一个保护私有信息的多边形相交判定协议

罗永龙[1,2],黄刘生[1],徐维江[1],荆巍巍[1]
(1. 中国科学技术大学计算机科学技术系,安徽合肥 230027; 2. 安徽师范大学计算机科学系,安徽芜湖 241000)

摘 要: 安全多方计算是信息安全领域的研究热点问题之一.保护私有信息的多边形相交判定是一个特殊的安全多方计算问题,在军事、商业等领域有着重要的应用前景.现有多边形相交判定算法的主要操作是执行点积协议,而目前的点积协议在安全性和计算效率上均难以同时满足该判定算法的要求.本文首先设计了一个常数时间的线段相交判定协议,在此基础上提出了一个保护私有信息的判定多边形相交的概率算法;证明了该算法是一个蒙特卡洛偏真算法,理论分析与实验结果均表明,该方法性能优于现有算法.

关键词: 安全多方计算; 计算几何; 点积协议; 算法
中图分类号: TP309 **文献标识码**: A **文章编号**: 0372-2112 (2007) 04-0685-07

A Protocol for Privacy-Preserving Intersect-Determination of Two Polygons

LUO Yong-long[1,2], HUANG Liu-sheng[1], XU Wei-jiang[1], JING Wei-wei[1]
(1. *Department of Computer Science, University of Science and Technology of China, Hefei, Anhui* 230027, *China*;
2. *Department of Computer Science, Anhui Normal University, Wuhu, Anhui* 241000, *China*)

Abstract: At present, research on secure multi-party computation is of great interest in the field of information security. Privacy-preserving intersect-determination of two polygons is a special secure multi-party computation problem, it can be applied in many fields, such as military field and commerce field. Scalar products protocol plays an important role in the known methods of privacy-preserving intersect-determination of two polygons, however, the current scalar products protocols aren't fit for the determination algorithm on the security and the complexity at the same time. In this paper, a protocol for privacy-preserving intersect-determination of two line segments is developed and a probability algorithm for privacy-preserving intersect-determination of two polygons is presented. Both of the theoretical analysis and the experiment results show that the new algorithms are more efficient than the current algorithm.

Key words: secure multi-party computation; computational geometry; scalar product protocol; algorithm

1 引言

安全多方计算(Secure Multi-Party Computation,简称SMC)[1]研究一组互不信任的参与者之间保护私有信息的合作计算问题,对解决网络环境下的信息安全具有重要价值.该问题自A C Yao[2]首次提出以来,已经得到了较多的理论研究[3~6].由于受到计算效率的限制,用一般的理论方法来解决安全多方计算中的一些特殊实例是不现实的,对于一些特殊问题需要用特殊方法才能达到高效性[1].高效实用的安全多方计算协议已经成为目前的热门研究课题之一[7~11].

保护私有信息的多边形相交判定是一个特殊的SMC问题,在军事、商业等领域有着重要的应用[8,12].为解决该问题,Atallah等[12]首先抽象出一个点积问题,并分别基于茫然传送(Oblivious Transfer)协议和同态加密方案(Homomorphic Encryption Schemes)设计了两个不同的点积协议(Scalar Product Protocol).若多边形 p_1、p_2 均有 n 条边,则Atallah的判定算法需要执行 $4n^2$ 次点积协议,其时间复杂性与通信复杂性都很高,该文也指出其所提出的协议性能有待进一步改进.另一方面,点积协议自提出以来,被迅速广泛应用于保护私有信息的各类安全多方计算,如数据挖掘[13~15]、统计分析[16,17]等,并已成为SMC的一个基本协议.尽管目前已经设计出了很多不同的高效点积协议,但它们在计算效率与安全性上难以同时满足文献[12]提出的多边形相交判定算法.为此,本文首先设计了一个常数时间的线段相交判定协议,在此基础上提出了一个保护私有信息的判定多边形相交的概率算法,我们的方法在性能上优于现有的判定算法.

收稿日期:2005-03-07;修回日期:2006-11-12
基金项目:国家自然科学基金(No. 60573171);教育部博士点基金(No. 20060358014);中国博士后科学基金(No. 20060390700);安徽省高校自然科学研究重点项目(No. 2006KJ024A,No. 2007KJ043A);安徽省自然科学基金(No. 070412043)

2 预备知识

2.1 基本概念

定义 1(计算模型) 安全两方计算是一种安全的分布式计算协议,在该协议中,两个成员 Alice 与 Bob 分别持有各自的秘密输入 x 与 y,对某个给定的函数 f,他们希望协作计算函数值 $f(x,y)$,但 Alice 与 Bob 都不愿意向对方暴露自己的输入[1]. 在计算过程中,一般 Alice 与 Bob 首先对各自的输入进行伪装,映射成另一个数据. 然后两方在伪装后的数据上进行计算,最后将计算的中间结果还原成原问题的解.

