+sk_n) \bmod p$. sk_i 称为 P_i 的私钥份额.

- 加密

任何参与者都可以应用 pk 进行加密, 加密过程和 Lifted ElGamal 加密系统相同. 将对应的加密算法记为 E_{pk}, 即对于明文 $m(2^m<p)$, 加密后得到密文 $E_{pk}(m)$为

$$E_{pk}(m)=(u,v)=(g^r \bmod p, 2^m h^r \bmod p) \tag{2}$$

- 联合解密

为了解密 $C=(u,v)$, 每个参与者 P_i 计算 $w_i=u^{sk_i} \bmod p$ 并公布, 则

$$m \equiv \log_2\left[v\left(\prod_{i=1}^{n} w_i\right)^{-1} \bmod p\right] \tag{3}$$

下文中, 将联合解密运算记为 D_{sk}, 上式的解密运算即可写成 $m=D_{sk}(C)$.

与 ElGamal 密码系统类似, 由离散对数问题的困难性假设可保证上面所构造的门限密码体制是语义安全的[20,24], 该门限密码体制具有加法同态性, 本文主要应用该门限密码体制构造协议.

注解 2. 在 Lifted ElGamal 门限密码体制中, 可以灵活应用联合解密方式使得 $P_1,\ldots,P_n$ 中仅有一人获得解密结果. 比如所有参与者合作解密密文 $C=(u,v)$, 则可按下面方法进行: 所有参与者 P_i, $i\in[1,n]$, $i\neq j$ 分别计算 $w_i=u^{sk_i}$ 并将其公布, 参与者 P_j 计算 $w_j=u^{sk_j}$ 并再根据式(3)进行计算, 如此即可保证仅有 P_j 得到最终解密结果 $D_{sk}(C)$. 在下文中, 如果涉及应用 Lifted ElGamal 门限密码体制联合解密且仅有一方参与者能够获得最终解密结果时, 都按此做法理解. 并且如果没有特殊说明, 本文中对于密文所做的乘法运算及指数运算都是在模 p 的意义下进行的.

2 数据范围有全集限制时的保密排序

假设参与者 $P_1,\ldots,P_n$ 的数据 $x_1,\ldots,x_n\in Z=\{z_1,z_2,\ldots,z_m\}$, 其中, $z_1<z_2<\ldots<z_m$. 此时, 对于每个 $i\in[1,n]$, 存在 $j\in[1,m]$, 使得 $x_i=z_j$. 为叙述方便, 将 j 称为 x_i 在全集 Z 中对应的指标, 记为 $J(x_i)$.

2.1 同数据并列计位法保密排序

- 问题描述

参与者 $P_i(i=1,\ldots,n)$拥有私密数据 $x_i\in Z$, 他们要合作保密计算, 计算完成后, P_i 只能获得 x_i 在联合序列 $X=[x_1,\ldots,x_n]$中按并列计位法的排序位次.

下面以计数排序法的基本思想设计协议.

- 计算原理

(i) 每个参与者 P_i 首先根据 x_i 在 Z 中对应的指标 $J(x_i)$构造一个 m 维向量 U_i:

$$U_i=(u_{i1},u_{i2},\ldots,u_{im}) \tag{4}$$

其中, 当 $s<J(x_i)$时, $u_{is}=0$; 当 $J(x_i)\leqslant s\leqslant m$ 时, $u_{is}=1$. 这样, x_i 与 U_i 一一对应.

(ii) 将向量 U_i 作为矩阵 U 的第 i 行, 构成矩阵 $U=(u_{ij})_{n\times m}$.

(iii) 对于每一个 $i\in[1,n]$, P_i 按下面方式计算 x_i 在联合序列 X 中按并列计位法定义的排序位次 r_i:

$$r_i=\begin{cases}1, & \text{if } x_i=z_1\\ 1+\sum_{s=1}^{n} u_{s(J(x_i)-1)}, & \text{if } x_i>z_1\end{cases} \tag{5}$$

由上式可知, 如果参与者 P_i 的数据为 z_1, 排序位次显然是 1; 对于数据 $x_i>z_1$, 其排序位次是对矩阵 U 中第 $J(x_i)-1$ 列的所有元素求和, 对求和结果再加 1 即可.

例 1: 假设参与者 P_1, P_2, P_3 和 P_4 分别拥有私密数据 $x_1=2$, $x_2=3$, $x_3=5$ 和 $x_4=3$. 令 $Z=\{1,2,\ldots,6\}$, P_i, $i=1,2,3,4$ 各自按照编码方法(4)构造一个 6 维向量, 并将其合成为下面的 4×6 阶矩阵 U.

$$U=\begin{pmatrix}0&1&1&1&1&1\\0&0&1&1&1&1\\0&0&0&0&1&1\\0&0&1&1&1&1\end{pmatrix}.$$

根据矩阵 U, 各参与者按照公式(5)确定各个数据的位次如下: 由于 X 中有互异数据 2, 3 和 5, 无论哪个参与者持有这些数据, 其排序位次是相同的. 其中, 数据 2 的位次是 1, 其值等于矩阵 U 第 1 列元素求和得到 0, 再加 1 而得到; 数据 3 的位次是 2, 其值等于 U 第 2 列元素求和得到 1, 再加 1 而得到; 数据 5 的位次是 4, 其值等于 U 第 4 列元素求和得到 3, 再加 1 而得到. 由于 X 排序后得到 T=[2,3,3,5], 直接检验可知上述排序位次是正确的.

命题 1. 对于并列计位排序方式, 由公式(5)给出的位次计算方法是正确的.

证明: 先将联合序列 X 表示为重数形式: $\hat{X}=[(t_1,b_1),...,(t_l,b_l)]$, 其中, $t_1,\ldots,t_l$ 为 X 中的所有互异数据, 满足 $t_1<t_2<\ldots<t_l$; b_i 为 t_i 的重数($b_1+\ldots+b_l=n$).

(i) 根据 $\hat{X}$ 的构造方式, X 中各数据的位次计算如下: t_1 的位次为 $r_1=1$, t_2 的位次为 $r_2=b_1+1$, …, t_j 的位次为 $r_j=b_1+b_2+\ldots+b_{j-1}+1$, …, t_l 的位次为 $r_l=b_1+\ldots+b_{l-1}+1$.

(ii) 根据 U_i 和 U 的构造, 可计算 U 中各列元素之和如下: 从第 1 列到 t_1-1 列, 各列元素之和皆为 $s_1=0$; 从 t_1 列到 t_2-1 列, 各列元素之和皆为 $s_2=b_1$; …; 从 t_j 列到 $t_{j+1}-1$ 列, 各列元素之和皆为 $s_{j+1}=b_1+b_2+\ldots+b_j$; …; 最后, 从 t_{l-1} 列到 t_l-1 列, 各列元素之和皆为 $s_l=b_1+\ldots+b_{l-1}$.

对于每一个 $j\in[1,l]$, 比较情形(i)中的 r_j 和情形(ii)中的 s_j 得到 $r_j=s_j+1$. 由此可知, t_j 的排序位次恰等于 U 中第 t_j-1 列元素求和后再加 1. 由于对任意的 $i\in[1,n]$, 总存在 t_j 使得 $x_i=t_j$, 因此, x_i 的排序位次满足公式(5). □

以上述原理为基础, 设计保密计算协议如下.

协议 1. 并列计位法保密排序协议.

输入: $P_i(i=1,\ldots,n)$的私密数据 x_i.

输出: 对于每一个 $i\in[1,n]$, P_i 输出 x_i 在联合序列 $X=[x_1,\ldots,x_n]$中按并列计位法的排序位次 $r_i=f_i(x_1,\ldots,x_n)$(下面假设 $x_i\neq z_1$, 否则直接输出 $r_i=1$).

准备: 所有参与者合作生成 ElGamal 门限密码系统的公钥 pk, 并记 P_i 持有的私钥份额为 sk_i.

(1) 参与者 $P_i(i\in[1,n])$操作如下.

(a) P_i 根据 x_i 按照公式(4)构造 m 维向量 $U_i=(u_{i1},\ldots,u_{im})$.

