文章编号:1002-1175(2011)02-0262-04

关于一个数值比较协议的安全性证明

邵秀凤[1],李荣花[2†]

(1 北京城市学院人工智能研究所, 北京 100083; 2 中国科学院研究生院信息安全国家重点实验室, 北京 100049)

(2009 年 12 月 21 日收稿; 2010 年 6 月 25 日收修改稿)

Shao X F, Li R H. On the security proof of a protocol for private integer comparison[J]. Journal of the Graduate School of the Chinese Academy of Sciences, 2011,28(2):262 - 265.

摘 要 Cachin 在 1990 年的 ACM 计算机和通信安全会议上提出了一个电子竞价和拍卖协议,并给出了协议的安全性证明. 我们分析发现 Cachin 的协议证明中存在一个错误,并纠正了这个错误.

关键词 信息安全, 安全协议, 可证明安全, 安全多方计算, 百万富翁问题

中图分类号 TP301

随着网络的日益发达和电子交易行业的不断发展, 保护信息的隐私性已经成为各种网络服务必须面对的问题. 比如, 在网络上有 2 个互不信任的成员, 各自有 1 个秘密的数 a、b, 能否在不泄漏 2 个秘密值的情况下比较一下是否 $a > b$? 即, 除了一个 bit 的信息外, 任何一个成员不能得到对方秘密值的其他信息. 这个有趣的问题就是 1982 年 Yao 在文献[1]中提出的“百万富翁问题”. 这个问题的提出标志着安全计算领域的诞生.

安全计算具有 2 层含义, 一是计算结果正确, 二是任何成员都不能从协议的运行中获得除了计算结果以外的关于其他成员秘密输入的信息. 如果网络中有一个可信机构, 那么安全计算问题就简单了, 各成员把自己的输入提供给这个可信机构, 由可信机构计算函数值, 并把结果分给各个成员, 这样就完成了计算, 而且保证了每个成员除了他应该得到的计算结果之外不能获得什么额外的信息. 安全计算协议就相当于实现了一个可信机构的功能.

秘密数值比较问题是安全计算问题中的一个子类, 可应用于电子交易. 比如, 比较大小协议在实际生活中最直接的应用就是电子竞价[2-3], 卖方和买方各自给出一个价格, 双方约定如果买方的价格高于卖方的价格, 那么就以平均价格成交, 否则不进行交易.

在文献[2]中, Cachin 提出了一个具有公平性的比较大小协议. 没有公平性的比较大小协议用于电子竞价或电子拍卖不能保护用户的隐私. 比如, 在一个不公平的在线竞价协议中, 卖方可能在买方还得不到任何信息的时候就知道了比较结果. 卖方一开始设个很高的价格, 在买方得知结果前, 卖方这样运行协议: 如果买方的价格高于卖方的价格则返回平均价格, 否则返回“没有交易”. 如果返回的是“没有交易”, 那么告诉买方因为网络原因此次竞价失败, 进行新一轮竞价, 并把自己的价格适当降低一点, 再如上运行协议. 这样, 卖方就能获得最大利益, 在极端的情况下, 可以得知买方的价格, 侵犯买方的隐私. 这样竞价协议就不具备安全性了.

本文研究了 Cachin 协议[2]的安全性证明, 发现其安全性证明中有错误, 并给出了正确的证明. 文献[4-5]的协议是根据 Cachin 的协议[2]构造的, 具体的协议构造和安全性证明基本与文献[2]类似, 因

† 通讯联系人,E-mail: lirhyh@ yahoo. com. cn

此为了方便理解，以 Cachin 的协议为例，首先给出具体协议步骤，然后分析其协议隐私性证明中的错误，并给出了正确的证明. 此证明也适用于文献[4-5]的协议.

1　Cachin 协议简介

1.1　Cachin 协议计算模型和假设

Cachin 在文献[2]中提出了一个公平比较大小的协议，协议包括 3 个参与者，比较大小的双方 A 和 B，以及用来保证公平性的第三方 T. 令 A 和 B 的输入空间为$[0, 2^l-1]$，第三方 T 的输入空间为$\{\varepsilon\}$. 要计算的函数为

$$f(a,b,\varepsilon)=\begin{cases}(1,0,\varepsilon), & \text{if } (a>b)\\ (0,1,\varepsilon), & \text{if } (a\leqslant b)\end{cases}.$$

A 的输入是 a，B 的输入是 b，a 和 b 都是长度为 l bit 的正整数，输出 1 个 bit 的信息；T 的输入和输出都是空. 以下假定输入长度为 l，安全参数为 k，文献[2]中还用到了其他 2 个安全参数 k'，k''. k'和 k 的关系：存在一个多项式 $p'(\cdot)$，$k'=\Theta(p'(k))$，即 k'的大小是 k 的常数倍. k''跟 k 的关系类似. l，k，k'，k''是所有成员的隐含输入.

Cachin 用到的假设：称 $m=p'q'$隐藏了一个素数 p，如果 p' 和 q'都是如下的大(比如长度为 500bit)素数：$p'=2q_1+1$（q_1 是素数），$q'=2pq_2+1$(p 和 q_2 是奇素数，p 长度较短，比如为 100bit)，显然 p 整除 $\phi(m)$. ϕ-隐藏假设如下：对于一个随机选择的数 m，m 满足隐藏一个素数 p_0 的条件，以及一个随机选择的短(比如为 100bit)素数 p_1，判断 p_0 或 p_1 是否是 $\phi(m)$的因子是困难的.

