保密替换及其在保密科学计算中的应用

杨晓艺 李顺东 亢 佳
(陕西师范大学计算机科学学院 西安 710062)

摘 要 安全多方计算是国际密码学界近年来的研究热点之一,也是网络社会隐私保护的关键技术.安全多方科学计算是安全多方计算的一个重要方面,最大(小)值的计算是一个基本的科学计算问题,具有重要的理论与实际意义.该文研究多个数据最大(小)值的保密计算问题.为解决此问题,该文首先利用概率加密算法的性质提出了保密替换的方法.其次,设计了一种新的编码方案,借助于保密替换、新的编码方案、概率加密以及门限解密密码系统,设计了三个最大(小)值保密计算协议.第一个协议可以用任何概率加密系统构造,使用中可以自由选择最高效的概率加密系统,适用于数据来自于一个小的稠密集;第二个方案应用类似的编码方案以及门限解密算法设计,可以抵抗任意合谋攻击,使用场合与第一个协议相同;第三个协议也能够抵抗任意合谋攻击,适用于保密数据来自于一个小的稀疏集.作为最大值问题的应用,该文进一步给出了多个保密数据的最小公倍数和最大公约数保密计算的解决方案并给出了最小公倍数的保密计算协议.最后应用模拟范例证明方案对于半诚实参与者是安全的,并给出了相应的效率分析与实验验证.

关键词 密码学;安全多方计算;概率加密;门限解密;最大(小)值;最小公倍数(最大公约数)
中图法分类号 TP309 **DOI 号** 10.11897/SP.J.1016.2018.01132

Private Substitution and Its Applications in Private Scientific Computation

YANG Xiao-Yi LI Shun-Dong KANG Jia
(*School of Computer Science, Shaanxi Normal University, Xi'an* 710062)

Abstract Secure multiparty computation is an important field of cryptography and a research focus in the international cryptographic community in recent years, and will become an integral part of computing science. It is a key privacy preserving technology in cooperative computation, cloud computing, electronic commerce, electronic voting, social activity etc. Secure multiparty scientific computation is an important field of secure multiparty computation, which studies how to preserve the privacy that may leak in cooperative scientific computation. Privately computing the maximum or the minimum of some private data, which naturally generalizes the famous millionaires' problem, is a basic problem in both scientific computation and secure multiparty computation. It is of great theoretical and practical significance, but has not been studied sufficiently. The existing solutions to this problem that use pair-wise comparison mechanism and have to invoke the protocols for millionaires' problems many times are either inefficient or insecure (will disclose more information). This paper studies how to privately compute the maximum or the minimum of some private data that uniformly distribute over some set with small cardinality. To solve this problem, we first propose a private substitution method which is based on the property of probabilistic encryption. Using this method, one can replace the ciphertexts of a data set to change or not to change the plaintexts of the data set but no party knows whether the data set is changed or not. Second, we design a new encoding scheme to simplify the private computation,

收稿日期:2016-12-06;在线出版日期:2017-11-20.本课题得到国家自然科学基金面上项目(61272435)资助.**杨晓艺**,女,1993 年生,博士研究生,主要研究方向为密码学与信息安全.E-mail:305965735@qq.com.**李顺东**(通信作者),男,1963 年生,博士,教授,中国计算机学会(CCF)会员,主要研究领域为密码学与信息安全.E-mail:shundong@snnu.edu.cn.**亢 佳**,女,1992 年生,硕士研究生,主要研究方向为密码学与信息安全.

which can reduce the computation on private data to the computation on some private vectors by encoding a private number to a vector. Using the private substitution method, the new encoding scheme, and a probabilistic encryption cryptosystem, we design our first protocol to privately compute the maximum or the minimum of some private data. This protocol can be constructed with any probabilistic cryptosystem so that one can choose the most efficient probabilistic encryption in practice, and is appropriate for the case in which the private data comes from a small dense set. The second protocol is designed using the private substitution, the new encoding scheme and a threshold decryption cryptosystem, it can resist any collusion attack, and is appropriate for the same case as the first protocol. The third protocol is constructed using the same building blocks as that of the second protocol, and is appropriate for the case in which the private data comes from a small sparse set. All the three protocols jump out the traditional pair-wise comparison mode to compute the maximum or the minimum, and therefore are efficient and secure. As application of these protocols, we discuss how to use these protocols to privately compute the greatest common divisor or the least common multiple of two private numbers, and give a complete protocol for privately compute the least common multiple. Finally, we prove that these protocols are secure in the semi-honest model using the simulation paradigm which is introduced by Goldreich and is well-accepted in secure multiparty computation research. We analyze the efficiency of the proposed protocols, and the result shows that our new protocols are efficient. We provide an experimental result to certify our efficiency analysis about the protocols.

Keywords　cryptography; secure multiparty computation; probabilistic encryption; threshold decryption; maximum (minimum) value; least common multiple (greatest common divisor)

1　引　言

在 20 世纪 80 年代 Yao[1] 以百万富翁问题引入了安全多方计算这一概念，经过 Goldreich、Micali、Wigderson[2] 等学者的发展，安全多方计算目前已经成为密码学中一个非常重要且活跃的研究领域，是国际密码学界近年来的研究热点. 安全多方计算是指 n 个参与者 $P_1, P_2, \cdots, P_n$ 合作在他们的私有数据 $x_1, x_2, \cdots, x_n$ 上计算函数 $f(x_1, x_2, \cdots, x_n)$，计算结束后所有参与者所能获得的关于其他参与者私有数据的信息不会超过从 $f(x_1, x_2, \cdots, x_n)$ 和自己的输入可以推断出的信息. 安全多方计算可以使得保密的私有数据被最大限度地利用又不泄露隐私信息，这正是其成为隐私保护与网络空间信息安全关键技术的重要原因. Goldwasser[3] 曾经预言，安全多方计算将成为计算科学中一个必不可少的组成部分. 这个预言激励许多研究人员投身于安全多方计算的研究中，并取得了许多重要的研究成果. Yao 用混淆电路[4] 的方法证明了所有的安全多方计算问题都是可解的，在后期研究中 Goldreich 等人[2,5] 给出了基于心理游戏(mental game)的通用解决方案. 然而，这两种通用解决方案的效率都很低，都只有理论意义. 因此 Goldreich 指出理论上证明所有安全多方计算问题的可解并不表示不再需要对具体问题进行研究. 相反，由于效率的原因，应用一般条件下导出的没有任何具体信息的解决方案来解决具体问题是不实际的，针对具体问题研究具体的解决方案可大大提高解决方案的效率. 目前，安全多方计算研究的问题主要包括保密的科学计算[1,6-9]、保密的计算几何[10-12]、保密的数据挖掘[13-14]、保密的统计分析[15-16] 和安全多方计算应用[17] 等，著名的百万富翁问题就是保密的科学计算问题之一.