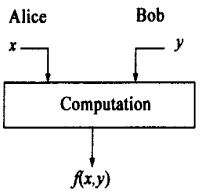


图 1 安全两方计算模型

图 1 描述了安全两方的计算模型,该模型很容易推广到多方计算情形.

定义 2(安全模型) 在安全两方计算中,假设以 I_A、I_B 分别表示 Alice 和 Bob 的输入实例,O_A、O_B 表示他们的输出结果,若用符号 C 表示两方所要进行的计算,则 $(O_A, O_B) = C(I_A, I_B)$. 一个能执行计算 C 的协议若满足如下两个条件则称作是安全的[1,17]:

(1) 存在一个无限集合 $D_A = \{(I_{Ai}, O_{Ai}) \mid i = 1,2,\cdots\}$, $\forall (I_A', O_A') \in D_A$, $(O_A', O_B) = C(I_A', I_B)$ 成立;

(2) 存在一个无限集合 $D_B = \{(I_{Bi}, O_{Bi}) \mid i = 1,2,\cdots\}$, $\forall (I_B', O_B') \in D_B$, $(O_A, O_B') = C(I_A, I_B')$ 成立.

定义 3(半信任方) 一个半信任方是指 SMC 中一个参与计算的用户,能严格执行协议的规程,不会中途强行退出或恶意掺入虚假数据,但在协议执行过程中他可能会保留所有能搜集到的关于对方的信息,以期望在协议结束后计算出对方的输入数据.

本文假设参与计算的双方都是半信任的.

定义 4 下面我们定义一组描述 SMC 协议的符号[9]:

①Ai:Alice 在本地执行协议的第 i 步,Bi:Bob 在本地执行协议的第 i 步;

②Ai|Bi:Alice 与 Bob 各自在本地执行协议的第 i 步;

③Ai ∧Bi:Alice 与 Bob 共同协作执行协议的第 i 步;

④*RANDOM SELECT*:选择某个随机数据;

⑤*GENERATE*:构造一个对象;

⑥*COMPUTE*:执行一个基本运算;

⑦*SEND*(*sender* →*receiver*, $s_1, s_2, \cdots, s_m$):发送方 sender 向接收方 receiver 发送 m 个消息 $s_1, s_2, \cdots, s_m$;

⑧*Protocol*_*Name*(parameter1,parameter2):双方调用某个子协议,parameter1 与 parameter2 分别为两方的输入.

2.2 两个基本协议

2.2.1 向量优先协议 VDP(Vector Dominance Protocol)

向量 $A = (a_1, a_2, \cdots, a_n)$ 优先于向量 $B = (b_1, b_2, \cdots, b_n)$ 是指对任意的 $i = 1,2,\cdots,n$, $a_i > b_i$ 都成立,此时记为 $A > B$. 而向量优先问题是指 Alice 输入一个私有向量 A,Bob 输入另一个私有向量 B,他们秘密判定 A 是否优先于 B,但双方都不能够知道对方的任何数据信息;当 A 不优先于 B 时,他们也不能够得到任意一对 a_i 与 b_i 的大小关系. 文献[12]设计了相应的计算协议,其通信复杂性为 $O(mn)$, m 表示 a_i 与 b_i 所需要的二进制位数.

2.2.2 点积协议 SPP(Scalar Products Protocol)

点积问题可以描述为:Alice 有一个私有向量 $X = (x_1, x_2, \cdots, x_n)$,Bob 有另一个私有向量 $Y = (y_1, y_2, \cdots, y_n)$,Alice 需要得到值 $u = X \cdot Y + v = \sum_{i=1}^{n} x_i y_i + v$,这里 v 是仅被 Bob 知道的随机数. 同时满足:(1) Alice 不能从 u 中得到 XY 的值,也不能从结果得到任何 y_i 的信息,(2) Bob 不能得到 u 的值,也不能得到任何 x_i 的信息.

近年来,点积协议得到了广泛的研究,基于不同程度的安全性与计算复杂性,目前已经提出了很多不同的点积协议. 文献[12]分别基于茫然传送协议和同态加密方案设计的两个点积协议泄露的信息量几乎为零(安全性极高),但它们的时间复杂性与通信复杂性都很高,因而并不实用. 文献[18]虽然提出了两个基于不可信第三方的点积协议,但由于无有效方法可以保证不可信第三方不会和参与计算的某一方串通,所以其安全性仍然较差. 文献[13～15,18]分别提出了一些实用的点积协议,它们都在不同程度上泄露了部分输入信息从而降低了时间复杂性和通信复杂性. 然而,对于一些特殊的输入,例如当点积协议中数据维数较小或数据取值范围较小时,这些协议就极不安全.

