



## 一种高效的百万富翁问题协议及其应用

张 静,何 铮,葛炳辉,汤永利,叶 青

(河南理工大学 计算机科学与技术学院,河南 焦作 454000)

**摘 要:** 百万富翁问题是安全多方计算的基础问题,但现有解决方案计算复杂度且效率较低,在两数相等时无法进行精确比较。针对该问题,提出一种基于0-1编码的百万富翁问题协议。使用改进的0-1保密数据编码规则构建向量,利用ElGamal同态加密变体算法的同态性质,将百万富翁问题转化为向量中两元素求和的问题,同时,在半诚实模型下利用模拟范例证明协议的正确性与安全性,并将其应用于安全两方集合交集个数问题的求解。实验结果表明,与采用ElGamal和Paillier同态加密算法的协议相比,该协议计算复杂度更低且效率更高,可在两数相等时进行准确对比。

**关键词:** 安全多方计算;百万富翁问题;0-1编码;同态加密;集合交集个数

开放科学(资源服务)标志码(OSID):



**中文引用格式:** 张静,何铮,葛炳辉,等.一种高效的百万富翁问题协议及其应用[J].计算机工程,2021,47(2):168-175.  
**英文引用格式:** ZHANG Jing, HE Zheng, GE Binghui, et al. An efficient protocol for millionaires' problem and its application[J]. Computer Engineering, 2021, 47(2): 168-175.

## An Efficient Protocol for Millionaires' Problem and Its Application

ZHANG Jing, HE Zheng, GE Binghui, TANG Yongli, YE Qing

(College of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, Henan 454000, China)

**[Abstract]** The existing solutions to the Millionaires' Problem (MP), a basic problem in Secure Multi-Party Computation (SMC), have high computational complexity and low efficiency, and the two numbers can not be compared accurately when they are equal. To solve the problems, this paper proposes a protocol for MP based on 0-1 coding. The improved 0-1 secret data coding rule is used to construct the vector. By using the homomorphic property of the ElGamal homomorphic encryption variant algorithm, the MP is transformed into the sum of two elements in the vector. In the semi-honest model, the simulation examples are used to prove the correctness and security of the protocol, and the protocol is applied to solving the number of intersection sets of two secure parties. Experimental results show that compared with the protocols using ElGamal and Paillier homomorphic encryption algorithms, the proposed protocol has lower computational complexity and higher efficiency, and the two numbers can be compared accurately when they are equal.

**[Key words]** Secure Multi-Party Computation (SMC); Millionaires' Problem (MP); 0-1 coding; homomorphic encryption; number of set intersections

**DOI:** 10. 19678/j. issn. 1000-3428. 0057131

### 0 概述

安全多方计算 (Secure Multi-Party Computation, SMC) 是指两个及两个以上的参与者在泄露各自隐私数据的情况下,利用隐私数据进行保密计算并共同完成某项计算任务。SMC 可满足人们利用隐私数据进行保密计算的需求,同时兼顾数据的保密

性与共享性,因此被广泛应用于机器学习<sup>[1]</sup>、数据分析<sup>[2]</sup>、社交网络<sup>[3]</sup>以及医疗信息等领域。

百万富翁问题 (Millionaires' Problem, MP) 是安全多方计算中的基本问题,其在1982年由YAO提出<sup>[4]</sup>后引起多方关注。近年来,研究人员相继提出多种解决该问题的方法。文献[5]将安全多方计算规约到智力游戏中,利用混淆电路解决百万富翁问

**基金项目:** 国家自然科学基金(61802117);河南省高等学校重点科研项目(18B520018, 19A520025);河南理工大学创新型科研团队支持计划(T2018-1)。

**作者简介:** 张 静(1978—),女,副教授、博士,主研方向为网络与信息安全、安全多方计算;何 铮、葛炳辉,硕士研究生;汤永利,教授、博士;叶 青,讲师、博士。

**收稿日期:** 2020-01-06 **修回日期:** 2020-02-19 **E-mail:** yeqing@hpu.edu.cn

题。文献[6]采用不经意传输工具对两方输入进行双重加密,设计一种解决百万富翁问题的安全双方计算协议。文献[7]使用不经意传输工具并通过简单异或运算解决百万富翁问题。文献[8-9]借助茫然第三方提出一种安全的百万富翁比较协议,解决第三方合谋问题。文献[10]利用零知识证明构造一种百万富翁问题协议。文献[11-13]通过私有置换操作提出基于卡片的密码协议,解决了百万富翁问题。文献[14-15]利用对称密码解决恶意模型下的百万富翁问题。

利用编码是解决百万富翁问题的有效措施之一。文献[16]采用0-1编码将双方待比较的数据转化为0/1集合,结合具有乘法同态性的加密算法解决百万富翁问题,但其计算复杂度较高且无法精确区分两数相等的情况。文献[17]利用基于二次剩余困难问题的GM加密算法,通过构造0-1编码将数据编码转换为向量,提出一种基于几何方法的有理数比较协议,但GM算法在解密过程中的时间开销随二次非剩余集合增大呈线性增长。文献[18]使用文献[16]中编码方式提出一种基于DDH假设的方案,但该方案仅适用于输入较小数据,当两个待比较数据较大时计算开销较高。文献[19]结合1-r编码方式和ELGamal同态加密算法解决数据比较问题,提出一种数据比较协议并将其应用于保密数据排序。文献[20-21]提出0-1-2编码方法,同时利用Paillier同态加密算法将百万富翁问题转化为向量问题求解。虽然文献[19-21]提出的方法能有效解决百万富翁问题中的两数相等问题,但计算效率均较低。

本文提出一种采用0-1编码的百万富翁问题协议。通过改进保密数据编码规则,利用ElGamal同态加密变体算法解决安全两数比较问题,在半诚实模型下证明协议的正确性和安全性,并从理论和实验两个角度对协议的计算复杂度与效率进行分析。