百万富翁问题[1] 简单地说就是保密比较两个私有数据大小的问题，这一问题的解决方案已经成为构造许多安全多方计算问题解决方案的基本模块. 百万富翁问题一个很自然的推广就是在参与者为多人时比较出他们私有数据中的最大或最小数值，这样的推广显然很有理论和现实意义. 理论上如果能解决求多个数据中最大值或最小值的问题就能解决百万富翁问题以及与其有关的问题；现实中最大值问题的解决方案可以应用在保密的电子招投标等电子商务活动以及电子选举等社会活动中. 设想这样一个场景，在投标者为多人的电子投标中，招标者需要决定投标价格最高(低)的投标者，但每个投标者

都不想泄露自己的投标价格,这样的问题就可以利用保密求最大值的安全多方计算协议来解决.由此可见,保密计算最大(小)值的问题,是安全多方计算的一个重要问题,自然引起了人们的注意.目前已有一些文献提出过该问题的解决方案,但这些方案都是通过对保密数据的两两比较得出结果,因而要多次调用百万富翁问题协议,如果求 n 个数中的最大值,那么就需要调用 $n-1$ 次百万富翁问题协议[18-22],或者将其转化为排序问题[23].这样的方案不但效率不高而且还会泄露不该泄露的信息(如两两数据之间的大小关系与总的顺序).最大(小)值的安全多方计算要求所有参与者只能得到最大(小)值而不能泄露其他任何信息.这正是解决最大(小)值保密计算问题的困难所在,也是现有的方案基本无法解决的问题.从这个意义上说,连续多次调用百万富翁问题协议进行两两比较求最大值的方法和转化为排序问题的方法都没有真正解决最大(小)值的保密计算问题.只有采用全新的思路才能使这个问题获得满意的解决.文献[24]应用编码的方法解决了这一问题,将最小值的运算转化为求对应编码位置乘积的运算.但由于其需要应用具有乘法同态性[25]的加密算法,所以给出的解决方案计算效率不高.本文主要的工作就是针对最大(小)值问题给出高效的解决方案,通过引入新的编码方法将计算最大值的问题归约到对于编码的运算,避免了两两比较的问题,从而跳出了多次调用百万富翁问题协议的传统思维,巧妙解决了最大值的保密计算问题.方案不需要调用百万富翁问题协议且不需要加密方案具有任何同态性,从而避免了不必要的信息泄露,提高了效率.进一步以此为基础解决了最小公倍数和最大公约数的保密计算问题.在数学中,求多个整数的最大公约数和最小公倍数能够帮助我们更好地研究整数的内在性质,例如,多个整数的最大公约数为1则说明它们是互素的.在密码学中,对整数最大公约数和最小公倍数的研究也是必不可少的.除此之外,解决最大公约数和最小公倍数问题也具有很重要的现实意义.但目前,还没有看到针对此问题的研究和相关的解决方案.

本文的主要贡献如下:

(1)利用概率加密方案的性质提出了一种保密替换的方法,这种方法可以用来解决一些保密的科学计算问题.

(2)针对保密数据属于一个稠密集的情况设计了一种编码方法,利用这种编码方法和概率密码系统、门限解密密码系统设计了两个保密计算多个数据最大(小)值的协议.这两个协议在半诚实模型下是安全的,可以抵抗不同程度的合谋攻击.

(3)针对保密数据属于一个稀疏集的情况利用类似编码方法[9]、门限解密、概率加密系统设计了两个保密计算多个数据最大(小)值的协议,这两个协议在半诚实模型下是安全的,可以抵抗任意数量的合谋攻击.

(4)以保密计算最大(小)值协议为基础,设计了保密计算多个数据的最小公倍数和最大公约数的协议并给出了最小公倍数保密计算的具体协议.同样,该协议在半诚实模型下是安全的且可以抵抗任意合谋攻击.

2 预备知识

2.1 理想模型

假设存在一个完全诚实的可信的第三方(Trusted Third Party),它在任何情况下都不会泄露任何不应该泄露的信息.那么 n 个参与者可以分别将自己的私有数据 $x_i(i\in\{1,2,\cdots,n\})$ 发送给它,它计算函数 $f(x_1,x_2,\cdots,x_n)$ 并将结果分别发给各个参与者.这种借助于可信的第三方完成的安全多方计算协议称为理想的安全多方计算协议(简称理想协议).理想协议是安全性最高且最简单的安全多方计算协议,任何一个计算函数 f 的实际协议的安全性都不可能超过这个协议;任意安全多方计算协议与理想协议相比,如果不泄露更多的信息都被认为是安全的.理想协议虽然简单且安全,但在应用中却经常受到限制.因为这样可信的第三者难以找到,如果找到也会成为网络合作计算活动的通信与安全瓶颈,而且可能还需要付出经济的代价,而这些经济、安全与时间代价往往是不可接受的.

2.2 半诚实模型

半诚实参与者.简单来说,半诚实参与者是指在协议执行过程中会忠实地履行协议但可能会记录下中间结果并试图从中推算出其他参与者私有信息的参与者.

设有 n 个半诚实参与者,分别拥有私有数据 x_i $(i\in\{1,2,\cdots,n\})$,他们执行保密计算函数 $f(x_1,x_2,\cdots,x_n)$ 的协议 Π 并在协议完成后得到最终结果.这种参与者都是半诚实的安全多方计算协议就被称为半诚实模型下的安全多方计算协议(简称半诚实协议),本文所设计协议都是半诚实协议.令 $X=(x_1,x_2,\cdots,x_n)$,在执行协议过程中,每一个参与者得到的消息序列记为

$$view_i^{\Pi}(X)=(x_i,r_i,M_i^1,\cdots,M_i^t),$$

其中 $M_i^j(j=1,2,\cdots,t)$ 表示第 i 个参与者收到的第 j 个信息，r_i 是第 i 个参与者产生的随机数.

定义 1.　半诚实参与者协议的安全性[26]. 设 $f:(\{0,1\}^*)^n\to(\{0,1\}^*)^n$ 是 n 元函数，$f_i(x_1,x_2,\cdots,x_n)$ 为 $f(x_1,x_2,\cdots,x_n)$ 中的第 i 个元素，$I=\{P_{i1},P_{i2},\cdots,P_{is}\}\subseteq\{P_1,P_2,\cdots,P_n\}$ 表示任意参与者的子集(这里 P_i 表示第 i 个参与者)，$f_I(x_1,x_2,\cdots,x_n)$ 为序列 $f_{i1}(x_1,x_2,\cdots,x_n),\cdots,f_{is}(x_1,x_2,\cdots,x_n)$，$output^{\Pi}(X)$ 为所有参与方执行协议 Π 的输出. 在参与者都是半诚实的情况下，如果存在概率多项式时间算法 S，使得对于任意的 I 均有下式成立：

$$\{S(I,(x_{i1},x_{i2},\cdots,x_{is}),f_I(X))\}_{X\in(\{0,1\}^*)^n}\overset{c}{\equiv}\{view_I^{\Pi}(X)\}_{X\in(\{0,1\}^*)^n}\quad(1)$$

那么称协议 Π 保密地计算 n 元函数 f. 其中 $\overset{c}{\equiv}$ 表示计算不可区分，$view_I^{\Pi}(X)=(I,view_{i1}^{\Pi}(X),\cdots,view_{is}^{\Pi}(X))$，即 $view_I^{\Pi}(X)$ 只包括 I 中参与者在协议执行过程中所得到的消息序列，不包括诚实方之间所传递的、I 中参与者看不到的信息.