(b) P_i 加密向量 U_i 的每个分量, 得到密文向量 C_i 并公布:

$$C_i=(E_{pk}(u_{i1}),\ldots,E_{pk}(u_{im})):=(c_{i1},\ldots,c_{im}) \tag{6}$$

(2) 将所有 C_i, $i\in[1,n]$合成矩阵 M:

$$M=\begin{pmatrix}c_{11}&c_{12}&\ldots&c_{1m}\\c_{21}&c_{22}&\ldots&c_{2m}\\\ldots&\ldots&\ldots&\ldots\\c_{n1}&c_{n2}&\ldots&c_{nm}\end{pmatrix}.$$

(3) 对于每个 $i\in[1,n]$:

(a) 参与者 P_i 计算下面 H_i 并公布:

$$H_i=E_{pk}(1)\prod_{s=1}^{n}c_{s(J(x_i)-1)}:=(u_i,v_i) \tag{7}$$

(b) 所有参与者合作解密 H_i: 参与者 $P_j(j\neq i)$计算 $w_{ij}=u_i^{sk_j}$ 并公布, P_i 计算 $w_{ii}=u_i^{sk_i}$ 以及

$$h_i=\log_2\left[v_i\cdot\left(\prod_{j=1}^{n}w_{ij}\right)^{-1}\bmod p\right],$$

P_i 获得最终解密结果 $D_{sk}(H_i)=h_i$.

- 协议 1 的正确性

定理 1. 协议 1 能够正确计算并列计位法的排序位次.

证明: 根据命题 1, 当 $x_i \neq z_1$ 时, 只需证明 $h_i = 1 + \sum_{s=1}^{n} u_{s(J(x_i)-1)}$ 即可.

首先, 在协议第(1)步, 每个 P_i 将自己的数据 x_i 按照公式(4)构造对应的 m 维向量 U_i 并加密得到 C_i. 协议第(2)步是参与者根据计算原理得到 U 的密文矩阵 M. 第(3)步根据给定的 $i \in [1,n]$, 计算 x_i 在联合序列 X 中的排序位次. 在步骤(3)(a)中, P_i 计算了 M 的第 $J(x_i)-1$ 列所有元素以及密文 $E_{pk}(1)$的乘积 H_i, 由加密算法的加法同态性可知, H_i 即为 r_i 的密文; 在步骤(3)(b)中, 所有参与者应用 D_{sk} 合作解密 H_i, P_i 最终得到 $h_i = D_{sk}(H_i)$, 根据计算原理, $h_i = r_i$ 即为 x_i 在 X 中的排序位次. □

- 协议 1 的安全性.

为了证明协议的安全性, 对于任意合谋者集合 $I \subseteq \{P_1,\ldots,P_n\}$, 需要构造满足公式(1)的模拟器 S. 在协议 1 中, 所有参与者的地位本质上是平等的, 由于在协议中应用了(n,n)门限密码体制, 进一步可证明, 协议 1 能够抵抗任意的合谋攻击. 这意味着对于每一个 P_i, 其他参与者全部合谋也无法获得 P_i 私密数据 x_i 及其排序位次 r_i 的任何信息. 具体地, 有下面的结论.

定理 2. 在半诚实模型下协议 1 是安全的, 并能抵抗任意的合谋攻击.

证明: 按照文献[3]中的模拟范例证明定理 2, 需要对任意 $n-1$ 个参与者子集构造模拟器 S, 使 S 满足公式(1)(这是最大合谋集合, 如果协议对这个合谋集合是安全的, 则对其任何合谋者子集也是安全的). 由于在协议中各参与者地位平等, 仅考虑 $I=\{P_1,\ldots,P_{n-1}\}$, 他们合谋想获知 P_n 的私密数据 x_n 的相关信息. 模拟器 S 按如下方式运行: 首先构造 ElGamal 门限密码系统, 设其公、私钥为 pk'/sk', n 个私钥份额分别为 sk_i', $i \in [1,n]$.

(i) 接收到输入$(I,x_1,\ldots,x_{n-1},f_I(x_1,\ldots,x_n))$后, S 随机选取 x_n', 使得:

$$f_I(x_1,\ldots,x_{n-1},x_n') = f_I(x_1,\ldots,x_{n-1},x_n).$$

(ii) S 根据 x_n' 按照公式(4)构造 m 维向量 $U_n' = (u_{n1}',\ldots,u_{nm}')$.

(iii) S 加密向量 $U_1,\ldots,U_{n-1}$ 以及 U_n' 的每个分量, 得到密文向量 $C_1',\ldots,C_{n-1}'$ 以及 C_n':

$$C_i' = (c_{i1}',\ldots,c_{im}'),\ i \in [1,n] \tag{8}$$

S 进一步构造矩阵:

$$M' = \begin{pmatrix} c_{11}' & c_{12}' & \ldots & c_{1m}' \\ c_{21}' & c_{22}' & \ldots & c_{2m}' \\ \ldots & \ldots & \ldots & \ldots \\ c_{n1}' & c_{n2}' & \ldots & c_{nm}' \end{pmatrix}.$$

(iv) S 计算 $H_1',\ldots,H_n'$ 如下:

$$H_i' = E_{pk'}(1)\prod_{s=1}^{n} c_{s(J(x_i)-1)}' = (u_i', v_i'),\ i \in [1,n-1],$$

$$H_n' = E_{pk'}(1)\prod_{s=1}^{n} c_{s(J(x_n')-1)}' = (u_n', v_n').$$

(v) S 计算 $w_{ij}' = (u_i')^{sk_j'}$, $i,j \in [1,n]$ 以及 $h_i' = \log_2\left[v_i' \cdot \left(\prod_{j=1}^{n} w_{ij}'\right)^{-1} \bmod p\right]$, 得到:

$$f_i(x_1,\ldots,x_{n-1},x_n') = h_i',\ i \in [1,n].$$

在协议执行中,

$$view_I^{\Pi}(\bar{x}) = \{x_1,\ldots,x_{n-1},c_{n1},\ldots,c_{nm},H_n,w_{1n},\ldots,w_{(n-1)n},f_I(\bar{x})\}.$$

令

$$S(I,x_1,\ldots,x_{n-1},f_I(\bar{x})) = \{x_1,\ldots,x_{n-1},c_{n1}',\ldots,c_{nm}',H_n',w_{1n}',\ldots,w_{(n-1)n}',f_I(x_1,\ldots,x_{n-1},x_n')\}.$$

因为 ElGamal 加密系统是语义安全的, 所有参与者合作才能正确解密. 由于对每个 $j \in [1,m]$, c_{nj} 是 P_n 加密的密文, 对于 I 中的参与者而言, c_{nj} 与 c_{nj}' 是计算不可区分的. 而 $H_n = E_{pk}(1)\prod_{s=1}^{n} c_{s(J(x_n)-1)}$ 中的 $E_{pk}(1)$是由 P_n 加

密的密文, 因此对 I 中的参与者而言, H_n 与 H'_n 也是计算不可区分的. 对每个 $i\in[1,n-1]$, $w_{in}=u_i^{sk_n}$ 是 P_n 为解密 H_i 而发送给 P_i 的部分解密结果, 其中, sk_n 为 P_n 的私钥份额. 根据离散对数问题的困难性, I 中参与者根据 $w_{1n},\ldots,w_{(n-1)n}$ 无法获得 sk_n 的任何信息, 因而无法解密 w_{in} 与 w'_{in}, 在不能解密的情况下, 它们是计算不可区分的. 进一步地, 由于 $f_I(x_1,\ldots,x_{n-1},x'_n)=f_I(\vec{x})$, 故有: $\{S(I,x_1,\ldots,x_{n-1},f_I(\vec{x}))\}_{x_n\in Z}\overset{c}{\equiv}\{view_I^{\Pi}(\vec{x})\}_{x_n\in Z}$. □

2.2 同数据增位次计位法保密排序

- 问题描述

参与者 P_i, $i\in[1,n]$分别拥有私密数据 $x_i(x_1,\ldots,x_n\in Z)$, 每个 P_i 有一个公开的初始位次 s_i(这里, $(s_1,\ldots,s_n)$为$(1,\ldots,n)$的一个置换). 他们合作保密计算, 使得P_i获得x_i在联合序列$X=[x_1,\ldots,x_n]$中按增位次计位法的排序位次 t_i. 要求该排序方法具有稳定性, 即要求对于任意 $i,j\in[1,n]$, $t_i>t_j$ 当且仅当 $x_i>x_j$ 或 $x_i=x_j$ 且 $s_i>s_j$.

- 计算原理

由于在增位次排序方法中要求在联合序列中不同数据按大小顺序排列, 而且相同数据的位次也要依次递增, 为了实现这个目的, 按下面方法进行计算.

(i) 每个参与者首先将自己的保密数据 $x_i\in Z$ 按照下面方式构造一个 m 维行向量 $\hat{V}_i=(\hat{v}_{i1},\ldots,\hat{v}_{im})$, 其中,

$$\hat{v}_{ij}=\begin{cases}1, & \text{if } j=J(x_i)\\ 0, & \text{if } j\neq J(x_i)\end{cases} \tag{9}$$

(ii) 将所有参与者的向量合成为一个 $n\times m$ 阶矩阵, 合成方法是, $\hat{V}_i$ 位于合成矩阵的第 s_i 行. 将合成后的矩阵记为 $V=(v_{ij})_{n\times m}$.