## 1 基础知识

### 1.1 安全多方计算的安全模型

在安全多方计算协议的执行过程中,半诚实参与者在忠实履行协议的同时会保留协议执行过程中的有效信息,并尝试推导出其他参与方的隐私信息。若安全多方计算协议中参与者均为半诚实参与者,则称该协议使用的计算模型为半诚实模型。由于本文协议的计算模型均为半诚实模型,因此以下文给出半诚实模型下两方计算模型的安全性定义。

设Alice和Bob分别拥有隐私数据 $x, y$ ,  $\pi$ 为计算函数 $f$ 的协议。Alice和Bob希望通过合作计算函数 $F: (x, y) \rightarrow (f_1(x, y), f_2(x, y))$ 且不泄露各自隐私数据,其中存在概率多项式函数 $f: \{0, 1\}^* \times \{0, 1\}^* \rightarrow \{0, 1\}^* \times \{0, 1\}^*$ ,  $f_1(x, y)$ 和 $f_2(x, y)$ 分别为Alice和Bob计算所得函数 $F$ 的两个分量。将Alice和Bob在执行 $P$ 协议过程中得到的消息序列分别记为 $\text{view}_1^\pi(x, y)$ 和 $\text{view}_2^\pi(x, y)$ ,其所得输出结果分别记为 $\text{output}_1^\pi(x, y)$ 和 $\text{output}_2^\pi(x, y)$ 。

**定义1**(半诚实参与者的保密性) 对于函数 $f$ ,如果存在概率多项式算法 $S_1$ 和 $S_2$ (也称为模拟器)满足式(1)与式(2),则称其为 $\pi$ 保密计算函数。

$$\left\{ S_1(x, f_1(x, y)), f_2(x, y) \right\}_{x, y} \stackrel{c}{=} \left\{ \text{view}_1^\pi(x, y), \text{output}_2^\pi(x, y) \right\}_{x, y} \quad (1)$$

$$\left\{ f_1(x, y), S_2(x, f_2(x, y)) \right\}_{x, y} \stackrel{c}{=} \left\{ \text{output}_1^\pi(x, y), \text{view}_2^\pi(x, y) \right\}_{x, y} \quad (2)$$

其中, $\stackrel{c}{=}$ 表示计算上不可区分。若要证明一个安全多方计算协议具备安全性,则需构造模拟器 $S_1$ 和模拟器 $S_2$ 使式(1)与式(2)成立。

### 1.2 ELGamal同态加密变体算法

ELGamal同态加密<sup>[22]</sup>变体算法如下:

1) 密钥生成(KeyGen)。给定安全参数 $k$ ,定义 $k$ 特别大,采用密钥生成算法生成1个大小为 $k$ 比特的素数 $p$ 以及有限域 $\mathbb{Z}_p^*$ 的1个生成元 $g$ ,随机选取 $x$ 作为私钥,其对应公钥 $h = g^x \bmod p$ 。

2) 加密阶段(Enc)。给定消息 $M \in \mathbb{Z}_p^*$ ,选择随机数 $r$ ,密文 $E(M) = (c_1, c_2) = (g^r \bmod p, g^M h^r \bmod p)$ 。

3) 解密阶段(Dec)。将密文 $E(M)$ 解密为 $g^M = c_2 \cdot c_1^{-x} \bmod p$ ,对明文 $m_1$ 和 $m_2$ 加密后得到:

$$E(m_1) = (c_1, c_2) = (g^r \bmod p, g^{m_1} h^r \bmod p) \quad (3)$$

$$E(m_2) = (c_1, c_2) = (g^r \bmod p, g^{m_2} h^r \bmod p) \quad (4)$$

由于存在以下关系式:

$$\begin{aligned} E(M_1) \times E(M_2) &= (g^{r_1} \bmod p, g^{M_1} h^{r_1} \bmod p) \times \\ &\quad (g^{r_2} \bmod p, g^{M_2} h^{r_2} \bmod p) = \\ &\quad (g^{r_1+r_2} \bmod p, g^{M_1+M_2} h^{r_1+r_2} \bmod p) = \\ &\quad E(M_1+M_2) \end{aligned} \quad (5)$$

$$\begin{aligned} E(M_1)^b &= \left( (g^{r_1} \bmod p)^b, (g^{M_1} h^{r_1} \bmod p)^b \right) = \\ &\quad (g^{r_1 b} \bmod p, g^{M_1 b} h^{r_1 b} \bmod p) = E(bM_1) \end{aligned} \quad (6)$$

因此, ELGamal 同态加密变体算法满足如下性质:

$$E(m_1)E(m_2)=E(m_1+m_2) \quad (7)$$

$$E(m_1)^b=E(bm_1) \quad (8)$$

## 2 MP 解决方案

### 2.1 改进的 0~1 编码规则

百万富翁问题的实质是在保密情况下比较两数大小, 即 Alice 有 1 个隐私数据  $x$ , Bob 有 1 个隐私数据  $y$ , 两人在不泄露  $x$  和  $y$  大小的前提下合作计算并比较两数大小。为解决该问题, 本文将文献[20]中保密数据编码规则改进为 0-1 编码规则, 并利用该规则构造基于 0-1 编码的百万富翁问题协议。

0-1 编码规则如下:

设  $x, y \in \{v_1, v_2, \dots, v_n\} = U$ , 其中  $v_1 < v_2 < \dots < v_n$ 。

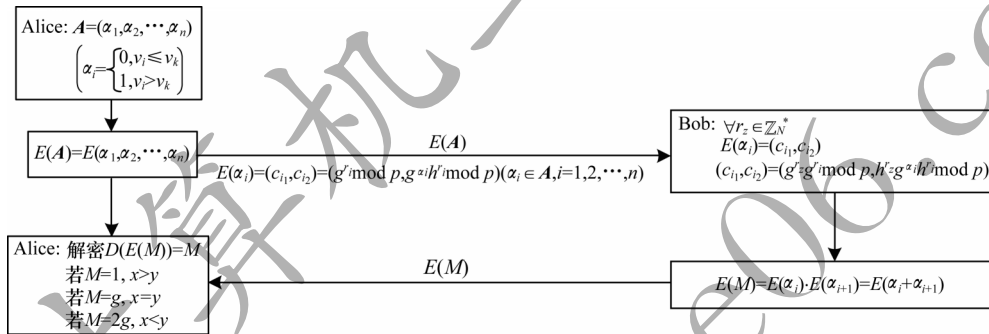


图1 基于0-1编码规则的MP协议框架

Fig.1 MP protocol framework based on 0-1 coding rule

基于0-1编码规则的MP协议实现流程如下:

#### 协议1 基于0-1编码的MP协议

输入 Alice:  $x$ , Bob:  $y$ ,  $x, y \in \{v_1, v_2, \dots, v_n\} = U$

输出  $M$

1. Alice:  $x, U \rightarrow A = (\alpha_1, \alpha_2, \dots, \alpha_n) // \alpha_i \in \{0, 1\} (i=1, 2, \dots, n)$

2. Alice  $\rightarrow$  Bob:  $E(A) = (E(\alpha_1), E(\alpha_2), \dots, E(\alpha_n))$

3. Bob:  $\forall r_i \in \mathbb{Z}_N^*, E(\alpha_i) = (c_{i1}, c_{i2}) = (g^{r_i} \bmod p, h^{r_i} g^{\alpha_i} \bmod p)$

//  $a_i$  为 Bob 在  $U$  中位置所对应  $A$  的值;  $E(M) = E(\alpha_i) \times E(\alpha_{i+1})$

//  $E(\alpha_{i+1}) = (g^{r_{i+1}} \bmod p, h^{r_{i+1}} g^{\alpha_{i+1}} \bmod p)$

4. Bob  $\rightarrow$  Alice:  $E(M)$

5. Alice:  $D(E(M)) = M$

该协议具体内容如下:

1) Alice 按照式(9)中的规则利用自身隐私数据  $x$  和隐私数据集  $U = \{v_1, v_2, \dots, v_n\}$  构造 1 个只含 0 和 1 的向量  $A = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 。

2) Alice 根据 ELGamal 同态加密算法生成公私钥对  $(pk, sk)$ , 选取不同的随机数  $r_i$ , 利用公钥  $pk$  将向量  $A$  中各

令  $x = v_k$  且  $y = v_l (1 \leq k, l \leq n)$ , 根据  $x$  与  $U$  构建只含 0 和 1 的向量  $A = (\alpha_1, \alpha_2, \dots, \alpha_n)$ , 具体规则如下:

$$\alpha_i = \begin{cases} 0, & v_i \leq v_k, i = 1, 2, \dots, n \\ 1, & v_i > v_k, i = 1, 2, \dots, n \end{cases} \quad (9)$$

根据  $y = v_l$  在  $U$  中的位置计算  $M = \alpha_l + \alpha_{l+1}$ 。若  $M = 0$ , 则  $k > l$ , 即  $x > y$ ; 若  $M = 1$ , 则  $k = l$ , 即  $x = y$ ; 若  $M = 2$ , 则  $k < l$ , 即  $x < y$ 。

### 2.2 基于0-1编码的MP协议

本文利用 0-1 编码规则提出一种解决百万富翁问题的协议, 以下介绍 MP 协议的具体设计方案, 并对其正确性与安全性进行分析。

#### 2.2.1 设计方案

MP 协议利用 0-1 编码规则将判断隐私数据  $x$  与  $y$  大小的问题归约到求解  $\alpha_i + \alpha_{i+1}$  值的问题, 主要通过两元素之和来判断两数大小, 该协议框架如图 1 所示。

个元素分别加密得到  $E(A) = (E(\alpha_1), E(\alpha_2), \dots, E(\alpha_n))$ ,  $E(\alpha_i) = (c_{i1}, c_{i2}) = (g^{r_i} \bmod p, g^{\alpha_i} h^{r_i} \bmod p)$ , 其中  $\alpha_i \in A$ ,  $i = 1, 2, \dots, n$ , 并将  $E(A)$  发送给 Bob。

3) Bob 根据  $y$  在隐私数据集  $U$  中的排列位置 (即  $a_i$  所在位置) 进行以下操作:

(1) 选取随机数  $r_i \in \mathbb{Z}_N^*$ , 计算  $E(\alpha_i) = (c_{i1}, c_{i2}) = (g^{r_i} \bmod p, h^{r_i} g^{\alpha_i} \bmod p)$ 。

(2) 计算  $E(M) = E(\alpha_i) E(\alpha_{i+1})$ 。

4) Bob 将  $E(M)$  发送给 Alice。

5) Alice 利用私钥  $sk$  对  $E(M)$  进行解密操作  $D(E(M))$  得到  $M$ 。若  $M = 1$ , 则  $x > y$ ; 若  $M = g$ , 则  $x = y$ ; 若  $M = g^2$ , 则  $x < y$ 。

#### 2.2.2 协议的正确性与安全性分析

本文对基于 0-1 编码规则的 MP 协议正确性与安全性进行分析。

**定理 1** 在半诚实模型下, 协议 1 是正确的, 同时也是安全的。

正确性证明:

1) Alice拥有密文信息 $E(\alpha_i) = (c_{i_1}, c_{i_2}) = (g^{r_i} \bmod p, g^{\alpha_i} h^{r_i} \bmod p)$ ,  $\alpha_i \in A, i = 1, 2, \dots, n$ 。

2) 基于ELGamal的同态加密变体算法具有加法同态性,计算公式如下:

$$\begin{aligned} E(M_1) \times E(M_2) &= (g^{r_1} \bmod p, g^{M_1} h^{r_1} \bmod p) \times \\ & (g^{r_2} \bmod p, g^{M_2} h^{r_2} \bmod p) = \\ & (g^{r_1+r_2} \bmod p, g^{M_1+M_2} h^{r_1+r_2} \bmod p) = \\ & E(M_1+M_2) \end{aligned} \quad (10)$$