2.3　概率加密算法

一个公钥加密方案[27] ε 由三个算法 $\mathrm{KeyGen}_\varepsilon$、$\mathrm{Encrypt}_\varepsilon$、$\mathrm{Decrypt}_\varepsilon$ 组成.

(1) $\mathrm{KeyGen}_\varepsilon$. 给定安全参数 λ，$\mathrm{KeyGen}_\varepsilon$ 算法输出一个私钥 sk 和对应的公钥 pk，明文空间 M 和相应的密文空间 C，即

$$(sk,pk,M,C)\leftarrow\mathrm{KeyGen}_\varepsilon(\lambda).$$

(2) $\mathrm{Encrypt}_\varepsilon$. 给定公钥 pk 加密明文 $m\in M$，$\mathrm{Encrypt}_\varepsilon$ 输出相应的密文 $c\in C$：

$$c\leftarrow\mathrm{Encrypt}_\varepsilon(pk,m).$$

(3) $\mathrm{Decrypt}_\varepsilon$. 给定密文 c 和私钥 sk，$\mathrm{Decrypt}_\varepsilon$ 输出相应的明文 m：

$$m\leftarrow\mathrm{Decrypt}_\varepsilon(sk,c).$$

概率公钥加密算法加密过程中需要额外选择一个随机数 $r\in M$ 参与运算，因此概率公钥加密算法的 $\mathrm{Encrypt}_\varepsilon$ 可以简要的表达为

$$c\leftarrow\mathrm{Encrypt}_\varepsilon(pk,m,r),$$

因此对于相同的明文 m 选用不同的随机数 r 会产生不同的密文 c.

2.4　门限解密

门限解密[28]是安全多方计算研究中对抗合谋攻击的一个重要工具. 在 (t,n) 门限解密密码体制中，n 个参与者联合生成一个公钥，解密密钥由 n 方联合持有. 加密可直接利用公钥完成，但解密必须由 n 个参与者中的 t 个 $(1<t\le n)$ 合作才能完成，而少于 t 个人合作将无法得到正确的明文，这样的密码体制称为 (t,n) 门限的密码体制. 因为本文需要抵抗尽可能多的合谋攻击，所以需要的是 (n,n) 门限密码系统，这样可以对抗 $n-1$ 个参与者的合谋攻击. RSA、Paillier 和 ElGamal 等密码系统都可以用来构造门限密码系统，下面以 ElGamal 为例给出一种门限密码系统的具体构造.

(1) KeyGen. 给定安全系数 k，KeyGen 产生一个大素数 p，并随机选择 Z_p^* 上的一个生成元 g. 每个参与者 P_i 随机地选取一个私钥 $sk_i\in Z_p^*$，并计算 $h_i=g^{sk_i}\bmod p$. 公钥则为

$$h=\prod_{i=1}^{n}h_i\bmod p=g^{\sum_{i=1}^{n}sk_i}\bmod p.$$

(2) Encrypt. 对于明文 M 加密，选择一个随机数 r，密文为

$$E(M)=(c_1,c_2)=(g^r\bmod p,Mh^r\bmod p).$$

(3) Decrypt. 对于密文 $c=(c_1,c_2)$ 解密，明文为

$$M=\frac{c_2}{\prod_{i=1}^{n}c_1^{sk_i}}\bmod p.$$

2.5　密文的自盲性

在概率加密系统中，一个明文可以有很多个不可区分的密文，很多密文都被解密为同一个明文. 若某加密系统加密的任一密文都可以在不知道私钥的条件下将其改变为同一个明文的不同密文，则称该加密系统具有自盲性[29]. 事实上，所有的概率加密系统都具有自盲性. 例如，在 Paillier 加密系统中，对于 $\forall m\in Z_n$ 和 $r\in N$，都有 $D(E(m)r^n\bmod n^2)=m$，$D(E(m)g^{nr}\bmod n^2)=m$. 此处，n 和 g 为 Paillier 加密系统的公钥，E,D 分别为 Paillier 加密系统的加密算法和解密算法. 容易看出，只需在原有密文 $E(m)$ 的基础上乘上任意随机数 r^n 或乘上公开参数 g^{nr} 就可将原有密文改变为另一个密文，而保持明文不变，而且若不公开所选随机数 r，利用自盲性计算得到的密文 $E'(m)$ 与调用加密算法计算得到的密文 $E(m')$ 是计算不可区分的. 因此，Paillier 加密系统具有自盲性，又称为再随机化特性(re-randomizing property).

3　基于概率加密方案的最大值保密计算

3.1　编码与替换原理

问题描述. n 个参与者 $P_1,P_2,\cdots,P_n$ 分别拥有私有数据 $x_1,x_2,\cdots,x_n$，假设 $1\le x_i\le m(i=1,2,\cdots,$

n). 他们要计算所有数据中的最大值 $y=\max\{x_1,x_2,\cdots,x_n\}$或最小值 $y=\min\{x_1,x_2,\cdots,x_n\}$,同时不愿泄露任何有关私有数据的信息.

编码方法 1. 每一个数据 x 都可按照如下编码方法被表示成一个与其唯一对应的数组 $X=(x_1,x_2,\cdots,x_m)$,其中

$$x_j=\begin{cases}0, & j\leq x\\ 1, & j>x\end{cases}.$$

第一个参与者 P_1 首先按照编码方法 1 构造数组 $X_1=(x_{11},x_{12},\cdots,x_{1m})$. 以这种方式进行编码后,第一个参与者 P_1 的私有数据 x_1 就被表示成了数组

$$X_1=(\overbrace{0,\cdots,0}^{x_1\text{个}},\overbrace{1,\cdots,1}^{m-x_1\text{个}}),$$

数组的前面有 x_1 个 0,后面接着有 $m-x_1$ 个 1. 为了叙述简洁,将 X_i 中 0(1)的个数记为$\langle X_i\rangle(\overline{\langle X_i\rangle})$,则有$\langle X_i\rangle=x_i,\overline{\langle X_i\rangle}=m-x_i$.

要计算 x_1 与 x_2 中的最大值,仅需参与者 P_2 用与自己的私有数据 x_2 相对应的 x_2 个 0 替换掉原有数组 X_1 中的前 x_2 个元素即可. 容易看出,若 $x_2\leq x_1$,经过替换后的数组 X_2 中 0 的个数没有变化;若 $x_2>x_1$,替换后的 X_2 中 0 的个数等于 x_2,也就是说 $\langle X_2\rangle=\max\{x_1,x_2\}$. 依次类推,参与者 P_{i+1} 可通过用 x_{i+1} 个 0 替换数组 X_i 中的前 x_{i+1} 个元素得到新的数组 X_{i+1} 而其他元素保持不变,最后可以根据 X_{i+1} 求得 $\max\{(x_1,x_2,\cdots,x_i),x_{i+1}\}=\langle X_{i+1}\rangle$.