(iii) 每一个 $P_i(i\in[1,n])$按下面方式计算, 即可得到其数据 x_i 按增位次方法的排序位次 t_i:

$$t_i=\sum_{j=1}^{J(x_i)-1}\sum_{s=1}^{n}v_{sj}+\sum_{s=1}^{s_i}v_{sJ(x_i)} \tag{10}$$

例 2: 参与者 P_1, P_2, P_3, P_4 分别拥有私密数据 2, 3, 5, 3, 其初始位次分别为 $s_1=2$, $s_2=1$, $s_3=4$, $s_4=3$, 令全集 $Z=\{1,2,\ldots,6\}$. 参与者 $P_1,\ldots,P_4$ 各自构造一个 6 维向量 $\hat{V}_1,\ldots,\hat{V}_4$, 并将其合成为 4×6 阶矩阵 V:

$$V=\begin{pmatrix}\hat{V}_2\\ \hat{V}_1\\ \hat{V}_4\\ \hat{V}_3\end{pmatrix}=\begin{pmatrix}0&0&1&0&0&0\\ 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&0&1&0\end{pmatrix}.$$

根据矩阵 V, 各参与者按照公式(10)确定其数据在排序后的位次 t_i 如下: P_1 的数据 $x_1=2$ 的位次 $t_1=1$, 其值为矩阵 V 的第 1 列及第 2 列第 1 行、第 2 行元素求和得到; P_2 的数据 $x_2=3$ 的位次是 $t_2=2$, 其值为矩阵 V 的第 1 列、第 2 列及第 3 列第 1 行元素求和得到; P_3 的数据 $x_3=5$ 的位次是 $t_3=4$, 其值等于矩阵 V 的第 1 列–第 5 列所有元素求和得到; P_4 的数据 $x_4=3$ 的位次是 $t_4=3$, 其值等于矩阵 V 的第 1 列、第 2 列及第 3 列第 1 行–第 3 行元素求和得到. 直接检验可知, 按照上述方法计算得到的位次计算结果是正确的.

命题 2. 对于增位次计位排序方式, 由公式(10)给出的位次计算方法是正确的.

证明: 根据矩阵 V 的构成方式, 对于每一个 $x\in Z$, V 的第 x 列所有 n 个元素之和是数据序列 $x_1,\ldots,x_n$ 中与 x 值相同的元素个数. 那么, V 的前 x 列元素之和即为序列 $x_1,\ldots,x_n$ 中小于等于 x 值的元素个数.

由此可知, 参与者 P_i 的数据 x_i 的位次计算公式应由两部分组成: 第 1 部分是 $x_1,\ldots,x_n$ 中小于 x_i 的元素个数, 即为 $\sum_{j=1}^{J(x_i)-1}\sum_{s=1}^{n}v_{sj}$; 第 2 部分是初始位次在 1 到 s_i 之间的参与者数据中, 其值等于 x_i 的元素个数, 即为 $\sum_{s=1}^{s_i}v_{sJ(x_i)}$. 因此, 计算公式(10)是正确的. □

命题 2 是增位次计位法计算参与者数据位次的计算原理, 保密排序协议设计如下.

协议 2. 增位次计位法保密排序协议.

输入: P_i, $i\in[1,n]$各自的私密数据 x_i 以及每个人公开的初始位次 s_i.

输出: 参与者 P_i, $i\in[1,n]$输出 x_i 在联合序列 X 中按增位次计位法的排序位次 $t_i=g_i(x_1,\ldots,x_n)$.

准备: 所有参与者合作生成 ElGamal 门限密码体制的公钥 pk, 参与者 P_i 具有的私钥份额为 sk_i.

(1) 每个参与者 $P_i(i\in[1,n])$操作如下.

(a) 按照公式(9)构造数据 x_i 对应的 m 维向量 $\hat{V}_i=(\hat{v}_{i1},\ldots,\hat{v}_{im})$.

(b) 加密 $\hat{V}_i$ 的每个分量, 得到下面的密文向量 C_i 并公布:

$$C_i=(E_{pk}(\hat{v}_{i1}),\ldots,E_{pk}(\hat{v}_{im})) \tag{11}$$

(2) 将所有 C_i, $i\in[1,n]$合成一个 $n\times m$ 阶密文矩阵, 合成方法是: 把向量 C_i 放在合成矩阵的第 s_i 行, 将合成的密文矩阵记为 $C=(c_{ij})_{n\times m}$.

(3) 对于每一个给定的 $i\in[1,n]$.

(a) 参与者 P_i 计算下面的 L_i 并公开:

$$L_i=\prod_{j=1}^{J(x_i)-1}\prod_{s=1}^{n}c_{sj}\times\prod_{s=1}^{s_i}c_{sJ(x_i)}E_{pk}(0)=(\alpha_i,\beta_i) \tag{12}$$

(b) 参与者合作解密 L_i: 参与者 P_j, $j\in[1,n]$, $j\neq i$ 计算 $\eta_{ij}=\alpha_i^{sk_j}$ 并公布, P_i 计算 $\eta_{ii}=\alpha_i^{sk_i}$ 以及

$$l_i=\log_2\left[\beta_i\cdot\left(\prod_{j=1}^{n}\eta_{ij}\right)^{-1}\bmod p\right].$$

P_i 获得最终解密结果 $D_{sk}(L_i)=l_i$.

关于协议 2 的正确性和安全性, 可类似应用定理 1 和定理 2 的方法进行证明, 仅叙述下面的定理, 证明省略.

定理 3. 在半诚实模型下协议 2 是安全的, 并能抵抗任意的合谋攻击.

3 基于关键词的保密排序问题

基于关键词的保密排序类似于第 2.2 节中的增位次保密排序, 这里, 每个参与者也具有初始位次, 不同的是: 此时初始位次 s_i 是参与者 P_i 的私密数据, 而在第 2.2 节中初始位次 s_i 为所有参与者的共享数据. 由于在下节研究无全集限制情形下增位次排序问题时需要应用基数排序的思想对各个数位依次排序, 要保证增位次排序方法的稳定性, 对较低数位的排序结果将要作为对较高数位排序时的“初始位次”, 而对每个数位得到的排序结果(即这些初始位次)显然是参与者的私密数据, 因此基于关键词的保密排序协议对无全集限制情形下设计增位次排序协议具有关键作用. 基于关键词保密排序问题的解决方案本身也具有独立的意义.

- 问题描述

假设参与者 P_i, $i\in[1,n]$分别拥有私密数据(x_i,s_i), 这里, x_i 属于全集 $Z=[0,m-1]$, 所有 $s_1,\ldots,s_n$ 互异, 均取自集合$\{1,\ldots,n\}$, 称 s_i 为 x_i 对应的关键词. 参与者希望通过合作保密计算, 使得 P_i 获知 x_i 在联合数据序列 $X=[x_1,\ldots,x_n]$中基于关键词 $S=[s_1,\ldots,s_n]$的排序位次 $t_i=T(x_i,s_i)$, 其中, $\{t_1,\ldots,t_n\}=[1,n]$, $t_i>t_j$ 当且仅当 $x_i>x_j$ 或 $x_i=x_j$ 且 $s_i>s_j$.

- 计算原理

由于在基于关键词的排序问题中, s_i 是 P_i 所拥有的私密数据, 无法像协议 2 那样由所有参与者合作构造保密矩阵 V 以解决问题. 对于这个问题需要设计全新的解决方案, 具体如下.

(i) 每个参与者 P_i 将其具有的私密数据(x_i,s_i)转变为 $y_i=\xi(x_i,s_i)=nx_i+s_i$, 如此得到的 $y_1,\ldots,y_n$ 属于全集$[1,nm]$. 下文中称此变换为ξ变换, 进一步的计算基于下面的结论:

命题 3: (a) 对(x_i,s_i)进行ξ变换后得到的$y_1,\ldots,y_n$互不相同; (b) 对变换后序列 $Y=[y_1,\ldots,y_n]$中元素由小到大进行排序, y_i 的排序位次即为 x_i 在序列 $X=[x_1,\ldots,x_n]$中基于关键词 $S=[s_1,\ldots,s_n]$的排序位次 $t_i=T(x_i,s_i)$.

证明:

(a) 显然, 对于任意 $i,j\in[1,n]$, $i\neq j$, 有 $y_i-y_j=n(x_i-x_j)+s_i-s_j$ 以及$-n<s_i-s_j<n$ 成立. 因此, $x_i>x_j$ 当且仅当 $y_i-y_j>0$;

$x_i<x_j$ 当且仅当 $y_i-y_j<0$; 当 $x_i=x_j$ 时, $s_i-s_j>0$ 当且仅当 $y_i-y_j>0$, $s_i-s_j<0$ 当且仅当 $y_i-y_j<0$. 故知 $y_1,\ldots,y_n$ 互不相同.