3) Bob利用公钥对 $M = \alpha_i + \alpha_{i+1}$ 计算过程进行加密:

$$\begin{aligned} E(\alpha_i) E(\alpha_{i+1}) &= (g^{r_i} g^{r_{i+1}} \bmod p, h^{r_i} g^{\alpha_i} h^{r_{i+1}} \bmod p) \times \\ & (g^{r_{i+1}} \bmod p, g^{\alpha_{i+1}} h^{r_{i+1}} \bmod p) = \\ & (g^{r_i+r_{i+1}} \bmod p, \\ & g^{\alpha_i+\alpha_{i+1}} h^{r_i+r_{i+1}} \bmod p) = \\ & E(\alpha_i + \alpha_{i+1}) = E(M) \end{aligned} \quad (11)$$

4) Bob对 $E(M)$ 进行解密:

$$\begin{aligned} D(E(M)) &= \frac{g^{\alpha_i+\alpha_{i+1}} h^{r_i+r_{i+1}}}{(g^{r_i+r_{i+1}})^x} \bmod p = \\ & \frac{g^{\alpha_i+\alpha_{i+1}} (g^x)^{r_i+r_{i+1}}}{(g^{r_i+r_{i+1}})^x} \bmod p = \\ & g^{\alpha_i+\alpha_{i+1}} \bmod p \end{aligned} \quad (12)$$

5) 由于选取的安全参数 $k$ 很大,因此生成大小为 $k$  bit的 $p$ 很大,生成元 $g$ 非常小且满足 $g^M \bmod p = g^M$ 。针对计算得到的 $g^M$ :当 $g^M = 1$ 时, $M = 0$ , $y$ 位置在 $x$ 左侧, $x > y$ ;当 $g^M = g$ 时, $M = 1$ , $y$ 位置与 $x$ 位置相同, $x = y$ ;当 $g^M = g^2$ , $M = 2$ , $y$ 位置在 $x$ 右侧, $x < y$ 。

安全性证明:

Alice根据自身隐私数据 $x$ 和双方的共有集合 $U = \{v_1, v_2, \dots, v_n\}$ 构建向量 $A = (\alpha_1, \alpha_2, \dots, \alpha_n)$ ,其中 $\alpha_i \in \{0, 1\}, i = 1, 2, \dots, n$ 。Alice拥有公钥 $pk_A$ 与私钥 $sk_A$ ,在计算每个 $E(\alpha_i) (i = 1, 2, \dots, n)$ 时,其对利用ELGamal同态加密算法选取的不同 $r$ 进行加密操作,即 $E(\alpha_i) = (c_{i_1}, c_{i_2}) = (g^{r_i} \bmod p, g^{\alpha_i} h^{r_i} \bmod p)$ ,其中 $\alpha_i \in A, i = 1, 2, \dots, n$ ,由于选取的 $r$ 不同造成密文不同,因此对于加密后的向量 $E(A) = (E(\alpha_1), E(\alpha_2), \dots, E(\alpha_n))$ ,Bob无法通过解密从中获取有用信息;Bob自身隐私数据 $y$ 在集合 $U$ 中位置为已知,其在向量 $A$ 中对应的位置 $\alpha_i$ 不变,Bob为混淆密文选取随机 $r_z \in \mathbb{Z}_N^*$ ,若其利用 $\alpha_i$ 直接计算 $E(M) = E(\alpha_i) E(\alpha_{i+1})$ ,则Alice可通过列举方式计

算每两个元素相乘的密文值 $E(M') = E(\alpha_i) E(\alpha_{i+1})$ ,当 $E(M') = E(M)$ 时,可确定Bob的隐私数据 $y$ 在集合 $U$ 中的位置,从而造成其隐私数据泄露。因此,双方在整个过程中均无法获得对方的隐私信息。以下通过构造模拟器 $S_1$ 和模拟器 $S_2$ 进一步证明协议的安全性。

1) 构造模拟器 $S_1$ 。

具体步骤如下:

(1) 模拟器 $S_1$ 接受输入 $(x, p(x, y))$ ,由 $p(x, y)$ 的值构造 $y'$ 且满足 $p(x, y') = p(x, y)$ ,并用 $x'$ 和 $y$ 进行模拟。根据 $x$ 与 $U$ 构建只含0和1的向量 $A = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 。

(2) 利用ELGamal同态加密算法选取不同 $r$ 对向量 $A = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 进行加密,得到:

$$\begin{aligned} E(A) &= (E(\alpha_1), E(\alpha_2), \dots, E(\alpha_n)) = \\ & ((g^{r_1} \bmod p, g^{\alpha_1} h^{r_1} \bmod p), \\ & (g^{r_2} \bmod p, g^{\alpha_2} h^{r_2} \bmod p), \dots, \\ & (g^{r_n} \bmod p, g^{\alpha_n} h^{r_n} \bmod p)) \end{aligned} \quad (13)$$

(3) 根据 $\alpha'_i$ 计算 $E(M') = E(\alpha'_i) E(\alpha'_{i+1})$ 。

(4) 对数据 $E(M')$ 进行解密得到 $M'$ 。

在协议1中, $\text{view}_1^\pi(x, y) = \{A, E(A), \text{令 } S_1(x, P(x, y)) = \{A, E(A), E(M'), P(x, y')\}$ ,由于 $P(x, y) = P(x, y')$ ,且协议计算所得值与模拟器计算所得值在计算上不可区分,即 $E(M)_{x,y} \stackrel{c}{=} E(M')_{x,y}$ ,同时ELGamal同态加密算法是语义安全的,因此 $\{S_1(x, P(x, y)), P(x, y)\}_{x,y} \stackrel{c}{=} \{\text{view}_1^\pi(x, y), \text{output}_2^\pi(x, y)\}_{x,y}$ 成立。

2) 构造模拟器 $S_2$ 。

具体步骤如下:

(1) 模拟器 $S_2$ 接受输入 $(p(x, y), y)$ ,根据 $p(x, y)$ 的值构造 $x'$ 且满足 $p(x', y) = p(x, y)$ ,并用 $x'$ 和 $y$ 进行模拟。根据 $x'$ 与 $U'$ 构建只含0和1的向量 $A' = (\alpha'_1, \alpha'_2, \dots, \alpha'_n)$ 。