例如,令 $x_1=10,x_2=14,x_3=6$,选取 $m=20$,那么 P_1 构造的 X_1 如下:

$X_1=(0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,1,1)$,

那么经由 P_2 替换后的 X_2 为

$X_2=(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1)$,

经由 P_3 替换后的 X_3 为

$X_3=(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,1)$,

显然有

$$\max(10,14,6)=\langle X_3\rangle=14.$$

若要得出 x_1 与 x_2 中的最小值,参与者 P_2 仅需用与自己的私有数据 x_2 相对应的 $m-x_2$ 个 1 替换掉原有数组 X_1 中的后 $m-x_2$ 个元素即可. 容易看出,若 $x_2<x_1$,必有$\overline{\langle X_2\rangle}>\overline{\langle X_1\rangle}$;若 $x_2\geq x_1$,必有$\overline{\langle X_2\rangle}=\overline{\langle X_1\rangle}$,也就是说$\overline{\langle X_2\rangle}=\min\{x_1,x_2\}$. 依次类推,参与者 P_{i+1} 可通过用 $m-x_{i+1}$ 个 1 替换数组 X_i 中的后 $m-x_{i+1}$ 个元素得到新的数组 X_{i+1},其他元素不变,因此 $\min\{(x_1,x_2,\cdots,x_i),x_{i+1}\}=\overline{\langle X_{i+1}\rangle}$.

例如,令 $x_1=10,x_2=14,x_3=6$,选取 $m=20$,那么 P_1 构造的 X_1 如下:

$X_1=(0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,1,1)$,

那么经由 P_2 替换后的 X_2 为

$X_2=(0,0,0,0,0,0,0,0,0,0,1,1,1,1,1,1,1,1,1,1)$,

经由 P_3 替换后的 X_3 为

$X_3=(0,0,0,0,0,0,1,1,1,1,1,1,1,1,1,1,1,1,1,1)$,

显然有

$$\min(10,14,6)=\overline{\langle X_3\rangle}=6.$$

以上提出的编码方法 1 就是本文计算在确定范围内的多个数据中最大值和最小值的基本原理. 直接这样计算是无法保密的,而在密文的条件下进行这样的操作则可以实现保密. 本文利用概率公钥加密算法对数组中的 0 和 1 进行概率加密,使得任何参与者都无法分辨出数组中 0 与 1 的个数. 做替换时,在用密文 $E(0)_{(i+1,j)}$ 替换原有密文 $E(0)_{(i,j)}$ 的同时利用概率加密方案本身具有的自盲性对 $m-x_{i+1}$ 个密文进行重随机化来保证本协议的保密性. 由于这种方法只是对密文进行了替换,所以选用的概率加密方案不需要任何同态性,可选用效率高的概率加密方案完成.

3.2 朴素的最大值保密计算方案

本小节将给出一种简单的最大值保密计算方案.

协议 1. 简单的最大值保密计算.

输入:$P_1,P_2,\cdots,P_n$ 各自的私有数据 $x_1,x_2,\cdots,x_n$

输出:$y=\max\{x_1,x_2,\cdots,x_n\}$

1. P_1 应用高效概率加密系统生成公钥 p_k 和私钥 s_k,并公布 p_k;

2. P_1 按照编码方法 1 将自己的数据 x_1 转化成数组 X_1,并将 $E(X_1)=(E(x_{1,1}),\cdots,E(x_{1,m}))$ 发送给 P_2;

3. 令 $c=(c_1,\cdots,c_m)=(E(x_{1,1}),\cdots,E(x_{1,m}))$;

4. 参与者 $P_i(i=2,\cdots,n)$ 计算如下:

For $i=2$ to $n-1$

P_i computes

$E_{i,k}\leftarrow E(0)(k=1,2,\cdots,x_i)$

$(c_1,\cdots,c_n)\leftarrow(E_{i,1},\cdots,E_{i,x_i},E'_{i-1,x_i+1},\cdots,E'_{i-1,m})$

sends to P_{i+1}

End

P_n computes

$E_{n,k}\leftarrow E(0)(k=1,2,\cdots,x_i)$

$(c_1,\cdots,c_n)\leftarrow(E_{n,1},\cdots,E_{n,x_n},E'_{n-1,x_n+1},\cdots,E'_{n-1,m})$

其中 $E'_{i-1,x_i+1}(i=2,\cdots,n)$ 表示利用加密方案的自盲性进行重随机化后的密文,若无特别说明以下协议中 E' 所表示内容与此处相同.

5. 参与者 P_n 与参与者 P_1 计算如下:

P_n sends c to P_1

P_1 computes $Y=D(c)$

$y\leftarrow\langle Y\rangle$

6. P_1输出 y 并公布结果.

类似地，可以给出相应的求最小值的安全多方计算协议.

3.3 协议 1 的分析

正确性分析. 协议 1 的正确性可由 3.1 节中所述的原理保证.

安全性分析. 协议 1 的安全性以概率加密方案的安全性为基础，概率加密方案都是语义安全的，应用模拟范例可以证明协议 1 的安全. 由于协议 1 只有 P_1可以解密，所以协议 1 可以抵抗没有 P_1参与的任何合谋攻击，但不能抵抗有 P_1参与合谋的攻击. 若 P_1与 P_{i-1}，P_{i+1}合谋则可以恢复出 P_i的数组 X_i.

定理 1.　保密计算最大值的协议 1 在半诚实模型下是安全的.

证明.　通过构造满足式(1)的模拟器 S 来证明本定理，此处选用可能参与合谋的最大合谋结构进行模拟. 分为以下 3 种情况：

(1) P_1不参与合谋，$P_2,\cdots,P_n$合谋要获得 P_1的保密数据的信息. 因为只有 P_1才能够解密密文. 如果他不合谋，除了 $P_2,\cdots,P_n$的数据和他们生成的密文之外，$P_2,\cdots,P_n$收到的关于x_1的信息只有 $E(X_1)=(E(x_{1,1}),\cdots,E(x_{1,m}))$. 由于加密算法是语义安全的，$(E(x_{1,1}),\cdots,E(x_{1,m}))$和 m 个随机数是计算不可区分的. 在此情况下

$$I=\{P_2,\cdots,P_n\},X=\{x_1,x_2,\cdots,x_n\},$$

$$f(X)=f_I(X)=\min\{x_1,x_2,\cdots,x_n\},$$

$$\{view_I^{\Pi}(X)\}_{X\in(\{0,1\}^*)^n}=(I,view_2^{\Pi}(X),\cdots,view_n^{\Pi}(X))$$

$$=\{I,x_2,R_2,E(X_1),\cdots,x_n,R_n,E(X_n),f_I(X)\}.$$

给定输入$(I,(x_2,\cdots,x_n),f_I(X))$，$S$ 随机选择 $x_1'\in\{1,2,\cdots,m\}$使得

$$f(X')=f_I(x_1',x_2,\cdots,x_n)=\min\{x_1',x_2,\cdots,x_n\}$$

$$=f_I(X)=f(x_1,x_2,\cdots,x_n)=\min\{x_1,x_2,\cdots,x_n\},$$

用 $x_1',x_2,\cdots,x_n$进行模拟. 首先按照协议要求构造数组 X_1'：

$$X_1'=(\overbrace{0,\cdots,0}^{x_1'个},\overbrace{1,\cdots,1}^{m-x_1'个}),$$

加密 X_1'得到

$$E(X_1')=(E(x_{11}'),\cdots,E(x_{1m}')).$$

模拟器 S 按照协议要求进行加密替换及更新. 解密最终数组 $E(X_n')$得到

$$f_I(x_1',x_2,\cdots,x_n)=\min\{x_1',x_2,\cdots,x_n\}.$$

令

$$S(I,(x_2,\cdots,x_n),f_I(X))$$

$$=\{I,x_2,R_2,E(X_1'),\cdots,x_n,R_n,E(X_n),f_I(X')\}.$$

因为概率加密方案是语义安全的，所以 $E(X_1)\overset{c}{\equiv}E(X_1')$，且 $f_I(X)=f_I(X')$，其他所有参数都是相等的，故

$$\{S(I,(x_2,\cdots,x_n),f_I(X))\}_{X\in(\{0,1\}^*)^n}$$

$$\overset{c}{\equiv}\{view_I^{\Pi}(X)\}_{X\in(\{0,1\}^*)^n}.$$

因此对于 x_1是安全的.