3 一个新的线段相交判定协议

判断两个多边形相交的一个重要步骤是判定它们是否存在一对边相交,因此线段相交判定是其中的关键算法,为此我们提出了一个新的线段相交判定协议.

3.1 问题描述

定义 5(线段表示) 假设 Alice 输入线段 l_1 的两个端点 (x_1, y_1) 和 (x_1', y_1'),Bob 输入线段 l_2 的两个端点 (x_2, y_2) 和 (x_2', y_2'). 我们把 l_1、l_2 对应的直线方程分别记为:

$$f(x,y) = a_1 x + b_1 y + c_1 = 0,\ g(x,y) = a_2 x + b_2 y + c_2 = 0$$

两条线段相交的充要条件是每条线段的两个端点分别位于另一条线段所在直线的两侧[19],文献[12]调

用点积协议分别判断 $f(x_2, y_2)$、$f(x_2', y_2')$、$g(x_1, y_1)$ 及 $g(x_1', y_1')$ 的符号,若双方都采用某个实用的点积协议,由于输入是一个二维数据(例如 Bob 输入 x_2 与 y_2,Alice 输入 a_1 与 b_1,他们协作计算 $f(x_2, y_2)$),则双方将都能够计算出对方的输入值,算法是不安全的. 故该算法只能采用基于茫然传送协议或基于同态加密方案的点积协议,算法性能很差.

若 (x_2, y_2) 与 (x_2', y_2') 在 $f(x, y)$ 的同一侧,

则 $\begin{cases} f(x_2, y_2) > 0 \\ f(x_2', y_2') > 0 \end{cases}$ 或 $\begin{cases} f(x_2, y_2) < 0 \\ f(x_2', y_2') < 0 \end{cases}$,

即 $f(x_2, y_2) f(x_2', y_2') > 0$,也就是

$$a_1^2 x_2 x_2' + b_1^2 y_2 y_2' + a_1 b_1 (x_2 y_2' + x_2' y_2) + a_1 c_1 (x_2 + x_2') + b_1 c_1 (y_2 + y_2') + c_1^2 > 0 \quad (1)$$

此时我们可以使用文献[18]介绍的基于可逆矩阵的点积协议判断式(1)是否成立,若成立,则 (x_2, y_2) 与 (x_2', y_2') 在 $f(x, y)$ 的同一侧,l_1 与 l_2 必定不相交,可以结束协议. 否则可以用类似方法继续判断 (x_1, y_1) 与 (x_1', y_1') 是否在 $g(x, y)$ 的同一侧.

3.2　线段相交判定协议

基于上述分析,我们提出的线段相交判定协议可描述如下:

Protocol 1 EIP(Edges-Intersect Protocol){

//Alice 输入线段 l_1,Bob 输入线段 l_2,他们判断 l_1 与 l_2 是否相交

A1| B1 :Alice *COMPUTE* the equation of $l_1: f(x, y) = a_1 x + b_1 y + c_1 = 0$;

Bob *COMPUTE* the equation of $l_2: g(x, y) = a_2 x + b_2 y + c_2 = 0$;

A2| B2 :Alice *GENERATE* $S_1 = (a_1^2, b_1^2, a_1 b_1, a_1 c_1, b_1 c_1, c_1^2)$;

Bob *GENERATE* $S_2 = (x_2 x_2', y_2 y_2', x_2 y_2' + x_2' y_2, x_2 + x_2', y_2 + y_2', 1)$;

A3 ∧B3 : *SPP*(S_1, S_2);

//Alice 与 Bob 执行点积协议,Alice 选取随机数 v_1,Bob 得到

// $u_1 = a_1^2 x_2 x_2' + b_1^2 y_2 y_2' + a_1 b_1 (x_2 y_2' + x_2' y_2) + a_1 c_1 (x_2 + x_2') + b_1 c_1 (y_2 + y_2') + c_1^2 + v_1$

A4 : *SEND*(Alice →Bob, v_1);

B5 :if $u_1 > v_1$ then{//即 $f(x_2, y_2) f(x_2', y_2') > 0$

Output l_1 与 l_2 不相交;结束协议;

} //endif

A6| B6 : Bob *GENERATE* $S_3 = (a_2^2, b_2^2, a_2 b_2, a_2 c_2, b_2 c_2, c_2^2)$;

Alice *GENERATE* $S_4 = (x_1 x_1', y_1 y_1', x_1 y_1' + x_1' y_1, x_1 + x_1', y_1 + y_1', 1)$;

A7 ∧B7 : *SPP*(S_3, S_4);

//Alice 与 Bob 执行点积协议,Bob 选取随机数 v_2,Alice 得到

// $u_2 = a_2^2 x_1 x_1' + b_2^2 y_1 y_1' + a_2 b_1 (x_1 y_1' + x_1' y_1) + a_2 c_2 (x_1 + x_1') + b_2 c_2 (y_1 + y_1') + c_2^2 + v_2$

B8 : *SEND*(Bob →Alice, v_2);

A9 :if $u_2 > v_2$ then {Output l_1 与 l_2 不相交;}

else{Output l_1 与 l_2 相交;} //endif

}// end of protocol

协议 1 的正确性和性能分析如下

定理 1(正确性)　协议 1 能正确判断两条线段是否相交.