(b) 对 $Y=[y_1,\ldots,y_n]$中元素由小到大进行排序, 对于任意 $i\in[1,n]$, 将 y_i 的排序位次记为 t'_i. 显然, $\{t'_1,\ldots,t'_n\}=[1,n]$ 并且满足: $t'_i>t'_j$ 当且仅当 $x_i>x_j$ 或 $x_i=x_j$ 且 $s_i>s_j$. 由 $t_i=T(x_i,s_i)$的定义可知, $t'_i=t_i$, $i\in[1,n]$. □

(ii) 由于转换后的数据 $y_1,\ldots,y_n$ 互异, 排序问题容易解决, 具体方案如下:

参与者 P_i 将自己的数据 y_i 编码成一个 nm 维向量:

$$W_i=(w_{i1},\ldots,w_{i(nm)}) \tag{13}$$

其中, 当 $j<y_i$ 时, $w_{ij}=0$; 当 $y_i\leqslant j\leqslant nm$ 时, $w_{ij}=1$.

参与者将这 n 个向量相加, 将所得的和向量记为 $W=(w_1,\ldots,w_{nm})$, 则易知: 对每个 $i\in[1,n]$, y_i 在 $Y=[y_1,\ldots,y_n]$ 中由小到大的排序位次为 w_{y_i}. 因此, x_i 在序列 $X=[x_1,\ldots,x_n]$中基于关键词 $S=[s_1,\ldots,s_n]$的排序位次为

$$t_i=T(x_i,s_i)=w_{y_i}.$$

例 3: 参与者 P_1, P_2, P_3, P_4 分别拥有私密数据 $x_1=1$, $x_2=2$, $x_3=4$, $x_4=2$, 假设其对应的关键词分别为 $s_1=2$, $s_2=1$, $s_3=4$, $s_4=3$. 令 $m=5$, P_i ($i=1,2,3,4$)应用ξ变换 $y_i=4x_i+s_i$ 得到 $y_1=6$, $y_2=9$, $y_3=20$, $y_4=11$. 再根据 y_i 构造相应的 20 ($nm=20$)维向量 W_i, 见表 1.

表 1 各参与者构造的向量

	1	…	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
W_1=(	0	…	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	)
W_{12}=(	0	…	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	)
W_3=(	0	…	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	)
W_4=(	0	…	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	)
W=(	0	…	0	1	1	1	2	2	3	3	3	3	3	3	3	3	3	4	)

根据向量 W, 各参与者按照计算原理确定其数据 x_i 在联合序列中按照相应关键词的排序位次如下: P_1 的数据 1 的位次是 $w_6=1$, P_2 的数据 2 的位次是 $w_9=2$, P_3 的数据 4 位次是 $w_{20}=4$, P_4 的数据 2 位次是 $w_{11}=3$. 直接检验可知, 这样得到的排序位次是正确的.

协议 3. 基于关键词的保密排序协议.

输入: P_i, $i\in[1,n]$各自的私密数据 $\hat{x}_i=(x_i,s_i)$.

输出: P_i, $i\in[1,n]$输出 x_i 在联合序列 X 中基于关键词 S 的排序位次 $t_i=T(\hat{x}_i)$.

准备: 参与者 P_i, $i\in[1,n]$首先合作生成 ElGamal 门限密码体制的公钥 pk 以及联合解密密钥 sk, P_i 持有的私钥份额为 sk_i; P_i 计算 $y_i=nx_i+s_i$.

(1) 参与者 P_i, $i\in[1,n]$将自己的数据 y_i 按照公式(13)构造 nm 维向量 $W_i=(w_{i1},\ldots,w_{i(nm)})$;

(2) P_i, $i\in[1,n]$加密向量 W_i 的每个分量, 得到密文向量 C_i 并公布:

$$C_i=(E_{pk}(w_{i1}),\ldots,E_{pk}(w_{i(nm)})):=(c_{i1},\ldots,c_{i(nm)}) \tag{14}$$

(3) 对于每个 $i\in[1,n]$:

(a) 参与者 P_i 把向量 $C_1,\ldots,C_n$ 的第 y_i 个分量相乘并重新随机化, 得到:

$$R_i=\prod_{k=1}^{n}c_{ky_i}E_{pk}(0):=(A_i,B_i) \tag{15}$$

(b) 参与者 $P_j(j\neq i)$计算 $q_{ij}=A_i^{sk_j}$ 并公布. P_i 计算 $q_{ii}=A_i^{sk_i}$ 以及下面的 r_i, 并输出 r_i:

$$r_i=\log_2\left[B_i\cdot\left(\prod_{k=1}^{n}q_{ik}\right)^{-1}\bmod p\right].$$

- 协议 3 的正确性与安全性

定理 4. 协议 3 是正确的. 在半诚实模型下, 协议 3 是安全的, 并能抵抗任意的合谋攻击.

证明: 前述计算原理保证协议 3 能正确计算基于关键词的排序位次. 可以用模拟范例证明协议 3 的安全性, 证明过程与定理 2 的证明类似, 这里省略. □

注解 3. 协议 3 准备工作之后, 实际上是对 $P_1,\ldots,P_n$ 计算得到的相异数据序列 $y_1,\ldots,y_n$ 进行保密排序, 得到的结果 r_i 也是 y_i 在序列 $y_1,\ldots,y_n$ 中由小到大的排序位次. 按此方式理解, 协议 3 也可看成是对相异数据序列 $y_1,\ldots, y_n$ 的排序协议. 协议 1 和协议 2 显然也适合相异数据序列的排序: 如果事先已知待排序的数据互异, 协议 1 和协议 3 本质上相同, 此时协议 2(对于任何初始位次)的排序结果也和协议 3 相同.

为叙述简单, 本节协议 3 中假设全集为 $Z=[0,m-1]$的形式, 对于一般形式的全集, 将协议计算原理中向量 W_i 的编码方式进行适当修改即可类似讨论, 不再赘述. 协议 1–协议 3 也可以推广应用于更复杂的数据情形, 比如每个参与者 P_i 具有一个私密数据序列 $x_i^{(1)},...,x_i^{(k_i)}$, 此时 P_i 像在协议 1–协议 3 中计算 x_i 的位次一样, 类似地计算每个 $x_i^{(j)}(j=1,...,k_i)$ 的位次即可.

4 数据范围无全集限制时的排序方案

假设有 n 个参与者 P_i, $i\in[1,n]$, P_i 具有私密数据 x_i, 在保护数据隐私性的前提下, 要求所有参与者能够确定一个适当的全集 Z, 使得所有数据属于全集 Z 有时很困难. 本节将研究在无全集限制条件下的排序问题, 这里仅需假设所有参与者能够商议确定一个正整数 l, 使得数据 x_i 的十进制(或二进制)位数不超过 l 即可. 如此, 每个参与者都把自己的私密数据通过高位补 0 的方式扩充为 l 位, 并假设参与者 P_i 的数据已表示为 $x_i=x_{il}\ldots x_{i2}x_{i1}$, 其中, 对于每一个 $i\in[1,n]$, $k\in[1,l]$, $x_{ik}\in\{0,1,2,\ldots,9\}$(或 $x_{ik}\in\{0,1\}$), 并将 x_{ik} 称为 P_i 的第 k 位数据. 下面分别对于同数据并列计位和增位次计位两种排序方式设计保密计算协议, 对这些 l 位数据实现从小到大的保密排序, 并要求增位次计位法排序方案具有稳定性. 在下文中, 参与者 P 对于数组(z,t)所做的ξ变换定义为$\xi(z,t)=nz+t$.

针对同数据并列(或增位次)计位方式设计保密排序方案时, 需要多次调用协议 1(或协议 3)以实现对于每个数位的保密排序.

4.1 同数据并列计位法保密排序方案(无全集限制)

方案 1. 计算的基本原理是: 对于参与者输入的 n 个 l 位数据, 从低位到高位逐位进行排序, 在保持较低位排序成果的基础上, 进一步对高位数据进行排序. 计算方案简述如下.

(1) $P_1,\ldots,P_n$ 分别以 $x_{11},\ldots,x_{n1}$ 为输入调用协议 1, $P_i(i\in[1,n])$得到自己对于 x_{i1} 的秘密排序位次 s_{i1}.

(2) 令 $j=2$:

(a) 对每一个 $i\in[1,n]$, P_i 计算 $y_{ij}=\xi(x_{ij},s_{i(j-1)})$.

(b) $P_1,\ldots,P_n$ 分别以 $y_{1j},\ldots,y_{nj}$ 为输入调用协议 1, 将 $P_i(i\in[1,n])$所获得的 y_{ij} 的排序位次记为 s_{ij}.