(2) P_1不参与合谋，$P_2,\cdots,P_n$的 $n-2$ 个合谋想得到其中一个的保密数据，因为这时 $P_2,\cdots,P_n$的地位是平等的，能力是相同的，不失一般性假设 $P_3,\cdots,P_n$合谋要获得 P_2的保密数据的信息. 这种情况与第一种情况相同，用类似的模拟可以证明对于 x_2是安全的.

(3) P_1参与合谋. 这种情况下的最大合谋结构仍然是包含 P_1的 $n-1$ 个参与者合谋，要得到某一个 $P_i\in\{P_2,\cdots,P_n\}$的某个参与者的保密数据 x_i的信息. 在此情况下，他们可以获得 $P_i\in\{P_2,\cdots,P_n\}$的保密数据 x_i的大小，而如果他们利用可信的第三者保密计算最大值，当 x_i是最大值时也会泄露 x_i的大小；如果 x_i不是最大值，他们将无法知道 x_i的具体数值. 因此这和使用可信第三者的安全性只有微小的差别.

综上所述，该协议对于半诚实参与者是安全的.

证毕.

4 基于门限解密的最大值保密计算

下面我们利用门限解密系统构造一个抗合谋攻击的最大值保密计算协议. 该协议的基本原理与 3.1 节中所述原理一致，即参与者 P_1首先将其私有数据 x_1转化为数组 X_1并进行加密，其他的参与者逐个按照规则对密文数组进行替换，替换数组中 0 的最大下标值即为最大值.

4.1 抗合谋攻击的最大值保密计算方案

协议 2.　基于门限解密的最大值保密计算.

输入：$P_1,P_2,\cdots,P_n$的私有数据 $x_1,x_2,\cdots,x_n$

输出：$y=\max\{x_1,x_2,\cdots,x_n\}$

1. $P_1,P_2,\cdots,P_n$首先协商一个公钥系统及其参数，并选择自己的私钥 p_{ki}，并联合生成公钥 h；

2. P_1按照编码方法 1 将自己的保密数据 x_1转化成数组 X_1并用公钥加密，将

$$E(X_1)=(E(x_{1,1}),\cdots,E(x_{1,m}))$$

发送给 P_2；

3. 令 $c=(c_1,\cdots,c_m)=(E(x_{1,1}),\cdots,E(x_{1,m}))$；

4. 参与者 $P_i(i=2,\cdots,n)$计算如下：

For $i=2$ to $n-1$

P_i computes

$E_{i,k}\leftarrow E(0)(k=1,2,\cdots,x_i)$

$(c_1,\cdots,c_n)\leftarrow(E_{i,1},\cdots,E_{i,x_i},E'_{i-1,x_i+1},\cdots,E'_{i-1,m})$

sends to P_{i+1}

End

P_n computes

$E_{n,k}\leftarrow E(0)(k=1,2,\cdots,x_n)$

$(c_1,\cdots,c_n)\leftarrow(E_{n,1},\cdots,E_{n,x_n},E'_{n-1,x_n+1},\cdots,E'_{n-1,m})$

5. 计算最大值的过程如下：

$P_1,P_2,\cdots,P_n$联合解密 $Y=D(c)$

$y\leftarrow\langle Y\rangle$

Outputs y

4.2 协议 2 的方案分析

正确性分析. 协议 2 的正确性由 2.5 节及 3.1 节中所述原理来保证.

安全性分析. 协议 2 的安全性是基于门限加密算法的安全性，由于门限公钥系统的公钥是由所有参与者共同产生的，而解密是所有参与者合作才能完成的. 而在计算过程中每个参与者 P_i 仅对外公布了加密后的信息，且若没有 P_i 的参与，所有加密信息是无法正确解密的，也就是说协议 2 可以抵抗合谋攻击.

定理 2. 基于门限解密的最大值保密计算协议 2 是安全的.

定理 2 的证明与定理 1 基本相同，所以我们只给出证明思路. 根据语义安全的概率门限加密算法的性质，若没有全部参与者合作，应用公钥加密的任何信息都是计算不可区分的. 因此，只要有一个参与者不参与合谋，对其他参与者来说实际执行协议时获得的 *view* 与用满足最大值不变的任意一组输入进行模拟所得到的信息序列是计算不可区分的，这种情况下，模拟器的构造是显而易见的.

5 基于门限解密的最大值高效保密计算

上一节中所给出的协议 1 与协议 2 很明显只能解决所有的数据都在一个确定范围内的最大值问题，比如 $x_1,\cdots,x_n\in\{k+1,\cdots,k+m\}(k\in Z_n)$. 但对于数据所在范围过大，即 m 的值很大时，这种方法的计算复杂性就很高了. 为解决这一问题，本节将给出适用于所有数据都属于某个稀疏集合内的最大值问题解决方案.

5.1 协议的基本原理

问题描述. n 个参与者 $P_1,P_2,\cdots,P_n$分别拥有私有数据 $x_1,x_2,\cdots,x_n$. 他们要计算所有数据中的最大值 $y=\max\{x_1,x_2,\cdots,x_n\}$或最小值 $y=\min\{x_1,x_2,\cdots,x_n\}$，同时不愿泄露任何有关私有数据的信息.

假设 $x_1,x_2,\cdots,x_n\in\{z_1,z_2,\cdots,z_m\}=U$，其中 $z_1<z_2<\cdots<z_m$. 进一步假设 $x_1=z_k,x_2=z_l,\cdots,x_n=z_p(1\le k,l,p\le m)$，那么 $x_1\le x_2$ 当且仅当 $k\le l$.

编码方法 2. 基于上述假设，每一个数据 $x=z_k$ 都可以按照如下编码方法被表示成一个与其唯一对应的数组 $X=(x_1,x_2,\cdots,x_m)$，其中

$$x_j=\begin{cases}0, & j\le k\\ 1, & j>k\end{cases}.$$

参与者 P_1 首先按照编码方法 2 构造一个数组 $X_1=(x_{11},\cdots,x_{1m})$. P_1 的私有数据 x_1 就被表示成了数组

$$X_1=(\overbrace{0,\cdots,0}^{k\text{个}},\overbrace{1,\cdots,1}^{m-k\text{个}}),$$

在数组中 0 元素的最大下标为 k.