(2) 利用ELGamal同态加密算法选取不同 $r'$ 对向量 $A' = (\alpha'_1, \alpha'_2, \dots, \alpha'_n)$ 进行加密,得到:

$$\begin{aligned} E(A') &= (E(\alpha'_1), E(\alpha'_2), \dots, E(\alpha'_n)) = \\ & ((g^{r'_1} \bmod p, g^{\alpha'_1} h^{r'_1} \bmod p), \\ & (g^{r'_2} \bmod p, g^{\alpha'_2} h^{r'_2} \bmod p), \dots, \\ & (g^{r'_n} \bmod p, g^{\alpha'_n} h^{r'_n} \bmod p)) \end{aligned} \quad (14)$$



(3) 根据  $\alpha'_l$  计算  $E(M') = E(\alpha'_l) E(\alpha'_{l+1})$ 。

(4) 对数据  $E(M')$  进行解密操作, 得到  $M'$ 。

在协议1中,  $\text{view}_2^\pi(x, y) = \{E(A), E(M), P(x, y)\}$ , 令  $S_2(P(x, y), y) = \{E(A'), E(M'), P(x', y)\}$ , 由于  $P(x, y) = P(x', y)$ , 且协议计算所得值与模拟器计算所得值在计算上不可区分, 即  $E(M)_{x, y} \stackrel{c}{=} E(M')_{x, y}$ , 同时 ElGamal 同态加密算法是语义安全的, 因此  $\{S_1(x, P(x, y)), P(x, y)\}_{x, y} \stackrel{c}{=} \{\text{view}_1^\pi(x, y), \text{output}_2^\pi(x, y)\}_{x, y}$ 。

### 3 协议性能分析

#### 3.1 计算效率分析

对协议1和文献[16, 19, 21]协议的计算复杂度、通信轮数以及适用范围进行对比。由于文献[16, 19]协议基于 ElGamal 同态加密算法, 文献[21]协议基于 Paillier 同态加密算法, 本文协议基于 ElGamal 同态加密变体算法, 因此为便于对比分析, 令 Paillier 同态加密算法模为  $N$ , 数据长度为  $n$ , ElGamal 同态算法及其变体算法的模为  $P$ , 比较结果如表1所示。

表1 3种协议的不同性能对比

Table 1 Performance comparison of three protocols			
协议	计算复杂度	通信轮数	适用范围
协议1	$(2n+1) \text{ lb } P+3$	3	$>, <, =$
文献[16]协议	$5n \text{ lb } P+4n-6$		$>, <$
文献[19]协议	$(2n+3) \text{ lb } P+2$		$>, <, =$
文献[21]协议	$2(n+1) \text{ lb } N+l$		$>, <, =$

从上述3种协议的计算复杂性来看, 协议1、文献[16, 19]协议都是基于 ElGamal 同态加密, 而采用 ElGamal 同态加密算法进行1次加密需  $2 \text{ lb } P$  次模乘计算, 进行1次解密需  $1 \text{ lb } P$  次模乘计算。协议1中 Alice 进行  $n$  次加密和1次解密, Bob 进行3次模乘计算, 总计算开销为  $(2n+3) \text{ lb } P+2$  次模乘计算; 文献[16]协议需  $n$  次加密、 $n$  次解密和  $2n \text{ lb } P+4n-6$  次模乘计算, 总计算开销为  $5n \text{ lb } P+4n-6$  次模乘计算。文献[19]协议中 Alice 需  $n$  次加密和1次解密, Bob 需1次加密和2次模乘计算, 总计算开销为  $(2n+3) \text{ lb } P+2$  次模乘计算。文献[21]协议需  $n$  次加密、1次解密与  $l$  次模乘计算, 由于该协议是基于 Paillier 同态加密, 因此由 Paillier 同态加密算法中每次加密和解密需  $2 \text{ lb } N$  次模乘计算可知, 文献[21]协议的总计算开销为  $2(n+1) \text{ lb } N+l$  次

模乘计算。

从通信轮数来看, 协议1中 Alice 将密文  $E(A)$  发送给 Bob, Bob 将处理后的密文  $E(M)$  反馈给 Alice, Alice 对其解密后得到结果  $M$  并告知 Bob, 在整个协议执行过程中双方共通信3次, 因此, 协议1通信轮数为3, 其他3种协议的通信轮数与协议1相同。

由上述分析可知, 虽然协议1的通信轮数与其他协议相同, 但是其计算复杂度较其他协议更低。此外, 与文献[16]协议相比, 协议1可更好地比较两数据相等的问题。因此, 协议1整体计算效率优于其他协议。

#### 3.2 计算耗时分析

将协议1和文献[16, 19, 21]协议加密和解密的计算耗时进行对比。实验采用 Windows10 64 位操作系统, Inter® Core™ i7-4720HQ 2.60 GHz CPU, 8 GB 内存以及 MyEclipse 2017CI 编译环境。假设上述协议中双方商定向量元素个数  $n=100$ , 令 Paillier 同态加密算法与 ElGamal 同态加密算法中模数相同, 则在模数为 128 bit、256 bit、512 bit 和 1 024 bit 时分别计算这2种同态加密算法加密和解密的耗时, 结果如表2所示。

表2 2种算法在不同模数下加密和解密的耗时

Table 2 Encryption and decryption time consumption of two algorithms under different modulus ms					
算法	操作	不同模数下的耗时			
		128 bit	256 bit	512 bit	1 024 bit
Paillier	加密	0.271	0.317	0.404	1.066
	解密	0.394	0.591	0.814	2.140
ELGamal	加密	0.124	0.128	0.241	0.398
	解密	0.185	0.317	0.883	3.310