(c) 当 $j<l$ 时, 返回步骤(2), 此时令 $j\leftarrow l+1$.

(3) 对每一个 $i\in[1,n]$, P_i 最终获得排序结果 s_{il}. s_{il} 即为 x_i 在 $X=[x_1,\ldots,x_n]$中按照同数据并列计位法所得的排序位次.

- 方案 1 的正确性与安全性

定理 5. 上述并列计位法保密排序方案 1 是正确的, 并在半诚实模型下是安全的.

证明:

(i) 方案 1 的正确性

在上面排序过程中, 当 $j=1$ 时, 首先对个位数应用协议 1 进行排序, 个位数相同者排序位次相同;

在个位数位次确定的基础上, 令 $j=2$, 此时参与者 $P_i(i\in[1,n])$对其私密数据的十位数 x_{i2} 与个位数排序结果 s_{i1} 进行ξ变换得到 y_{i2}. 所有参与者再对相异数据 $y_{12},\ldots,y_{n2}$ 调用协议 1 进行排序, 结果是十位数小的排在前; 十位数相同的情况下, 个位数小的排在前; 末两位均相同的情况下, 末两位的排序位次相同.

对于百位数的排序方法类似, 此时令 $j=3$, 在末两位排序位次确定的基础上, 对于百位数与前面得到的末两位排序位次 s_{i2} 进行ξ变换得到新的 y_{i3}, 再对相异数据 $y_{13},\ldots,y_{n3}$ 调用协议 1 进行排序, 结果是百位数小的排在前; 百位数相同的情况下, 末两位小的排在前; 末三位均相同的情况下, 末三位的排序位次相同.

如此做法一直进行到最高位, 这时所有参与者对于 $l-1$ 位数据 $z_i=x_{i(l-1)}\ldots x_{i2}x_{i1}$, $i\in[1,n]$已经完成排序, z_i 的

排序结果为 $s_{i(l-1)}$. 最后令 $j=l$, 对于最高位与已有排序结果 $s_{i(l-1)}$进行ξ变换得到新的 y_{il}, 再对相异数据 $y_{1l},\ldots,y_{nl}$ 调用协议 1 进行排序, 结果是最高位小的数据排在前; 最高位数据相同的情况下, z_i 小的排在前; 而对于 $x_1,\ldots,x_n$ 中相同的数据, 排序位次亦相同. 根据上面的分析, 方案 1 能够正确计算并列计位方式下的排序位次.

(ii) 方案 1 的安全性

方案是对于 n 个 l 位的输入数据从低位到高位逐位排序, 对于每一数位的排序是调用协议 1 完成的. 由于协议 1 在半诚实模型下是安全的, 根据 SMC 的组合定理[3], 保密排序方案 1 是安全的. □

4.2 同数据增位次计位法保密排序方案(无全集限制)

方案 2. 假设每个参与者 P_i 具有数据 $x_i=x_{il}\ldots x_{i2}x_{i1}$, 并且每个数据 x_i 有一个公开的初始位次 s_{i0}. 对于 $x_1,\ldots,x_n$ 按照同数据增位次计位法排序, 要求排序结果对于初始位次具有稳定性, 即如果在 $x_1,\ldots,x_n$ 中有数据相等的情形发生, 则初始位次较小者的最终排序位次也较小. 计算过程简述如下.

(1) 所有参与者对他们的第 1 位数据 $x_{11},\ldots,x_{n1}$ 按初始位次$(s_{10},\ldots,s_{n0})$调用协议 2 进行排序, 每个参与者 P_i 得到自己对于 x_{i1} 的秘密排序位次 t_{i1}.

(2) 参与者 $P_i(i\in[1,n])$以(x_{i2},t_{i1})为输入调用协议 3, 得到第 2 位数据 x_{i2} 的秘密排序位次 t_{i2}.

(3) 类似于对 $x_{12},\ldots,x_{n2}$ 的排序方式, 所有参与者依次对 $x_{13},\ldots,x_{n3}$ 一直到 $x_{1l},\ldots,x_{nl}$ 进行排序, 最终 P_i 得到秘密排序结果 t_{il}. $t_{1l},\ldots,t_{nl}$ 即为 $x_1,\ldots,x_n$ 按照增位次方式的排序结果.

- 方案 2 的正确性与安全性

关于增位次计位法保密排序方案的正确性与安全性, 我们有下面定理 6. 定理 6 的证明类似于定理 5, 在此省略.

定理 6. 同数据增位次计位法保密排序方案 2 是正确的, 并在半诚实模型下是安全的.

下面给出一个具体实例, 解释说明在无全集限制情形下, 按照两种计位方式进行排序的实际操作过程.

例 4: 假设 $P_1,\ldots,P_5$ 分别具有数据 351, 251, 421, 153, 251, 在增位次计位排法中, 设初始位次为 s_{10}=1, s_{20}=4, s_{30}=5, s_{40}=3, s_{50}=2. 对这组数据按照上述两种排序方法进行逐位排序. 为叙述简单, 对第 $j(j$=1,2,3)个数位的排序简称为第 j 轮, 将同数据并列计位法以及增位次计位法分别简称为并列法以及增位法. 排序结果见表 2.

表 2 两种不同排序方法的排序结果

参与者	待排数据	初始位次	第1轮		第2轮		第3轮	
			并列法	增位法	并列法	增位法	并列法	增位法
P_1	351	1	1	1	2	2	4	4
P_2	251	4	1	3	2	4	2	3
P_3	421	5	1	4	1	1	5	5
P_4	153	3	5	5	5	5	1	1
P_5	251	2	1	2	2	3	2	2

在表 2 中,

- 第 1 轮并列法排序结果是对数据[1,1,1,3,1]应用并列计位法得到的排序结果, 第 1 轮增位法排序结果是对数据[1,1,1,3,1]以[1,4,5,3,2]为初始位次应用增位次计位法得到的排序结果.
- 第 2 轮并列法排序结果是对数据[25+1,25+1,10+1,25+5,25+1]应用并列计位法得到的排序结果, 第 2 轮增位次排序结果是对于(互异的)数据[25+1,25+3,10+4,25+5,25+2]的排序结果.
- 第 3 轮并列法排序结果是对数据[15+2,10+2,20+1,5+5,10+2]应用并列计位法得到的排序结果, 第 3 轮增位次排序结果是对于(互异的)数据[15+2,10+4,20+1,5+5,10+3]的排序结果.

5 保密拍卖与招投标

目前广泛应用的拍卖方式基本分为 4 种: (i) 英国式拍卖, 也称“增价拍卖”; (ii) 荷兰式拍卖, 也称“降价拍卖”; (iii) 第一价格密封报价拍卖, 由投标价最高者赢得拍卖, 成交价格即为其所投标价; (iv) 第二价格密封

报价拍卖, 又称 Vickrey 拍卖, 由投标价最高者赢得拍卖, 但成交价格是所有投标价中第二高的标价[25].

工程招标是拍卖理论的反向操作形式. 第二价格密封报价招标中, 投标人向招标人提交密封的投标书, 投标人中报价最低者赢得工程, 并以所有报价中第二低价的价格(未中标人中的最低报价)与中标人签订合同. 本节中, 我们以 Vickrey 招标拍卖理论为基础, 应用前面所构造的保密排序协议, 设计构造安全、高效的招投标保密计算协议, 使得招标投标过程执行更安全、更高效, 以及尽量公平公正.

5.1 问题描述与计算方案

- 问题描述

假设有 n 个投标人要对某项工程进行投标. 招标人按照投标者的资质、信誉给每位投标人分配一个编号, 此编号即为招标人的私密输入数据; 每个投标人各有自己对投标工程的报价, 此报价作为投标人的私密输入数据. 招标人和投标人根据他们的私密输入数据进行合作保密计算, 计算结果是招标人得到中标人身份(即最低报价人)以及成交价(次低价), 对于所有投标人的其他信息全部保密; 而每个投标人仅知道自己的报价在整个报价中的排序位次, 得不到招标人分配给各个投标人的编号和其他投标人报价的任何信息.

上述问题的解决, 可保证在招投标过程中对招标人的秘密编号以及各投标人私密报价的隐私性给予最大程度的保护. 下面给出解决问题的具体实施策略和保密计算方案.

- 具体实施策略和计算方案

(i) 招标投标过程一般分为技术评标和商务评标这两个阶段. 在技术评标过程中, 招标人(记为 S)根据既定的评价内容对每一投标人进行具体分析和评价, 估算出有统一基础的评定分值, 选出符合技术标准要求的若干投标人(假设有 n 个人, 记为 $P_i, i\in[1,n]$). 按照评定分值的大小顺序对评标人从小到大进行编号, 这个编号即为招标人 S 对投标人分配的秘密编号, 对投标人 P_i 的编号记为 s_i. 这个过程结束后, 进入商务评标阶段.