若要计算 $\min\{x_1,x_2\}$，P_2 仅需用 l 个 0 替换掉原有数组 X_1 中的前 l 个元素即可. 容易看出，若 $x_2\le x_1$，则有 $l\le k$，必有$\langle X_2\rangle=\langle X_1\rangle$；若 $x_2>x_1$，必有$\langle X_2\rangle>\langle X_1\rangle$，也就是说 $z_{\langle X_2\rangle}=\max\{x_1,x_2\}$. 依次类推，参与者 P_{i+1} 可通过替换得到新的数组 X_{i+1}，据其求得 $\max\{(x_1,x_2,\cdots,x_i),x_{i+1}\}=z_{\langle X_{i+1}\rangle}$.

例如，令 $x_1=8,x_2=19,x_3=4$，全集 $U=\{1,4,6,8,12,13,17,19,25,40\}$，那么

$$X_1=(0,0,0,0,1,1,1,1,1,1),$$

那么经由 P_2 替换后的 X_2 为

$$X_2=(0,0,0,0,0,0,0,0,1,1),$$

经由 P_3 替换后的 X_3 为

$$X_3=(0,0,0,0,0,0,1,1,1,1),$$

显然有

$$\max(8,19,4)=z_{\langle X_3\rangle}=z_8=19.$$

若要得出 x_1 与 x_2 中的最小值，P_2 仅用 $m-l$ 个 1 替换掉原有数组 X_1 中的后 $m-l$ 个元素即可. 容易看出，若 $x_2<x_1$，必有$\overline{\langle X_2\rangle}>\overline{\langle X_1\rangle}$；若 $x_2\ge x_1$，必有$\overline{\langle X_2\rangle}=\overline{\langle X_1\rangle}$，也就是说 $z_{\overline{\langle X_2\rangle}}=\min\{x_1,x_2\}$. 依次类推，参与者 P_{i+1} 可通过替换得到新的数组 X_{i+1}，据其求得 $\min\{(x_1,x_2,\cdots,x_i),x_{i+1}\}=z_{\overline{\langle X_{i+1}\rangle}}$.

例如，令 $x_1=8,x_2=19,x_3=4$，全集 $U=\{1,4,6,8,12,13,17,19,25,40\}$，那么

$$X_1=(0,0,0,0,1,1,1,1,1,1),$$

那么经由 P_2 替换后的 X_2 为

$$X_2=(0,0,0,0,1,1,1,1,1,1),$$

经由 P_3 替换后的 X_3 为

$$X_3=(0,0,1,1,1,1,1,1,1,1),$$

显然有

$$\min(8,19,4)=z_{\langle X_3\rangle}=z_2=4.$$

以上提出的编码方法 2 就是本文针对计算在某个稀疏集合内的多个保密数据的最大值和最小值的基本原理.同样,直接这样计算会泄露很多信息,本文利用概率公钥加密方案对数组中的 0 和 1 进行概率加密,使得无法从密文中分辨出数组中 0 与 1 的个数,因而替换称为保密替换,以此保密计算最大(小)值.替换方法同 3.1 节中所述,这种方法同样可以选用效率高的概率加密方案完成.

5.2　抗合谋的最大值高效保密计算方案

协议 3.　基于门限解密的高效最大值保密计算.

输入:$P_1,P_2,\cdots,P_n$ 各自的私有信息 $x_1,x_2,\cdots,x_n$,全集 U

输出:$y=\max\{x_1,x_2,\cdots,x_n\}$

1. $P_1,P_2,\cdots,P_n$ 首先选取一种公钥系统的公开参数,并选择自己的私钥 p_{ki},之后联合生成公钥 h;

2. P_1 按照编码方法 2 将自己的数据 x_1 转化成数组 X_1 并用公钥加密,将

$$E(X_1)=(E(x_{1,1}),\cdots,E(x_{1,m}))$$

发送给 P_2;

3. 令 $c=(c_1,\cdots,c_m)=(E(x_{1,1}),\cdots,E(x_{1,m}))$;

4. 令 $x_i=z_{ki}$ 对应参与者 $P_i(i=2,\cdots,n)$ 计算如下:

For $i=2$ to $n-1$

P_i computes

$E_{i,q}\leftarrow E(0)(q=1,2,\cdots,ki)$

$(c_1,\cdots,c_n)\leftarrow(E_{i,1},\cdots,E_{i,ki},E'_{i-1,ki+1},\cdots,E'_{i-1,m})$

sends to P_{i+1}

End

P_n computes

$E_{n,q}\leftarrow E(0)(q=1,2,\cdots,kn)$

$(c_1,\cdots,c_n)\leftarrow(E_{n,1},\cdots,E_{n,kn},E'_{n-1,kn+1},\cdots,E'_{n-1,m})$

5. 计算最大值的过程如下:

$P_1,P_2,\cdots,P_n$ 联合解密 $Y=D(c)$

$y\leftarrow\langle Y\rangle$

Outputs z_y

5.3　协议 3 的方案分析

正确性分析. 协议 3 的正确性可由 2.5 节门限解密的基本原理及 5.1 节协议的基本原理得到保证.

安全性分析. 协议 3 的安全性可由协议 2 的安全性证明得到保证.

定理 3.　抗合谋的最大值高效保密计算协议 3 是安全的.

证明.　定理 3 的证明思路与定理 2 的证明思路相同,这里省略具体证明过程.　　证毕.

6　效率分析

现在对上面保密计算最大值的三个协议的效率进行全面的分析比较.

计算效率分析. 首先分析协议的计算复杂性,每一个协议都假设有 n 个参与者.

在协议 1 中,假设每个参与者的数据不超过 m,由于协议不需要任何同态性且只需加密 0 和 1,因此选用效率较高的 Goldwasser-Micali(GM)加密算法完成,也可以选用椭圆曲线或者 NTRU 加密算法实现.假设采用 GM 加密算法,参与者 P_1 需要对 0 或 1 进行加密,这个过程最多需要计算 $2m$ 次模乘运算.除 P_1 外,P_i 需要加密对应的 x_i 个 0,然后更新 $m-x_i$ 个原有密文.由于 GM 算法更新原有密文需要计算两次模乘运算,所以 n 个参与者最多需要计算 $2nm$ 次模乘运算完成所有替换.最后,参与者 P_1 对数组 $E(X_n)$ 进行解密,GM 算法解密一次需要 $\log p$ 次模乘运算(其中 p 为 GM 加密系统中的大素数私钥).所以协议 1 最多共需要 $2nm+\log p$ 次模乘运算.

在协议 2 中,同样假设每个参与者的数据不超过 m,由于协议要求使用门限密码,为了方便描述,此处我们以 ElGamal 加密算法为例进行说明.当然,若选用效率更高的门限密码系统,本协议的效率也就会更高.首先,参与者各方合作产生公钥 $h=g^{k_1+\cdots+k_n} \bmod p$(这里的 p 是 ElGamal 加密算法中的公开参数),共需要 n 次模指数运算.加密过程所有参与者最多共需要 $2nm$ 次模指数运算,解密过程最多需要 nm 次模指数运算.所以协议 2 最多需要 $m(3n+1)$ 次模指数运算.

在协议 3 中,假设所有参与者的数据都包含在集合 U 中,且 U 的势为 m.由于协议 3 也需要使用门限密码,所以协议 3 最多也需要 $m(3n+1)$ 次模指数运算.