证明　显然 $u_1 > v_1$ 蕴含式(1)成立,而

$$a_1^2 x_2 x_2' + b_1^2 y_2 y_2' + a_1 b_1 (x_2 y_2' + x_2' y_2) + a_1 c_1 (x_2 + x_2') + b_1 c_1 (y_2 + y_2') + c_1^2 > 0$$

$$\Leftrightarrow (a_1 x_2 + b_1 y_2 + c_1)(a_1 x_2' + b_1 y_2' + c_1) > 0$$

$$\Leftrightarrow \begin{cases} f(x_2, y_2) > 0 \\ f(x_2', y_2') > 0 \end{cases} 或 \begin{cases} f(x_2, y_2) < 0 \\ f(x_2', y_2') < 0 \end{cases}$$

即点 (x_2, y_2) 与 (x_2', y_2') 在直线 $f(x, y) = 0$ 的同一侧,此时可以判定两条线段必定不相交. 同理可证 $u_2 > v_2$ 时,点 (x_1, y_1) 与 (x_1', y_1') 必定位于 $g(x, y)$ 的同一侧,两条线段也不相交. 否则,每条线段的两端点必然分别位于另一条线段所在直线的两侧,故两线段相交.

因此协议 1 是正确的.

定理 2(复杂性)　协议 1 的时间复杂性及通信复杂性均为常数.

证明:协议 1 中所使用的点积协议 SPP 是基于可逆矩阵的协议,而该点积协议通信次数(轮复杂性)为 2,通信的总位数(位复杂性)为 $2nd$,计算的时间复杂性为 $O(n^2)$. 由于协议 1 在 A3 ∧B3 和 A7 ∧B7 中各调用了 SPP 一次,而 A4 与 B8 中各进行了一次通信,故通信总次数为 6,又因为两次调用 SPP 所用向量(S_1 与 S_2 及 S_3 与 S_4)的维数 n 是一个固定的常数值 6,所以协议 1 通信的总位数为常数. 同理,时间复杂性亦为常数.

定理 3(安全性)　Alice 与 Bob 在协议 1 中均不能够得到任何关于对方线段的信息.

证明:若 A3 ∧B3 中使用基于可逆矩阵的点积协议计算,Alice 与 Bob 分别输入 $(a_1^2, b_1^2, a_1 b_1, a_1 c_1, b_1 c_1, c_1^2)$ 与 $(x_2 x_2', y_2 y_2', x_2 y_2' + x_2' y_2, x_2 + x_2', y_2 + y_2', 1)$,协议执行结束后,尽管 Alice 在 A4 步向 Bob 传送了随机数 v_1,此时 Bob 能够知道关于 $(a_1^2, b_1^2, c_1^2, a_1 b_1, a_1 c_1, b_1 c_1)$ 的方程,但由于得到的是二次方程,Bob 也不能推导出 (a_1, b_1, c_1) 的具体信息. 同样,Alice 也不能够得到 (x_2, y_2, x_2', y_2') 的值. 根据安全模型的定义(定义 2),协议 1

是安全的.

4 多边形相交判定的概率算法

4.1 问题表示

定义 6(多边形表示) 不妨设两个多边形 p_1、p_2 均有 n 条边,则可将其表示为:

$p_1=\{(f_i,(x_{1i},y_{1i}),(x_{1i}',y_{1i}'))\mid f_i(x,y)=a_{1i}x+b_{1i}y+c_{1i},1\le i\le n\}$ 与 $p_2=\{(g_j,(x_{2j},y_{2j}),(x_{2j}',y_{2j}'))\mid g_j(x,y)=a_{2j}x+b_{2j}y+c_{2j},1\le j\le n\}$,

其中 $f_i(x,y)=0$ 表示 p_1 的第 i 条边对应的直线方程,(x_{1i},y_{1i}) 与 (x_{1i}',y_{1i}') 表示它的两个端点,p_2 中 $(g_j,(x_{2j},y_{2j}),(x_{2j}',y_{2j}'))$ 的意义类同. 更为一般地,两个多边形的边数不一定相等. 为描述简洁,本文假设 p_1、p_2 的边数均为 n,其分析很容易推广到一般情形.