由表2可计算得到这2种算法加密和解密的平均耗时, 结合3.1节中对4种协议的效率分析可得到各协议在不同模数下的时间开销, 结果如图2所示(文献[21]协议中  $l$  值取决于 Bob 的隐私数据在商定向量中的位置,  $l$  取值范围为  $[1, 100]$ , 由于实验假定所商定向量的长度  $n=100$ , 为便于对比, 设定  $l=50$ )。由图2可以看出: 文献[16]协议的时间开销最高, 协议1与文献[19]协议的时间开销最低且两者很接近。对协议1与文献[19]协议在不同模数下的时间开销进行对比, 结果如表3所示。可以看出, 协议1在不同模数下的耗时均低于文献[19]协议。由上述分析可知, 协议1时间开销低于其他协议, 此结果与协议计算效率分析结果一致。

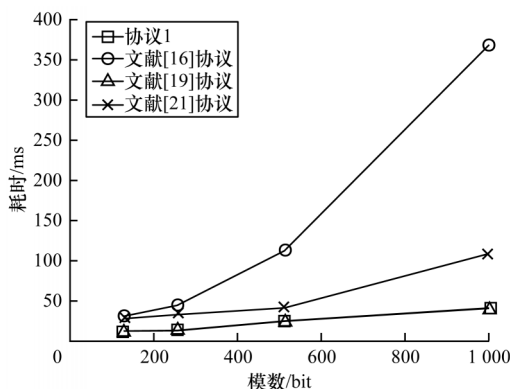


图2 4种协议在不同模数下的时间开销

Fig.2 Time cost of four protocols under different modulus

表3 协议1与文献[19]协议在不同模数下的时间开销

Table 3 Time cost of the protocol 1 and the protocol in reference [19] under different modulus

协议	模数/bit	耗时/ms
协议 1	128	12.585
	256	13.117
	512	24.983
	1 024	43.110
文献[19]协议	128	12.709
	256	13.245
	512	25.224
	1 024	43.508

## 4 集合交集的势

### 4.1 问题描述

当前社交网络已深入人们的日常生活,为扩大用户交友范围,云服务器会向每个用户推荐可能认识的好友,其推荐时判断依据为用户之间相同好友数量。然而在服务器与用户交互过程中,服务器在精准计算不同用户之间相同好友数量的同时,还要保证用户隐私不被泄露。如果将计算相同好友数量的过程视为安全两方的计算问题,则该问题可转化为求解安全两方集合交集个数的问题,即:Alice拥有隐私数据集  $W_1 = \{x_1, x_2, \dots, x_n\}$ , Bob拥有隐私数据集  $W_2 = \{y_1, y_2, \dots, y_m\}$ , Alice与Bob在不泄露自身隐私数据集情况下求解两集合交集的势。

### 4.2 协议设计

本文结合协议1,设计出求解保密集合交集势的协议,其具体内容如下:

1) Alice和Bob利用自身隐私数据集与共有隐私数据集  $U = \{v_1, v_2, \dots, v_z\}$  ( $z \geq m+n$ ) 构造0-1编码向量  $A = (a_1, a_2, \dots, a_z)$  和  $B = (b_1, b_2, \dots, b_z)$ , 编码规则如下:

$$a_i = \begin{cases} 1, & v_i \in U \\ 0, & v_i \notin U \end{cases}, i = 1, 2, \dots, z \quad (15)$$

$$b_i = \begin{cases} 1, & v_i \in U \\ 0, & v_i \notin U \end{cases}, i = 1, 2, \dots, z \quad (16)$$

2)  $(G, D, E)$  为 ElGamal 同态加密算法,  $k$  为设定的安全参数, Alice 运行  $G(k)$  生成算法的公私钥对  $(pk, sk)$ 。Alice 用公钥  $pk$  加密向量  $A$  得到  $E(A) = (E(a_1), E(a_2), \dots, E(a_z))$ , 并将  $E(A)$  发送给 Bob。

3) Bob 计算  $E(M) = (E(a_1)^{b_1} E(a_2)^{b_2} \dots E(a_z)^{b_z})$ , 并将其发送给 Alice。

4) Alice 通过私钥  $sk$  对  $E(M)$  进行解密得到数据  $\omega = g^M$ , 两集合交集的势  $M = \log_g \omega$ 。

求解保密隐私数据集交集势的协议实现流程如下:

协议2 求解保密隐私数据集交集势的协议

输入 Alice:  $W_1 = \{x_1, x_2, \dots, x_n\}$ ; Bob:  $W_2 = \{y_1, y_2, \dots, y_m\}$

输出  $M$

1. Alice, Bob:  $A, U \rightarrow A = (a_1, a_2, \dots, a_z)$ ;  $B, U \rightarrow B =$

$(b_1, b_2, \dots, b_z)$

2. Alice  $\rightarrow$  Bob:  $E(A) = (E(a_1), E(a_2), \dots, E(a_z))$

3. Bob  $\rightarrow$  Alice:  $E(M) = (E(a_1)^{b_1} E(a_2)^{b_2} \dots E(a_z)^{b_z})$

4. Alice:  $D(E(M)) = \omega / \omega = g^M, M = \log_g \omega$

定理2 在半诚实模型下, 协议2是正确的, 同时也是安全的。

正确性证明:

1) Alice 拥有密文信息:  $E(a_i) = (c_{i_1}, c_{i_2}) = (g^{r_i} \bmod p, g^{a_i h^{r_i}} \bmod p), i = 1, 2, \dots, z$ 。

2) 基于 ElGamal 加密变体算法具有加法同态性, 计算公式如下:

$$\begin{aligned} E(M_1)^b &= ((g^{r_1} \bmod p)^b, (g^{M_1 h^{r_1}} \bmod p)^b) = \\ & (g^{r_1 b} \bmod p, g^{M_1 b h^{r_1}} \bmod p) = \\ & E(bM_1) \end{aligned} \quad (17)$$

$$\begin{aligned} E(M_1) \times E(M_2) &= (g^{r_1} \bmod p, g^{M_1 h^{r_1}} \bmod p) \times \\ & (g^{r_2} \bmod p, g^{M_2 h^{r_2}} \bmod p) = \\ & (g^{r_1+r_2} \bmod p, g^{M_1+M_2 h^{r_1+r_2}} \bmod p) = \\ & E(M_1+M_2) \end{aligned} \quad (18)$$