通信效率分析. 协议 1 中,每一个参与者将数组加密或替换后的密文发送给下一个参与者,这期间参与者需要 $n-1$ 次通信.在最后解密时,P_n 将最终数组 $E(X_n)$ 发送给 P_1,这仅需要一次通信.因此协议共需要 n 次通信.

协议 2、协议 3 所有参与者共同构造公钥和加密过程各需要 $n-1$ 次通信,解密过程需要 $n-1$ 次通信,所以共需要 $2n-2$ 次通信.

实验模拟. 我们通过模拟实验来测试本文中协议 1、协议 2 和协议 3 执行所用的时间,通过协议执行的时间来验证效率.实验的测试环境:Windows 10 64 位操作系统,处理器参数为 Intel(R) Core(TM) i5-6600 CPU@3.30GHz,8GB 内存,用 JAVA 语言编程实现.本实验以保密求三个数据的最大值为例,

数据范围为 100 以内的数据到 1000 以内的数据，协议 1 使用的是 GM 加密系统，协议 2 和协议 3 使用的是 ElGamal 门限加密系统进行模拟，都忽略了预处理时间.

实验结果. 从图 1 可知在 P_1 不参与合谋的情况下要保密求某一确定范围内多个数据的最大值，协议 1 的效率是最高的. 在无法确定 P_1 诚实性(即是否会参与合谋)的情况下要保密求在某一确定集合内的多个数据最大值，协议 3 的效率是最高的. 在本实验中，协议 2 与协议 3 使用的都是效率一般的 ElGamal 门限加密系统进行模拟，若选用其他效率更高的门限加密系统，协议的效率会更高.

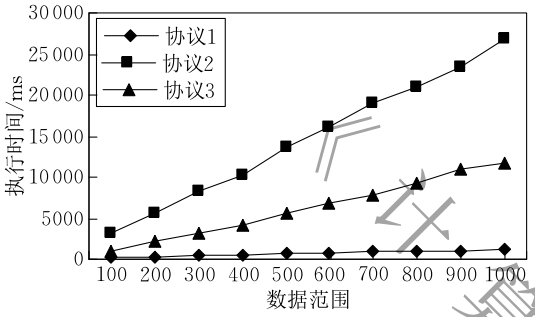

图 1 保密求最大值的执行时间与数据范围增长的变化规律

7 最小公倍数和最大公约数的保密计算

前面给出的协议 1～3 解决了若干个私有数值的最大值保密计算问题，对这些协议稍加修改就可以解决最小值的问题. 事实上，在协议的基本原理部分已经给出了解决最小值问题的原理. 本节中，我们将以已经给出的协议为基础，进一步研究若干个数值的最小公倍数和最大公约数问题，并给出最小公倍数问题的具体解决方案.

7.1 保密计算多个数的最小公倍数

问题描述. 假设 n 个参与者 $P_i(i=1,2,\cdots,n)$ 各有一个私有数据 $x_1,x_2,\cdots,x_n$，他们希望保密计算这 n 个数的最小公倍数 $\mathrm{lcm}(x_1,x_2,\cdots,x_n)$，但都不想泄露自己的私有数据.

解决方法的基本原理. 由算术基本定理可知，每一个整数都可以被唯一地表示成若干素数乘积的形式，也就是因子分解. 即 x_i 可以被表示成

$$x_i=2^{x_{i1}}\cdot 3^{x_{i2}}\cdot\cdots\cdot q_m^{x_{im}} \tag{2}$$

那么 $x_1,x_2,\cdots,x_n$ 的最小公倍数就可以表示为

$$\mathrm{lcm}(x_1,x_2,\cdots,x_n)=2^{\max(x_{11},\cdots,x_{n1})}\cdot 3^{\max(x_{12},\cdots,x_{n2})}\cdot q_m^{\max(x_{1m},\cdots,x_{nm})}.$$

由上式可知，求多个数据最小公倍数的问题就可以这样归约到求多个集合对应位置最大值的问题，进而归约到求多个数据的最大值问题.

同理，$x_1,x_2,\cdots,x_n$ 的最大公约数就可以表示为

$$\gcd(x_1,x_2,\cdots,x_n)=2^{\min(x_{11},\cdots,x_{n1})}\cdot 3^{\min(x_{12},\cdots,x_{n2})}\cdot q_m^{\min(x_{1m},\cdots,x_{nm})}.$$

因此，求多个数据最大公约数的问题就可以这样归约到求多个集合对应位置最小值的问题，进而归约到求多个数据的最小值问题.

协议 4. 基于门限解密的最小公倍数保密计算.

输入：$P_1,P_2,\cdots,P_n$ 各自的私有信息 $x_1,x_2,\cdots,x_n$

输出：$y=\mathrm{lcm}(x_1,x_2,\cdots,x_n)$

1. $P_1,P_2,\cdots,P_n$ 首先共同商定一个 q_m，并将自己的私有数据用式(2)表示出来并生成自己的私有数组 $X_i=(x_{i1},\cdots,x_{im})$，共同商定 x_{ij} 的范围后(假定 $x_{ij}\in\{0,1,\cdots,k\}$)将 X_i 编码为矩阵形式

$$X_i'=\begin{bmatrix}\overbrace{0,\cdots,0}^{x_{i1}\text{个}},\overbrace{1,\cdots,1}^{k-x_{i1}\text{个}}\\ \cdots\\ \overbrace{0,\cdots,0}^{x_{im}\text{个}},\overbrace{1,\cdots,1}^{k-x_{im}\text{个}}\end{bmatrix}_{m\times k}$$

2. $P_1,P_2,\cdots,P_n$ 共同调用协议 2 得到结果

$$Y=\begin{bmatrix}Y_1\\ \cdots\\ Y_m\end{bmatrix}=\max(X_1',\cdots,X_n'),$$

这里 $\max(X_1',\cdots,X_n')$ 是指以每个数组对应位置的最大值为元素组成的数组.

3. $P_1,P_2,\cdots,P_n$ 各自计算

$$y=\mathrm{lcm}(x_1,x_2,\cdots,x_n)=2^{Y_1}\cdot 3^{Y_2}\cdot\cdots\cdot q_m^{Y_m}.$$

7.2 协议 4 的分析

正确性分析. 协议 4 的正确性可由协议 2 的正确性及 7.1 节中协议的基本原理得到保证.

安全性分析. 协议 4 的安全性可由协议 2 的安全性得到保证.

定理 4. 基于门限解密的最小公倍数保密计算是安全的.

证明. 定理 4 的证明思路与定理 2 的证明思路相同. 证毕.

8 结 论

本文利用概率加密算法设计了一种保密替换方法，基于一种新的编码方法和门限解密算法构造了三个解决最大(小)值问题的安全多方计算协议，并以此为基础构造了一个解决最小公倍数问题的安全多方计算协议，所给出的协议在半诚实模型下都是

安全的.在抗合谋攻击方面,协议 1 能抵抗不含参与者 P_1 的攻击,协议 2、协议 3 和协议 4 可以抵抗任何数量的合谋攻击.今后将在现有研究的基础上进一步研究更高效的最大值问题解决方案和任意多个数的最大值问题解决方案,并尝试开展对恶意模型下的保密科学计算问题的研究.