定义 7(界限框) 一个多边形的界限框是指包含这个多边形的最小矩形,如图 2. 我们用矩形 $(\hat{p}_{11},\hat{p}_{12})$ 表示多边形 p_1 的界限框,其中 $\hat{p}_{11}=(\hat{x}_1,\hat{y}_1)$、$\hat{p}_{12}=(\hat{x}_2,\hat{y}_2)$ 分别表示界限框的左下角点与右上角点,这里

$$\begin{cases}\hat{x}_1=\min\{x_{1i}\mid 1\le i\le n\}\\ \hat{y}_1=\min\{y_{1i}\mid 1\le i\le n\}\end{cases},\quad\begin{cases}\hat{x}_2=\max\{x_{1i}\mid 1\le i\le n\}\\ \hat{y}_2=\max\{y_{1i}\mid 1\le i\le n\}\end{cases}\tag{2}$$

类似地,用 $(\hat{p}_{21},\hat{p}_{22})$ 表示 p_2 的界限框,其中 $\hat{p}_{21}=(\hat{x}_3,\hat{y}_3)$ 表示左下角点,$\hat{p}_{22}(\hat{x}_4,\hat{y}_4)$ 表示右上角点.

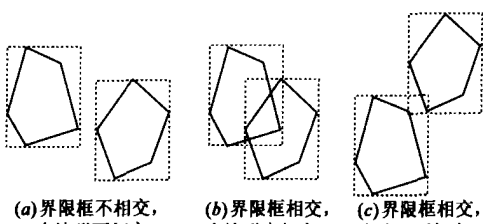
4.2 界限框测试

两个多边形不相交时,可能相距较远,我们可以用界限框测试预先排除一些特殊情况. 该步骤首先确定 p_1 与 p_2 的界限框,然后判断界限框是否相交. 如图 2(a)所示,若界限框不相交,则 p_1 与 p_2 必定也不相交. 而 p_1 与 p_2 的界限框相交当且仅当下述表达式为真:

$$(\hat{x}_2\ge\hat{x}_3)\wedge(\hat{x}_4\ge\hat{x}_1)\wedge(\hat{y}_2\ge\hat{y}_3)\wedge(\hat{y}_4\ge\hat{y}_1)\tag{3}$$

在 SMC 中式(3)可以通过向量优先协议计算.

界限框相交只是多边形相交的必要条件,但不是充分条件. 例如,图 2(c)所示的界限框相交,但两个多边形并不相交. 因此,当两个界限框相交时,还必须进行边的相交测试才能判断多边形是否相交.



(a)界限框不相交,多边形不相交　(b)界限框相交,多边形也相交　(c)界限框相交,多边形不相交

图 2 界限框测试的三种情况

4.3 多边形相交判定协议

若能找到 p_1 中某条边 $(f_i,(x_i,y_i),(x_i',y_i'))$ 与 p_2 中某条边 $(g_j,(x_j,y_j),(x_j',y_j'))$ 相交,我们就可以判定 p_1 与 p_2 相交. 假若不存在这样的一对边,则 p_1 与 p_2 必定不相交. 这里我们不考虑多边形相互包含的特殊情形.

为了在各对边的相交判定过程中不向对方泄露信息,我们采用概率算法来实现. Alice 与 Bob 每次从自己的边集中随机选取一条边,然后使用线段相交判定协议(协议 1)判断这对边是否相交. 若相交,则结束算法,否则继续判定,直到找到一对边相交或者重复次数达到某个阈值 k 为止.

根据上述分析,多边形相交判定协议可以用协议 2 描述如下:

Protocol2 (Polygons-Intersect Protocol) {

// Alice 输入多边形 p_1 的 n 个顶点,Bob 输入多边形的 n 个顶点

// Alice 与 Bob 共同判断 p_1 与 p_2 是否相交

A1| B1 :Alice *COMPUTE* $p_1=\{(f_i,(x_{1i},y_{1i}),(x_{1i}',y_{1i}'))\mid f_i(x,y)=a_{1i}x+b_{1i}y+c_{1i},1\le i\le n\}$;

Bob *COMPUTE* $p_2=\{(g_j,(x_{2j},y_{2j}),(x_{2j}',y_{2j}'))\mid g_j(x,y)=a_{2j}x+b_{2j}y+c_{2j},1\le j\le n\}$;

// 通过输入的顶点计算各条边对应的方程

A2| B2 :Alice *COMPUTE* $(\hat{p}_{11},\hat{p}_{12})=((\hat{x}_1,\hat{y}_1),(\hat{x}_2,\hat{y}_2))$;

GENERATE $S_1=(\hat{x}_2,-\hat{x}_1,\hat{y}_2,-\hat{y}_1)$;