(ii) 在商务评标阶段, 符合投标要求的投标人秘密报出自己的投标价, 将 P_i 的投标价记为 x_i.

(iii) 以招标人 S 的私密编号 $s_1,\ldots,s_n$ 以及投标人 $P_i(i\in[1,n])$的私密报价 $x_i(i\in[1,n])$为输入进行保密计算. 计算结果得到报价最低者的编号以及所有报价中的次低价. 假设所有 n 个投标报价从小到大排列为 $x_{i_1}\leqslant x_{i_2}\leqslant\ldots\leqslant x_{i_n}$, 如果 $x_{i_1}<x_{i_2}$, 此时 P_{i_1} 中标, 且成交价为 x_{i_2}; 如果有 $x_{i_1}=x_{i_2}=\ldots=x_{i_k}:=x_{\min}$ 的情形发生(即前 k 个报价相同), 此时规定 $P_{i_1},P_{i_2},\ldots,P_{i_k}$ 中对应编号最小的投标人中标(即有多个报价均为最低价时, 规定技术水平最优的投标人中标), 此时成交价应为 $x_{\min}$(即为未中标人中的最低报价).

(iv) 在商务评标过程中, 招标人可要求投标人仅能选取全集 $Z=\{z_1,z_2,\ldots,z_m\}$(满足 $z_1<z_2<\ldots<z_m$)中的某个数值作为投标价. 这样的要求在很多应用场景中是合理的: Z 中的最小值 z_1 可理解为招标商的保留价, 它是保证工程质量的最低成本价(类似于荷兰式拍卖中的起叫价); 最大值 z_m 可理解为招标商对于该工程的底价, 将高于该价格的成交价视为无效(类似于英式拍卖中的起叫价); 若记 $z_i-z_{i-1}=d_i$, 如果对于所有 $i=2,\ldots,m$, d_i 均相同(记为 d), 这时 d 可理解为英式(荷兰式)拍卖中的最低增幅(降幅).

在下面协议中, 假设招标人 S 已根据所有投标人的技术评价得分高低从 1 到 n 进行秘密编号, 对 P_i 的秘密编号记为 s_i, 该编号的作用是: 当有多个投标人均报出相同的最低价时, 此时要保证编号最小者(即技术评价最优者)中标. 招标人事先确定一个全集 $Z=\{z_1,z_2,\ldots,z_m\}$, 满足 $z_1<z_2<\ldots<z_m$. 根据计算方案, 输入数据 $s_1,\ldots,s_n$ 的作用类似于协议 2 中各个参与者的初始位次. 在协议 2 中, 初始位次 $s_1,\ldots,s_n$ 是所有参与者的共享数据, 而招投标问题中的编号 $s_1,\ldots,s_n$ 为招标人 S 的私密数据. 除此之外, 这里的招投标问题除了要保密计算各投标人报价高低的排序位次外, 还要保密计算中标者编号以及中标价. 若要解决这个问题, 需以协议 2 为基础, 并对其进行适当的修改和补充.

5.2 保密计算协议

协议 4. 保密招投标协议.

输入: S 输入对于 $P_1,\ldots,P_n$ 的私密编号$[s_1,\ldots,s_n]$; $P_i(i\in[1,n])$输入私密投标报价 $x_i\in Z$.

输出: 每个 P_i 输出 x_i 在报价序列$[x_1,\ldots,x_n]$中按增位次计位法的排序位次; S 输出中标人编号以及次低价.

准备: 招标人 S 和投标人 $P_1,\ldots,P_n$ 合作生成 ElGamal 门限密码体制的公钥 pk. 共享联合解密密钥 sk.

(1) 每个投标人 $P_i(i\in[1,n])$按照协议 2 的方法构造 m 维向量 $\hat{V}_i$, 并将 $\hat{V}_i$ 逐分量加密得到向量 C_i.

(2) 所有投标人由 $C_1,\ldots,C_n$ 合成一个矩阵 C, 合成方法是, C_i 位于矩阵 C 的第 i 行. 将合成后的矩阵记为 $C=(c_{ij})_{n\times m}$.

(3) 对每个 $i\in[1,n]$:

(a) P_i 对矩阵 C 的第 x_i 列元素进行重加密, 得到一个 n 维的列向量 $W_i=(w_{1x_i},...,w_{nx_i})^{\mathrm{T}}$, 并将 W_i 发送给 S.

(b) S 计算 $E_{pk}(0)$以及

$$Z_i=\prod_{s_k\leqslant s_i} w_{kx_i}E_{pk}(0) \tag{16}$$

并将 Z_i 发送给 P_i.

(c) P_i 计算 $E_{pk}(0)$以及

$$R_i=E_{pk}(0)Z_i\prod_{j=1}^{x_i-1}\prod_{k=1}^{n}c_{kj} \tag{17}$$

(d) 所有投标人 $P_1,\ldots,P_n$ 以及 S 合作解密 R_i, P_i 得到最终解密结果 $r_i=D_{sk}(R_i)$;
假设 $r_{i_1}=1$, $r_{i_2}=2$;

(4) S 和所有投标人 $P_1,\ldots,P_n$ 合作计算 i_1, 具体如下:

(a) 投标人 P_{i_1} 计算 $E_{pk}(i_1)$, 其他投标人计算 $E_{pk}(0)$, 并计算 n 个密文的乘积, 乘积结果记为 H_1.

(b) S 和 $P_1,\ldots,P_n$ 合作解密 H_1, S 得到最终解密结果 $h_1=D_{sk}(H_1)$.

(5) S 和所有投标人 $P_1,\ldots,P_n$ 合作计算 x_{i_2}, 具体如下:

(a) 投标人 P_{i_2} 计算 $E_{pk}(x_{i_2})$, 其他投标人计算 $E_{pk}(0)$, 并计算 n 个密文的乘积, 乘积结果记为 H_2.

(b) S 和 $P_1,\ldots,P_n$ 合作解密 H_2, S 得到最终解密结果 $h_2=D_{sk}(H_2)$.

5.3 协议4的正确性与安全性

- 协议 4 的正确性

1) 协议 4 步骤(1)–步骤(3)是由招标人和所有投标人合作计算每个 x_i 在序列$[x_1,\ldots,x_n]$中的排序位次(以$[s_1,\ldots,s_n]$为初始位次, 按增位次计位方式计算). 根据公式(10), P_i 的数据 x_i 的位次计算公式应由两部分组成: 一部分是初始位次在 1 到 s_i 之间, 投标价等于 x_i 的投标人数目(满足条件 $s_k\leqslant s_i$ 的 k 的数目), 步骤(3)(b)中所计算的 Z_i 即为该数目的密文; 另一部分是投标价小于 x_i 的投标人数目, 公式(17)中的连乘积 $\prod_{j=1}^{x_i-1}\prod_{k=1}^{n}c_{kj}$ 即为该数目的密文. 根据加密算法的加法同态性, 解密结果 r_i 等于由公式(10)所计算的 t_i, 即 r_i 为 x_i 在序列 $X=[x_1,\ldots,x_n]$中以$[s_1,\ldots,s_n]$为初始位次按增位次计位法获得的排序位次. 进一步, 根据增位次计位法的定义, $\{r_1,\ldots,r_n\}=\{1,\ldots,n\}$. 因此存在 $i_1,i_2\in\{1,\ldots,n\}$, 使得 $r_{i_1}=1$, $r_{i_2}=2$. 这说明 P_{i_1} 和 P_{i_2} 的投标价分别为最低价和次低价.

2) 协议 4 步骤(4)(或步骤(5))是为合作计算 i_1(或 x_{i_2}), 即确定中标人身份(或中标价). 根据加密算法的加法同态性, 解密结果 $h_1=i_1$(或 $h_2=x_{i_2}$), 说明中标人为 P_{i_1} (或中标价为 x_{i_2}).

- 协议 4 的安全性

关于协议 4 的安全性, 需要对不同参与者数据的安全性分别进行分析, 具体如下.

(i) 招标人数据的安全性. 在整个协议中, 招标人 S 仅由公式(12)计算了 $Z_1,\ldots,Z_n$. 在计算每个 Z_i 的过程中, S 都乘了一个 $E_{pk}(0)$. 由于加密算法的语义安全性, 在解密前, 即使所有投标人合谋, Z_i 的值与随机数也是不可区分的. 在解密后, P_i 所得到的 r_i 是协议规定的输出结果. 因此在协议 4 中, S 的私密数据 $s_1,\ldots,s_n$ 是安全的.