3) Bob 利用编码后的隐私数据集  $B = (b_1, b_2, \dots, b_z)$  对  $E(a_i)$  进行加密:

$$\begin{aligned} \prod_{i=1}^z E(a_i)^{b_i} &= \prod_{i=1}^z ((g^{r_i} \bmod p)^{b_i}, (g^{a_i h^{r_i}} \bmod p)^{b_i}) = \\ & \prod_{i=1}^z (g^{r_i b_i} \bmod p, g^{a_i b_i h^{r_i}} \bmod p) = \\ & \prod_{i=1}^z E(a_i b_i) = E\left(\sum_{i=1}^z a_i b_i\right) = E(M) \end{aligned} \quad (19)$$

4) Bob 对  $E(M)$  进行解密:

$$D(E(M)) = \frac{g^M h^r}{(g^r)^x} \bmod p = \frac{g^M (g^x)^r}{(g^r)^x} \bmod p = g^M \bmod p = \omega \quad (20)$$

5) 由于选取的安全参数  $k$  很大, 因此生成大小为  $k$  比特的  $p$  很大, 生成元  $g$  非常小且满足  $g^M \bmod p = g^M$ 。对解密后  $D(E(M))$  的数据  $\omega$  进行计算得到  $M = \log_g \omega$ ,  $M$  即集合的势。

安全性证明:

采用模拟器  $S_1$  和模拟器  $S_2$  证明定理 2, 首先构造模拟器  $S_1$ 。

1) 模拟器  $S_1$  接受输入  $(x, p(x, y))$ , 由  $p(x, y)$  的值构造  $y'$  且满足  $p(x, y') = p(x, y)$ , 并用  $x'$  和  $y$  进行模拟。根据  $x$  与  $U$  构建只含 0 和 1 的向量  $A = (a_1, a_2, \dots, a_z)$ 。通过模拟器  $S_1$  构造向量  $B' = (b'_1, b'_2, \dots, b'_z)$ 。

2) 利用 ELGamal 同态加密算法选取不同的  $r$  对向量  $A = (a_1, a_2, \dots, a_z)$  进行加密, 得到:

$$E(A) = (E(a_1), E(a_2), \dots, E(a_z)) = ((g^{r_1} \bmod p, g^{a_1} h^{r_1} \bmod p), (g^{r_2} \bmod p, g^{a_2} h^{r_2} \bmod p), \dots, (g^{r_z} \bmod p, g^{a_z} h^{r_z} \bmod p)) \quad (21)$$

3) 根据  $B' = (b'_1, b'_2, \dots, b'_z)$  计算  $E(M')$ , 计算公式如下:

$$E(M') = (E(a_1)^{b'_1}, E(a_2)^{b'_2}, \dots, E(a_z)^{b'_z}) = ((g^{r_1} \bmod p, g^{a_1} h^{r_1} \bmod p)^{b'_1}, (g^{r_2} \bmod p, g^{a_2} h^{r_2} \bmod p)^{b'_2}, \dots, (g^{r_z} \bmod p, g^{a_z} h^{r_z} \bmod p)^{b'_z}) = ((g^{r_1 b'_1} \bmod p, g^{a_1 b'_1} h^{r_1 b'_1} \bmod p), (g^{r_2 b'_2} \bmod p, g^{a_2 b'_2} h^{r_2 b'_2} \bmod p), \dots, (g^{r_z b'_z} \bmod p, g^{a_z b'_z} h^{r_z b'_z} \bmod p)) = (E(a_1 b'_1), E(a_2 b'_2), \dots, E(a_z b'_z)) \quad (22)$$

4) 对数据  $E(M')$  进行解密得到  $M'$ 。

在协议 2 中,  $\text{view}_1^\pi(x, y) = \{A, E(A), E(M'), P(x, y)\}$ , 令  $S_1(x, P(x, y)) = \{A, E(A), E(M'), P(x, y)\}$ , 由于  $P(x, y) = P(x, y)$ , 且协议计算所得值与模拟器计算所得值在计算上不可区分, 即  $E(M) \stackrel{c}{=} E(M')_{x, y}$ , 同时 ELGamal 同态加密算法是语义安全的, 因此

$\{S_1(x, P(x, y)), P(x, y)\} \stackrel{c}{=} \{\text{view}_1^\pi(x, y), \text{output}_2^\pi(x, y)\}_{x, y}$  成立。

采用与上述类似的方法构造模拟器  $S_2$  得到  $\{S_1(x, P(x, y)) P(x, y)\} \stackrel{c}{=} \{\text{view}_1^\pi(x, y), \text{output}_2^\pi(x, y)\}_{x, y}$ 。

在协议 2 中, Alice 需执行  $n$  次加密、1 次解密和 1 次对数计算。Bob 需要执行  $z$  次模幂计算和  $z$  次模乘运算, 定义 Mul、Exp、lb 分别代表 1 次模乘计算、1 次模幂计算和 1 次对数计算。因此, 协议 2 的计算复杂度为  $((2n+1) \text{ lb } N + z) \text{ Mul} + z \text{ Exp} + 1 \times \text{lb}$ 。

在协议 2 中, Alice 将编码后的隐私数据集元素  $A$  进行加密, 将加密结果  $E(A)$  发送给 Bob, Bob 对  $E(A)$  进行处理得到  $E(M)$  并将其反馈给 Alice, Alice 对  $E(M)$  解密并向 Bob 公布结果。在整个协议执行过程中双方共通信 3 次, 因此通信轮数为 3。

## 5 结束语

百万富翁问题作为安全多方计算的基本模块, 常见于数据挖掘、数据查询和计算几何等方面, 而现有解决方案计算效率与安全性较低。本文提出一种基于 0-1 编码的百万富翁问题协议, 利用 ELGamal 同态加密的性质解决百万富翁问题, 在半诚实模型下证明其安全性, 并用协议求解安全两方集合交集个数。实验结果表明, 与采用 ElGamal 和 Paillier 同态加密算法的协议相比, 该协议计算效率更高。后续将在两数比较的基础上进行多数比较, 以解决带隐私保护的多集合交集问题。