Bob *COMPUTE* $(\hat{p}_{21},\hat{p}_{22})=((\hat{x}_3,\hat{y}_3),(\hat{x}_4,\hat{y}_4))$;

GENERATE $S_2=(\hat{x}_3,-\hat{x}_4,\hat{y}_3,-\hat{y}_4)$;

// Alice 与 Bob 分别由式(2)确定 p_1 与 p_2 的界限框

A3 ∧B3 : *VDP*(S_1,S_2); // Alice 与 Bob 用向量优先协议比较两个向量是否满足式(3)

if Not $(S_1>S_2)$ then {

output p_1 与 p_2 不相交;return false;

}// endif

A4 ∧B4 : $t\leftarrow 1$;

while $t\le k$ do{ // 边的相交测试,k 为预设的重复次数

A4 - 1| B4 - 1 :Alice *RANDOM SELECT* $(f_i,(x_{1i},y_{1i}),(x_{1i}',y_{1i}'))\in p_1$;

Bob *RANDOM SELECT* $(g_j,(x_{2j},y_{2j}),(x_{2j}',y_{2j}'))\in p_2$;

// Alice 与 Bob 分别在多边形中随机挑选边

A4 - 2 ∧B4 - 2 : EIP(f_i,g_j); // 用协议 1 判断 f_i 与 g_j 是否相交

if f_i 与 g_j 相交 then{

Output p_1 与 p_2 相交;结束协议;

}// endif

$t \leftarrow t+1$;

}// endwhile

Output p_1 与 p_2 不相交;return false;

}// end of protocol

5　协议的性能分析

5.1　理论分析

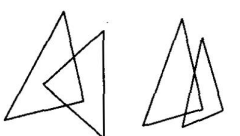
定理 4(正确性)　协议 2 所实现的多边形相交判定算法是一个 $1-\left(1-\frac{m}{n^2}\right)^k$ 正确、偏真的蒙特卡洛算法,$m \geq 2$.

证明　若将两个多边形 p_1 和 p_2 相交定义为真,不相交定义为假,则因为协议 2 返回真时,p_1 和 p_2 至少有一条相交的边,此时 p_1 和 p_2 必相交,因此结论正确. 但是,协议 2 返回假时,结论未必正确. 协议中第一处返回假是由于 p_1 和 p_2 所在的界限框不相交,导致了 p_1 和 p_2 必不相交,此时返回假的结论是正确的;然而,第二处返回假时,p_1 和 p_2 不相交的结论未必正确. 其原因是,若 p_1 和 p_2 相交,但在 while 循环中相交的边未必被随机选中,则算法返回的解是错误的. 因为仅当算法返回 false 时才有可能产生错误的解,故该算法是一个蒙特卡洛偏真(true-biased)算法[20].

算法的概率分析如下:

(1)若 p_1 与 p_2 不相交,则算法每次执行均返回 false,其解正确.

(2)若 p_1 与 p_2 相交,不妨设 p_1 与 p_2 有 m 个交点,即有 m 对边相交. 因为两个多边形相交至少有两对边相交,故 $m \geq 2$. 一般地,两对边相交会涉及 4 条边,如图 3(a)所示,但最少可以只涉及到 3 条边,其中一个多边形的一条边与另一多边形的两条边相交,如图 3(b)所示. 显然,相交边的数目越少时,算法出错的概率越大. 当从 p_1 和 p_2 中随机挑选一条边进行 EIP 测试时,其相交的概率至少是 $\frac{m}{n^2}$,而不相交(即出错)的概率至多为 $1-\frac{m}{n^2}$. 由此可知,k 次 EIP 判定出错的概率是 $\left(1-\frac{m}{n^2}\right)^k$,返回真的概率是 $1-\left(1-\frac{m}{n^2}\right)^k$[20]. 即协议 2 是一个偏真的 $1-\left(1-\frac{m}{n^2}\right)^k$ 正确的蒙特卡洛算法.



(a) 4 条边相交　(b) 3 条边相交

图 3　两多边形相交示例

(3)因为 $\left(1-\frac{m}{n^2}\right)^k \leq \left(1-\frac{2}{n^2}\right)^k$,当我们选取 $k=n^2$ 时,

由 $\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e$,

可得 $\lim_{x\to\infty}\left(1-\frac{2}{n^2}\right)^{n^2} = \frac{1}{e^2} \approx 0.135$,

也就是说,当 $k=n^2$ 时,协议是一个 0.865 正确的偏真的蒙特卡洛算法. 因此重复调用该协议 c 次,即可得到一个 $1-(1-0.865)^c$ 正确的算法[20]. 例如,若调用该协议 3 次,正确的概率将是 99.75 %. 故我们能得到如下推论:

推论 1　当 $k=cn^2$ 时,协议 2 出错的概率可以控制为 $\varepsilon \leq 0.135^c$,这里对于给定的 $0 \leq \varepsilon \leq 1$,因为 c 一般为常数,故协议 2 的时间复杂性与通信复杂性均为 $O(n^2)$.