参 考 文 献

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

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

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

[4] Yao A. How to generate and exchange secrets//Proceedings of the 27th Annual Symposium on Foundations of Computer Science. Toronto, Canada, 1986: 162-167

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

[6] Fagin R, Naor M, Winklery P. Comparing information without leaking it. Communications of the ACM, 1996, 39(5): 77-85

[7] Boneh D, Goh E J, Nissim K. Evaluating 2-DNF formulas on ciphertexts. Theory of Cryptography. Berlin, Germany: Springer, 2005: 325-341

[8] Kissner L, Song D. Privacy-preserving set operations//Proceedings of the Advances in Cryptology-CRYPTO 2005. Berlin, Germany, 2005: 241-257

[9] Li Shun-Dong, Wang Dao-Shun. Efficient secure multiparty computation based on homomorphic encryption. Acta Electronica Sinica, 2013, 41(4): 798-803(in Chinese)
(李顺东，王道顺. 基于同态加密的高效多方保密计算. 电子学报，2013，41(4)：798-803)

[10] Atallah M J, Du W. Secure multi-party computational geometry. Algorithms and Data Structures. Berlin, Germany: Springer, 2001: 165-179

[11] Li Shundong, Wu Chunying, Wang Daoshun, Dai Yiqi. Secure multiparty computation of solid geometric problems and their applications. Information Sciences. 2014, 282: 401-413

[12] Qin J, Duan H, Zhao H, et al. A new Lagrange solution to the privacy-preserving general geometric intersection problem. Journal of Network and Computer Applications, 2014, 46: 94-99

[13] Kantardzic M. Data Mining: Concepts, Models, Methods, and Algorithms. New York, UK: John Wiley & Sons, 2011

[14] Aggarwal C C. Privacy-Preserving Data Mining. New Delhi, India: Springer International Publishing, 2015: 663-693

[15] Du W L, Atallah M J. Privacy-preserving cooperative statistical analysis//Proceedings of the 17th Annual Conference of Computer Security Applications. Piscataway, USA, 2001: 102-110

[16] Jawurek M, Kerschbaum F. Fault-tolerant privacy-preserving statistics: USA Patent 8,880,867. 2014-11-4

[17] Du W L, Atallah M J. Protocols for secure remote database access with approximate matching. Advance of E-Commerce and Privacy. New York, UK: Springer, 2001: 87-111

[18] Du W. A Study of Several Specific Secure Two-Party Computation Problems[Ph. D. dissertation]. Purdue University, West Lafayette, USA, 2001

[19] Elmehdwi Y, Samanthula B K, Jiang W. Secure *k*-nearest neighbor query over encrypted data in outsourced environments//Proceedings of the 2014 IEEE 30th International Conference on Data Engineering (ICDE), Chicago, USA, 2014: 664-675

[20] Brickell J, Shmatikov V. Privacy-preserving graph algorithms in the semi-honest model//Proceedings of the International Conference on the Theory and Application of Cryptology and Information Security. Berlin, Germany, 2005: 236-252

[21] Smaragdis P, Shashanka M. A Framework for Secure Speech Recognition. IEEE Transactions on Audio Speech & Language Processing, 2007, 15(4): 1404-1413

[22] Liu A, Zhengy K, Liz L, et al. Efficient secure similarity computation on encrypted trajectory data//Proceedings of the 2015 IEEE 31st International Conference on Data Engineering (ICDE). Seoul, Korea, 2015: 66-77

[23] Tang Chun-Ming, Shi Gui-Hua, Yao Zheng-An. Secure multy-party computation protocol for sequencing problem. Science China Information Science, 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：信息科学，2011，41(7)：789-797)

[24] Dou Jia-Wei, Ma Li, Li Shun-Dong. Secure multi-party computation for minimum and its applications. Acta Electronica Sinica, 2017, 45(7): 1715-1721(in Chinese)
(窦家维，马丽，李顺东. 最小值问题的安全多方计算及其应用. 电子学报，2017，45(7)：1715-1721)

[25] ElGamal T. A public key cryptosystem and a signature scheme based on discrete logarithms. Workshop on the Theory and Application of Cryptographic Techniques. Berlin, Germany: Springer, 1984: 10-18

[26] Goldreich O. Foundations of Cryptography: Volume 2, Basic Applications. London, UK: Cambridge University Press, 2009

[27] Reddy K A N, Vishnuvardhan B. The probabilistic encryption algorithm using linear transformation//Proceedings of the 49th Annual Convention of the Computer Society. Hyderabad, India, 2015: 389-395

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

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

YANG Xiao-Yi, born in 1993, Ph. D. candidate. Her main research interests include modern cryptography and information security.

LI Shun-Dong, born in 1963, Ph. D., professor, Ph. D. supervisor. His main research interests include modern cryptography and information security.

KANG Jia, born in 1992, M. S. candidate. Her main research interests include modern cryptography and information security.

Background

Secure multiparty computation (SMC), an important field of cryptography, was first introduced by Yao in the 1980s, and has become one of the most active research fields of modern cryptography and a focus in the international cryptographic community. 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 need to be studied; many previously addressed problems also require further effort to develop more efficient solutions.

The "millionaires' problem" is a basic problem of secure multi-party computation, which can be simply abstracted as privately comparing two numbers. The protocol for it can be used as building blocks (sub-protocols) to construct protocols for many other secure multiparty computation problems. A natural generalization of the millionaires' problem is to privately determine the maximum or the minimum of some private data owned by different parties. This problem is of great theoretical and practical significance. Its solution can be used to solve many other secure multiparty computation problems, or be directly used in electronic commerce, electronic voting etc. However, the existing protocols based on pair-wise comparison have to invoke the protocol for the millionaires' problem or reduce this problem into a sorting problem. If they are applied to privately compute the maximum, they are inefficient or insecure. They can disclose much more information than expected. This paper tries to solve this problem efficiently and securely. To solve this problem efficiently, we first propose a private substitution method which uses the property of probabilistic public key cryptosystem. Using this method, one can change the ciphertexts of a private data set to change or not to change the plaintexts of the data set, but no parties knows whether the plaintext of the data set has been changed or not. Then we propose a new encoding scheme to encode a private number into a private vector such that the maximum computation on the private data set can be performed on the private vectors to jump out the pair-wise comparison and to preserve the relationship of data. We finally propose our protocol for computing the maximum of private numbers. As the application of the protocol for privately computing the maximum or the minimum, we show how to use it to privately compute the least common multiple (the greatest common divisor) of some private numbers. The protocol for computing the maximum of private numbers can be directly in seal-bid to privately determine the winner.

Three different protocols proposed in this paper are appropriate for different cases. The first and the second are appropriate for the cases in which the private numbers are belong to a small dense set of which the cardinality is not very large. The second and the third are constructed based threshold decryption cryptosystem, and thus resist any collusion attack. The third is appropriate for the cases in which the private numbers are belong to a small sparse set of which the cardinality is not very large. The first protocol can be implemented with any probabilistic encryption. Therefore one can chooses the most efficient encryption cryptosystem in practice. We use the well-accepted simulation paradigm to prove that all the protocols are secure in the semi-honest model. The analysis shows that these schemes are efficient. The experimental result also demonstrates that our protocols are efficient and practical.

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