## 参考文献

- [1] FRITCHMAN K, SAMINATHAN K, DOWSLEY R, et al. Privacy-preserving scoring of tree ensembles; a novel framework for AI in healthcare[C]//Proceedings of 2018 IEEE International Conference on Big Data. Washington D. C., USA: IEEE Press, 2018: 2413-2422.
- [2] SUNDARI S, ANANTHI M. Secure multi-party computation in differential private data with data integrity protection [C]//Proceedings of 2015 International Conference on Computing and Communications Technologies. Washington D. C., USA: IEEE Press, 2015: 180-184.
- [3] CUI Weirong, DU Chenglie, CHEN Jinchao. PSP: proximity-based secure pairing of mobile devices using WIFI signals [J]. Wireless Networks, 2019, 25 (2): 733-751.
- [4] YAO A C. Protocols for secure computations [C]//Proceedings of the 23rd IEEE Symposium on Foundations of Computer Science. Washington D. C., USA: IEEE Press, 1982: 160-164.
- [5] GOLDBREICH O, MICALI S, WIGDERSON A. How to play any mental game [C]//Proceedings of the 19th Annual ACM Symposium on Theory of Computing. New York, USA: ACM Press, 1987: 218-229.

- [ 6 ] YAO A C. How to generate and exchange secrets[C]// Proceedings of the 27th Annual Symposium on Foundations of Computer Science. Washington D. C. , USA; IEEE Press, 1986: 162-167.
- [ 7 ] IOANNIDIS I, GRAMA A. An efficient protocol for Yao' s millionaires' problem [C]//Proceedings of the 36th Annual Hawaii International Conference on System Sciences. Washington D. C. , USA; IEEE Press, 2003: 6-9.
- [ 8 ] QIN Jing, ZHANG Zhenfeng, FENG Dengguo, et al. A protocol of comparing information without leaking [J]. Journal of Software, 2004, 15(3): 421-427. (in Chinese) 秦静,张振峰,冯登国,等. 无信息泄露的比较协议[J]. 软件学报, 2004, 15(3): 421-427.
- [ 9 ] JIA Hengyue, WEN Qiaoyan, SONG Tingting, et al. Quantum protocol for millionaire problem [J]. Optics Communications, 2011, 284(1): 545-549.
- [ 10 ] JAWUREK M, KERSCHBAUM F, ORLANDI C. Zero-knowledge using garbled circuits: how to prove non-algebraic statements efficiently [C]//Proceedings of 2013 ACM SIGSAC Conference on Computer and Communications Security. New York, USA: ACM Press, 2013: 955-966.
- [ 11 ] NAKAI T, TOKUSHIGE Y, MISAWA Y, et al. Efficient card-based cryptographic protocols for millionaires' problem utilizing private permutations [C]//Proceedings of 2016 International Conference on Cryptology and Network Security. Berlin, Germany; Springer, 2016: 500-517.
- [ 12 ] MIYAHARA D, HAYASHI Y, MIZUKI T, et al. Practical and easy-to-understand card-based implementation of Yao' s millionaire protocol [C]//Proceedings of 2018 International Conference on Combinatorial Optimization and Applications. Berlin, Germany; Springer, 2018: 246-261.
- [ 13 ] ONO H, MANABE Y. Efficient card-based cryptographic protocols for the millionaires' problem using private input operations [C]//Proceedings of 2018 Asia Joint Conference on Information Security. Washington D. C. , USA; IEEE Press, 2018: 23-28.
- [ 14 ] MOHASSEL P, ROSULEK M, ZHANG Y. Fast and secure three-party computation; the garbled circuit approach [C]// Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. New York, USA: ACM Press, 2015: 591-602.
- [ 15 ] ZHAO Chuan, ZHAO Shengnan, ZHANG Bo, et al. Secure comparison under ideal/real simulation paradigm [J]. IEEE Access, 2018, 6(5): 31236-31248.
- [ 16 ] LIN H Y, TZENG W G. An efficient solution to the millionaires' problem based on homomorphic encryption [C]// Proceedings of 2005 International Conference on Applied Cryptology and Network Security. Berlin, Germany: Springer, 2005: 456-466.
- [ 17 ] LIU Xin, LI Shundong, LIU Jian, et al. Secure multiparty computation of a comparison problem [J]. SpringerPlus, 2016, 5(1): 1489-1497.
- [ 18 ] LIU M, NANDA P, ZHANG X, et al. Asymmetric commutative encryption scheme based efficient solution to the millionaires problem [C]//Proceedings of 2018 IEEE International Conference on Big Data Science and Engineering. Washington D. C. , USA; IEEE Press, 2018: 990-995.
- [ 19 ] LI Zhanli, CHEN Lichao, CHEN Zhenhua, et al. Design and applications of efficient protocol of millionaires' problem based on 1-r encoding [J]. Journal of Cryptologic Research, 2019, 6(1): 50-60. (in Chinese) 李占利,陈立朝,陈振华,等. 基1-r编码的高效百万富翁问题协议及应用[J]. 密码学报, 2019, 6(1): 50-60.
- [ 20 ] LI Shundong, ZUO Xiangjian, YANG Xiaoli, et al. Secure vector dominance protocol and its applications [J]. Acta Electronica Sinica, 2017, 45(5): 1117-1123. (in Chinese) 李顺东,左祥建,杨晓莉,等. 安全向量优势协议及其应用[J]. 电子学报, 2017, 45(5): 1117-1123.
- [ 21 ] ZUO Xiangjian, LI Shundong, YANG Xiaoli. An efficient homomorphic encryption based solution to millionaires' problem [J]. Journal of Chinese Computer Systems, 2017, 38(3): 455-459. (in Chinese) 左祥建,李顺东,杨晓莉. 同态加密的百万富翁问题高效解决方案[J]. 小型微型计算机系统, 2017, 38(3): 455-459.
- [ 22 ] FREEDMAN M J, HAZAY C, NISSIM K, et al. Efficient set intersection with simulation-based security [J]. Journal of Cryptology, 2016, 29(1): 115-155.