(ii) 投标人数据的安全性. 中标人 P_{i_1} 在协议执行中加密 $\hat{V}_{i_1}$, i_1 和 0 得到密文 C_{i_1}, $E_{pk}(i_1)$ 和 $E_{pk}(0)$, 并计算

了密文 W_{i_1} 以及 R_{i_1}. 根据加密算法的语义安全性, 即使其他参与者全部合谋, 在解密前也得不到关于这些密文对应明文的任何信息. 所有参与者合作解密 R_{i_1} 和 H_1, P_{i_1} 和 S 分别得到 $r_{i_1}=1$ 和 i_1, 这是协议规定的输出结果. 因此在协议执行中, P_{i_1} 的数据是安全的.

关于投标价为次低价的投标人 P_{i_2} 数据的安全性, 完全类似于 P_{i_1} 数据的安全性分析, 可知其数据是安全的. 类似地, 其他投标人的数据也是安全的.

关于协议 4 的安全性有下面的定理 7, 应用模拟范例可对定理 7 进行严格证明, 在此省略.

定理 7. 在半诚实模型下协议 4 是安全的, 并能抵抗任意的合谋攻击.

6 效率分析

本文首先设计了在数据范围有全集限制条件下, 以同数据并列位次(或增位次)方式进行保密排序的协议 1(协议 2)以及基于关键词的保密排序协议 3. 进一步设计了数据范围无全集限制情况下两种排序方式的保密排序方案, 最后设计了保密的招投标协议. 数据范围无全集限制情况下的排序方案以及招投标协议主要都是以协议 1–协议 3 为基础进行设计, 因此在效率分析部分, 主要对协议 1–协议 3 进行分析讨论, 并假设在这 3 个协议中, 参与者人数均为 n, 全集的势均为 m.

- 计算效率分析

在分析计算复杂性时, 忽略协议执行中需要的乘法运算, 只考虑最费时的模指数运算, 应用 ElGamal 门限密码体制加密(解密)一次需要进行两次(n 次)模指数运算.

协议 1 和协议 2 的计算复杂性相同: 每个参与者首先需要逐分量加密一个 m 维向量(U_i 或 $\hat{V}_i$), 这部分需要进行 nm 次加密; 其次, 对每个 $i\in[1,n]$, P_i 加密一个 $E_{pk}(1)$或 $E_{pk}(0)$, 并进行密文乘积运算得到一个密文(H_i 或 L_i), 所有参与者再对得到的密文进行合作解密. 所以, 协议 1、协议 2 各需要 $2nm+n^2+2n$ 次模指数运算.

在协议 3 中, 每个参与者 P_i 首先加密一个 nm 维向量 W_i; 其次, P_i 加密一个 $E_{pk}(0)$, 并进行密文乘积运算得到一个密文 R_i, 所有参与者再对得到的密文进行合作解密. 所以, 协议 3 共需要 $2n^2m+n^2+2n$ 次模指数运算. 由于协议 3 主要是为应用于无全集限制时的增位次排序(方案 2)中各个数位的排序, 这时 m 的值是确定的(二[或十]进制时, m=2[或 m=10]), 此时, 协议 3 所需模指数运算次数是参与者人数的二次函数.

由于在协议 1–协议 3 中, 每个参与者加密的所有数据都是 0 或 1, 这些加密任务可以由各参与者提前独立完成(或由云完成), 由此, 协议 1–协议 3 的在线复杂性仅为参与者合作解密 n 个密文所需要的 n^2 次模指数运算.

- 通信效率分析

本文以协议执行所需要的通信次数衡量通信复杂性. 协议 1–协议 3 的通信复杂性相同: 对每个 $i\in[1,n]$, 参与者 P_i 首先需要将所加密的向量 C_i 公布; 其次, P_i 需要将自己所计算的密文(H_i, L_i 或 R_i)公布; 最后, 为了解密其他参与者 P_j 所计算的密文(H_j, L_j 或 R_j), P_i 需要公布相应的部分解密结果. 因此在协议 1–协议 3 中, 每个参与者需要公布 3 组信息, 协议的通信次数为 $3n$.

假设各参与方具有的数据为 l 位, 在方案 1 中需要调用 l 次协议 1; 在方案 2 中, 首先需要调用一次协议 2, 再调用 l–1 次协议 3. 由于大规模数据中各数位的取值范围较小(十进制时全集 Z={0,1,2,…,9}, 二进制时全集为 Z={0,1}), 由此可知, 方案 1 及方案 2 的计算复杂性和通信复杂性分别为 $O(n^2l)$和 $O(nl)$.

- 与已有相关结果的分析比较

本文协议 1、协议 2 分别设计了在有全集限制条件下的同数据并列计位(增位次计位)排序问题, 这两个协议研究的问题和研究方法与文献[16,17,19]关系密切, 因此将这两个协议与这些文献的工作进行详细的比较. 这些协议主要是应用 Paillier 加密系统、RSA 加密系统、椭圆曲线加密系统, 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协议计算复杂性的基准, 用 M_a 表示模加运算, M_e 表示模指数运算; 另外, 测评协议通信复杂性的基准统一为通信次数, 并统一用 m 表示全集的势, n 表示参与者人数, CF 表示计算功能, C_1 (C_2)表示计算(通信)复杂性, CD 表示是否能够完全解密. 分析结果见表 3.

表 3 协议效率分析和功能比较

	CF	C_1	C_2	*CD*
本文协议 1	并列排序	$2nm+n^2+2n[M_e]$	O(n)	是
本文协议 2	增序排序	$2nm+n^2+2n[M_e]$	O(n)	是
文献[16]协议 2	并列排序	$2n(m+2)[M_e]$	O(n^2)	是
文献[17]协议	并列排序	$3nm+2n(n+1)[M_e]$	O(n^2)	是
文献[19]协议 2	并列排序	$(nm+2n)\log r[M_a]$	O(n)	否
文献[19]协议 3	增序排序	$(nm+2n)\log r[M_a]$	O(n)	否

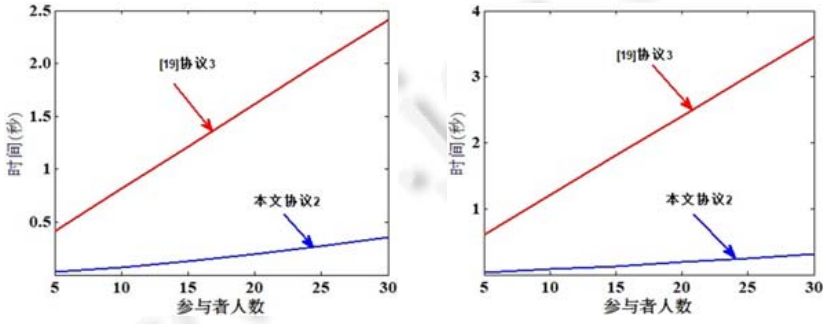
由表 3 可知: 本文协议 1 与文献[16]协议 2 相比计算复杂性差别不大, 但本文协议通信复杂性较低; 而与文献[17]的协议相比, 本文协议 1 计算复杂性和通信复杂性都比较低. 本文协议 1(协议 2)研究的问题和文献[19]的协议 2(协议 3)相同, 本文协议 1 和协议 2 在参与者仅有一个数据的情形下, 对文献[19]的相应结果进行了改进和补充, 经过修改后的协议, 能够在设计无全集限制情形下的保密计算方案时作为基础协议调用. 具体地,在文献[19]协议 2 的合作解密过程中, 要求所有参与者将其对密文 R_{ik} 的部分解密结果公布, 如此做法将可能使参与者 P_i, $j\neq i$ 获知 P_i 数据的排序位次. 在本文协议 1 中修改了参与者的合作解密方式, 以保证仅有 P_i 才能获得其私密数据的排序位次. 本文协议 2 与文献[19]中协议 3 的编码方式不同, 协议的设计原理也不同(本文是以计数排序思想为基础进行设计), 最大区别是, 文献[19]中没有考虑增位次排序中参与者的排序结果关于其初始位次的稳定性. 在本文协议 2 和协议 3 中均假设每个参与者具有初始位次, 要求排序结果对于初始位次具有稳定性, 这一点在后续设计无全集限制情形下的保密排序方案时非常重要.

- 实验测试

为了测试本文协议的实际执行效率以及与文献[19]的效率比较, 我们采用的编程语言是 PyCharm+Python 3.8, 测试环境是台式机, 处理器为 Intel(R) Core(TM) i5-9400 CPU @ 2.90 GHz, RAM 8.00 GB, Windows 10 64 位操作系统. 椭圆曲线密码系统选择 NIST 推荐的 P-256 椭圆曲线密码系统参数, 并采用 GitHub 网站上的 Python 程序. ElGamal 密码系统的参数选择如下: 系统参数 p 为 512 比特, 私钥和公钥长度均为 64 比特, ElGamal 程序是自己编写的程序. 所有数据都取 1 000 次测试的平均值.