定理 5(安全性)　Alice 与 Bob 不能够从协议 2 中推导出对方多边形的信息.

证明　(1)协议 2 调用了 k 次协议 1 来判断每次产生的一对边是否相交,由定理 3,该判定过程不会泄露信息. (2)由于每次测试的边对是由 Alice 与 Bob 随机选取的,尽管双方都能知道这对边是否相交,但由于不能够知道对方边的信息,因此也就不能够得到对方多边形的信息. 根据定义 2,协议 2 是安全的.

5.2　性能比较

我们将协议 2 的性能与文献[12]的两种协议进行对比,表 1 从通信的次数、通信的总位数及计算时间复杂性三方面给出了比较结果. 表中 n 为多边形的顶点数,d 表示每个数据所占的比特数,p 与 t 为安全参数①. 从表 1 可以看出,文献[12]通信的次数为 $4n^2t$,而我们的算法仅为 $3n^2$,尽管它们的数量级相同,但在分布式系统中,通信次数对系统的性能影响非常大,故我们的算法在性能上具有很大的改进.

表 1　算法性能比较

算法	通信次数	通信总位数	时间复杂性
基于同态加密算法	$4n^2t$	$4n^2td$	$O(n^2t)$ 次加(解)密运算
基于茫然传送算法	$4n^2t$	$4n^2ptd$	$O(n^2p)$ 次茫然传送协议
本文的概率算法	$3n^2$	$3n^2d$	$O(n^2)$ 次基本算术运算

5.3　实验方法与实验结论

(1)实验方法

Step1:随机选取两个多边形 p_1、p_2,边数 $n=10$,交点数 $m=2$,分别设定 $k=100,110,\cdots,230$ 时,对协议 2 进行 10000 次模拟实验,表 2 统计了相应的出错概率.

表 2　误差率与循环次数的关系

k	100	110	120	130	140	150	160	170	180	190	200	210	220	230
ε(%)	13.4	11.3	9.1	7.3	6.0	4.8	3.9	3.6	2.7	2.4	2.0	1.6	1.3	0.9

Step2:分别设定交点数 $m=2,3,4,5,6$,重复 Step1,

① 文献[9]介绍了安全参数的取值,要求 p^t 是一个足够大的数值,确保执行 p^t 次运算是不可行的.

图 4 描述了误差率随循环次数的变化规律.

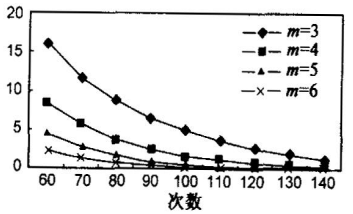


图 4 误差率随循环次数的变化规律

Step3:分别取 $n = 10, 15, 20, 25, \cdots, 50$,交点数 $m = 2, 3, 4, 5, 6$,设定重复次数 $k = 3n^2$,对协议 2 进行 10000 次模拟实验.实验结果表明,出错概率 ε 均小于 0.5%,图 5 统计了返回 true 时实际所使用的平均测试次数.

(2)实验结论

Step1 的结果(表 2)说明协议 2 是一个 0.865 正确的偏真的蒙特卡洛算法;从图 4 可以看出,交点数越多,出错的概率越低.

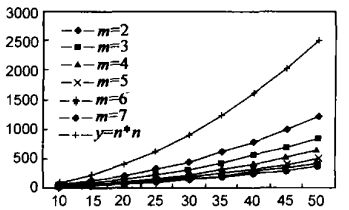


图 5 实际平均测试次数

从理论上分析,本文的概率算法需要重复 $3n^2$ 次才能使出错概率低于 0.01,但从图 5 的实验结果可以看出,当两个多边形相交时,算法实际执行的次数远远低于 n^2,协议 2 降低了计算双方的通信代价及计算成本.

6 结束语

保护私有信息的多边形相交判定是一个重要的安全多方计算问题,在军事、商业等领域都有着重要的应用前景.本文设计了一个常数时间的线段相交判定协议,基于该协议,提出了一个保护私有信息的判定多边形相交的概率算法,理论分析与实验结果均表明,该算法是高效实用的.另外,本文使用的概率算法也为解决特殊的安全多方计算问题提供了一种新的思路.

参考文献:

[1] O Goldreich. Foundations of Cryptography: Volume Ⅱ, Basic Applications [M]. Cambridge: Cambridge University Press, 2004.

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

[3] C Cachin. Efficient private bidding and auctions with an oblivious third party [A]. In Proceedings of the 6th ACM Conference on Computer and Communications Security [C]. Singapore, 1999. 120 - 127.