由于本文协议 1 与协议 2 计算复杂性相同, 而文献[19]的协议 2 与协议 3 计算复杂性相同, 因此仅对本文协议 2 与文献[19]协议 3 进行测试比较. 对于本文协议 2 与文献[19]协议 3 的对比实验测试中, 设置 m=10, 30. 在每组测试中, 对两个协议选取相同的测试数据(满足 $x_i\in[1,m]$), 实验结果如图 1 所示. 由于本文协议 3 主要是为应用于无全集限制时的方案 2 中各个数位的排序, 此时 m 的值是确定的(二[或十]进制时 m=2[或 m=10]), 我们选取 m=2 和 m=10 进行实验测试, 实验结果如图 2 所示.

由图 1 可知, 当全集的势 m 固定时, 协议执行时间随参与者人数 n 的增加而增加; 而对于固定的 n, 随着 m 的增加, 协议执行时间也相应增加. 图 1 表明, 本文协议 2(协议 1)比文献[19]中的协议 3(协议 2)的效率高. 实验结果表明, 本文协议 1–协议 3 的计算效率较高.



(a) m=10 (b) m=30

图 1 当 m=10, 30 时, 协议执行时间

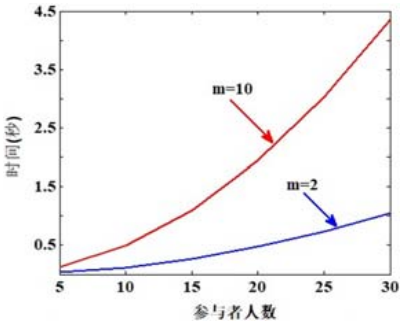


图 2 当 m=2, 10 时，协议 3 执行时间

理论上，协议的通信比特数即为协议需要传送的密文数乘以每个密文的平均长度(在我们的实验中平均长度为 256 比特). 在实验中，我们记录的通信比特数与理论分析结果大致相同.

7 结 论

本文首先设计了在数据范围有全集限制条件下，以同数据并列位次(或增位次)方式进行保密排序的协议 1(协议 2)以及基于关键词的保密排序协议 3. 效率分析和实验测试表明，本文的这些基本协议简单、高效. 以这些基本协议为基础，进一步研究了无全集限制情形下的安全排序问题，这是一类新的问题，目前还未见到有相关研究报道. 本文所设计的无全集限制情形下的计算方案由于要多次调用基本协议，协议效率较低. 在后续的工作中，我们将进一步深入思考，寻求优化方案，以期改进这些方案的计算效率；我们也将进一步研究恶意模型下，有关排序问题的安全计算.

References:

[1] Yao AC. Protocols for secure computations. In: Proc. of the 23rd Annual Symp. on Foundations of Computer Science. Los Alamitos: IEEE Computer Society, 1982. 160−164.

[2] Goldreich O, Micali S, Wigderson A. How to play any mental game. In: Alfred VA, ed. Proc. of the 19th Annual ACM Symp. on Theory of Computing. New York: ACM, 1987. 218−229.

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

[4] Lipmaa H, Toft T. Secure equality and greater-than tests with sublinear online complexity. In: Fedor VF, Freivalds R, Kwiatkowska M, *et al*., eds. Proc. of the Int'l Colloquium on Automata, Languages, and Programming. New York: Springer, 2013. 645−656.

[5] Li SD, Guo YM, Zhou SF, Dou JW, Wang DS. Efficient protocols for the general millionaires' problem. Chinese Journal of Electronics, 2017, 26(4): 696−702.

[6] Li SD, Wu CY, Wang DS, Dai YQ. Secure multiparty computation of solid geometric problems and their applications. Information Sciences, 2014, 282(1): 401−413.

[7] Liu L, Chen XF, Lou WJ. Secure three-party computational protocols for triangle area. Int'l Journal of Information Security, 2016, 15(1): 1−13.

[8] Lindell Y, Pinkas B. Privacy preserving data mining. Journal of Cryptology, 2008, 9(8): 616−621.

[9] Egert R, Fischlin M, Gens D, Jacob S, Senker M, Tillmanns J. Privately computing set-union and set-intersection cardinality via bloom filters. European Journal of Operational Research, 2015, 139(2): 371−389.

[10] Dou JW, Liu XH, Zhou SF, Li SD. Efficient secure multiparty set operations protocols and their application. Chinese Journal of Computers, 2018, 41(8): 1844−1860 (in Chinese with English abstract).

[11] Yuan J, Ye QS, Wang HX, Pieprzyk J. Secure computation of the vector dominance problem. In: Chen LQ, Mu Y, Susilo W, eds. Proc. of the Int'l Conf. on Information Security Practice and Experience. London: Springer, 2008. 319−333.

[12] Liu W, Luo SS, Wang YB. Secure two-party vector dominance statistic protocol and its applications. Acta Electronica Sinica, 2010, 38(11): 2573−2577 (in Chinese with English abstract).

[13] Hamada K, Kikuchi R, Dai I, Chida K, Takahashi K. Practically efficient multi-party sorting protocols from comparison sort algorithms. In: Kwon T, Lee MK, Kwon D, eds. Proc. of the Int'l Conf. on Information Security and Cryptology. Berlin: Springer, 2012. 202−216.

[14] Tang CM, Shi GH, Yao ZA. Secure multi-party computation protocol for sequencing problem. Science China Information Sciences, 2011, 54(8): 1654−1662.

[15] Zhang B. Generic constant-round oblivious sorting algorithm for MPC. In: Boyen X, Chen XF, eds. Proc. of the Int'l Conf. on Provable Security. New York: Springer, 2011. 240−256.

[16] Li SD, Kang J, Yang XY, Dou JW, Liu X. Secure multiparty characters sorting. Chinese Journal of Computers, 2018, 41(5): 206−222 (in Chinese with English abstract).

[17] Qiu M, Luo SS, Liu W, Chen P. A solution secure multi-party multi-data ranking problem based on RSA encryption scheme. Journal of Electronics, 2009, 37(5):1119−1123 (in Chinese with English abstract).

[18] Liu W, Luo SS, Wang YB, Jiang ZT. A protocol of secure multi-party multi-data ranking and its application in privacy preserving sequential pattern mining. In: Yu L, ed. Proc. of the 4th Int'l Joint Conf. on Computational Sciences and Optimization. Los Alamitos: IEEE Computer Society, 2011. 272−275.

[19] Li SD, Du RM, Yang YJ, Wei Q. Secure multiparty multi-data ranking. Chinese Journal of Computers, 2020, 43(8): 1448−1462 (in Chinese with English abstract).

[20] ElGamal T. A public key cryptosystem and a signature scheme based on discrete logarithms. In: Blakley GR, Chaum D, eds. Proc. of the Advances in Cryptology. New York: Springer, 1984. 10−18.

[21] Tsiounis Y, Yung M. On the security of ElGamal based encryption. In: Imai H, Zheng YL, eds. Proc. of the Public Key Cryptography. Berlin: Springer, 1998. 117−134.

[22] Zhang G, Qin J. Lattice-based threshold cryptography and its applications in distributed cloud computing. Int'l Journal of High Performance Computing and Networking, 2015, 8(2): 176−185.

[23] Long Y, Chen K, Mao X. New constructions of dynamic threshold cryptosystem. Journal of Shanghai Jiao Tong University (Science), 2014, 19(4): 431−435.

[24] Chaum D, Evertse JH, Van De Graaf J. An improved protocol for demonstrating possession of discrete logarithms and some generalizations. In: Chaum D, ed. Proc. of the Workshop on the Theory and Application of Cryptographic Techniques. Berlin: Springer, 1987. 127−141.

[25] Vickrey W. Counterspeculation, auctions, and competitive sealed tenders. The Journal of Finance, 1961, 16(1): 8−37.

附中文参考文献:

[10] 窦家维, 刘旭红, 周素芳, 李顺东. 高效的集合安全多方计算协议及应用. 计算机学报, 2018, 41(8): 1844−1860.

[12] 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.

[16] 李顺东, 亢佳, 杨晓艺, 窦家维, 刘新. 多个字符排序的安全多方计算. 计算机学报, 2018, 41(5): 206−222.

[17] 邱梅, 罗守山, 刘文, 陈萍. 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.

[19] 李顺东, 杜润萌, 杨颜璟, 魏琼. 安全多方多数据排序. 计算机学报, 2020, 43(8): 1448−1462.

窦家维(1963－), 女, 博士, 副教授, 主要研究领域为密码学, 信息安全.

汪榆淋(1997－), 女, 硕士, 主要研究领域为密码学, 信息安全.