[4] A C Yao. How to generate and exchange secrets [A]. In Proceedings 27th IEEE Symposium on Foundations of Computer Science [C]. Toronto, Ontario, Canada, 1986. 162 - 167.

[5] O Goldreich, S Micali, A Wigderson. How to play any mental game [A]. In Proceedings of the 19th Annual ACM Symposium on Theory of Computing [C]. New York City, 1987. 218 - 229.

[6] S Goldwasser. Multi-party computations: Past and present [A]. In Proceedings of the 16th Annual ACM Symposium on Principles of Distributed Computing [C]. Santa Barbara, CA, USA, 1997. 1 - 6.

[7] 秦静,张振峰,冯登国,李宝. 无信息泄漏的比较协议[J]. 软件学报. 2004, 15(3): 421 - 427.
Qin J, Zhang ZF, Feng DG, Li B. A protocol of comparing information without leaking [J]. Journal of Software, 2004, 15(3): 421 - 427. (in Chinese)

[8] 罗永龙,黄刘生,等,空间几何对象相对位置判定中的私有信息保护[J]. 计算机研究与发展. 2006, 43(3): 410 - 416.
Luo Yong-Long, Huang Liu-Sheng, et al, Privacy protection in the relative position determination for two spatial geometric objects [J]. Chinese Journal of Computer Research and Development, 2006, 43(3): 410 - 416. (in Chinese)

[9] Luo Yong-Long, Huang Liu-Sheng, et al, Privacy-preserving distance measurement and its applications [J]. Chinese Journal of Electronics. 2006, 15(2): 237 - 241.

[10] Luo Yong-Long, Huang Liu-Sheng, et al, A secure protocol for determining whether a point is inside a convex polygon [J]. Chinese Journal of Electronics. 2006, 15(4): 578 - 582.

[11] 罗永龙,黄刘生,荆巍巍,姚亦飞,陈国良. 一个保护隐私的布尔关联规则挖掘算法[J]. 电子学报. 2005, 33(5): 900 - 903.
Luo Yonglong, Huang Liusheng, et al, An algorithm for privacy-preserving boolean association rule mining [J]. Acta Electronica Sinica. 2005, 33(5): 900 - 903. (in Chinese)

[12] Mikhail J Atallah, Wenliang Du. Secure multi-party computational geometry [A]. In Proceedings of 7th International Workshop on Algorithms and Data Structures [C]. Springer Verlag, 2001. 165 - 179.

[13] J Vaidya, C Clifton. Privacy preserving association rule mining in vertically partitioned data [A]. In Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [C]. Edmonton, Canada, 2002. 639 - 644.

[14] I Ioannidis, A Grama, M Atallah. A secure protocol for computing dot-products in clustered and distributed environments [A]. In The 2002 International Conference on Parallel Processing [C]. Los Alamitos: IEEE Computer Society Press, 2002. 379 - 384.

[15] C Clifton, M Kantarcioglu, et al. Tools for privacy preserving distributed data mining [A]. In SIGKDD Explorations [C]. New York: ACM Press, 2002, 4(2): 28 - 34.

[16] Wenliang Du, Mikhail J Atallah. Privacy-preserving cooperative statistical analysis[A]. In Proceedings of the 17th Annual Computer Security Applications Conference[C]. New Orleans, Louisiana, USA. 2001. 102 - 110.

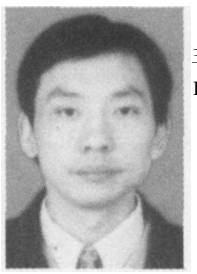
[17] Wenliang Du, Yunghsiang S Han, Shigang Chen. Privacy-preserving multivariate statistical analysis: linear regression and classification[A]. In Proceedings of the 4th SIAM International Conference on Data Mining[C]. Lake Buena Vista, Florida, 2004. 222 - 233.

[18] Wenliang Du, Zhijun Zhan. A practical approach to solve secure multi-party computation problems[A]. In New Security Paradigms Workshop[C]. Virginia Beach, Virginia, USA, 2002. 127 - 135.

[19] Thomas H Cormen, Charles E Leiserson, et al. Introduction to Algorithms. Second Edition[M]. Massachusetts: The MIT Press, 2001.

[20] Gilles Brassard, Paul Brattey. Algorithm: Theory and Practice[M]. London: Prentice Hall, 1988.

作者简介:



罗永龙　男,1972 年 4 月生,博士,副教授.主要研究方向为信息安全、分布式算法.
E-mail: ylluo@ustc.edu

黄刘生　男,1957 年 4 月生,中国科学技术大学计算机系教授、博士生导师.主要研究方向为分布式计算、信息安全、无线传感网络等.
E-mail: lshuang@ustc.edu.cn