假设 a 是 $l$bit 的数，令 a_{l-1}，…，a_0 是 a 的二进制表示，即 $a=\sum_{j=0}^{l-1}a_j2^j$.

假设 m 是一个正整数，QR_m 代表 Z_m^* 的二次剩余子群.

1.2　Cachin 协议步骤

初始化阶段　　设 S 是一个语义安全的加法同态加密方案，$\boldsymbol{M}$ 是消息空间，满足$|\boldsymbol{M}|>2^{k'}$. T 产生加密方案 S 的一对密钥 (pk, sk)，把加密算法 E_{pk}向 A 和 B 公开. 令 λ 是一个映射，该映射以确定、有效(多项式时间内可计算)、可逆的方式将每一个 k'bit 串 x 与一个 k'bit 素数 p 对应起来. 令 $\lambda^{-1}(\cdot)$代表一个函数，当输入 p 后返回 $\{x'\in(\{0,1\}^{k'}|\lambda(x')=p\}$ 中的一个元素. 这样的函数存在，比如 x 被看作一个自然数，$\lambda(x)$ 是大于 x 的最小素数.

第 1 步　　A 随机地选择 x_0，x_1，…，$x_l\in\boldsymbol{M}$，s_0，s_1，…，$s_{l-1}\in\boldsymbol{M}$，和一个哈希函数 H_k 的密钥 K，A 随机选择 $t_1\in\{0,1\}^{k'}$，并计算 $p_1=\lambda(t_1)$. A 产生一个随机数 $m_1\in\boldsymbol{R}_{k''}$，$m_1$ 隐藏了 p_1. 这里 $\boldsymbol{R}_{k''}$是所有满足下面条件的数 m 的集合：$m=p'q'$，p'和 q' 是 k''bit 的素数，$p'=2q_1+1$，q_1 是素数，$q'=2pq_2+1$，p 和 q_2 是奇素数，p 的长度是 k'bit. 对于 $j=0,\cdots,l-1$，A 计算 $\delta_{1,j}=H_k(K, x_j+s_j)\oplus t_1$，A 随机选择 $u_j\in\{0,1\}^k$，并计算 $y_j=E_{pk}(u_j, x_j-x_{j+1}+a_js_j)$. A 把 x_0，x_1，…，x_l，s_0，s_1，…，s_{l-1}，K，y_0，…，y_{l-1}，$\delta_{1,0}$，…，$\delta_{1,l-1}$，m_1 发送给 B.

第 2 步　　B 随机选择 $t_2\in\{0,1\}^{k'}$，并计算 $p_2=\lambda(t_2)$. B 产生一个隐藏 p_2 的随机数 $m_2\in\boldsymbol{R}_{k''}$. 对于 $j=0,\cdots,l-1$，B 计算：$\delta_{2,j}=H_k(K, x_j+s_j)\oplus t_2$，B 随机选择 $u'_j\in\{0,1\}^k$ 并计算：$z_j=E_{pk}(u'_j, -b_js_j)\cdot y_j$. B 把 K，$x_l$，$z_0$，…，$z_{l-1}$，$\delta_{1,0}$，…，$\delta_{1,l-1}$，$\delta_{2,0}$，…，$\delta_{2,l-1}$，$m_1$，$m_2$ 发送给 T.

第 3 步　　T 令 $c_l=x_l$. T 随机选择 $g_{1,l}\in QR_{m_1}$，$g_{2,l}\in QR_{m_2}$. T 计算：$c_j=c_{j+1}+D_{sk}(z_j)$，$q_{i,j}=\lambda(H_k(K, c_j)\oplus\delta_{i,j})$，$g_{i,j}=(g_{i,j+1})^{q_{i,j}}$，这里 $j=l-1,\cdots,0$，$i=1,2$. T 随机选择 $r_i\in Z_{m_i}$，并计算 $h_i=g_{i,0}^{r_i}$，T 把 h_1 发送给 A，把 h_2 发送给 B.

获知比较结果　　A 检验是否 $h_1^{\phi(m_1)/p_1}\equiv1(\bmod m_1)$. 如果成立则 A 输出 1，A 得知 $a>b$，否则 A 输出 0；A 得知 $a\leqslant b$，原因见随后的协议正确性分析. B 检验是否 $h_2^{\phi(m_2)/p_2}\equiv1(\bmod m_2)$. 如果成立则 B 输出 1，B 得知 $a<b$，否则 B 输出 0，B 得知 $a\geqslant b$，原因见随后的协议正确性分析.

协议的正确性　对于输入 a 和 b，假设 $a\geqslant b$，令 j^* 代表 a 和 b 的二进制表示中对应比特不相等的

最高位数(比如 $a_{j^*}=1$ 且 $b_{j^*}=0$, $a_j=b_j$, 这里 $0\leqslant j<j^*$),如果 $a=b$ 则令 $j^*=-1$;对于 $j=l, l-1, \cdots, j^*+1$, 有 $c_j=x_j$, 因此 $q_{1,j}$和 $q_{2,j}$是随机的素数;对于 j^*, T 得到的是 $c_{j^*}=x_{j^*}+s_{j^*}$ 和 $q_{1,j^*}=p_1$, 但是 q_{2,j^*}是随机的素数;对于 $j=j^*-1, \cdots, 0$, $q_{1,j}$和 $q_{2,j}$是随机的素数. 可见对于 h_1 来说, h_1 有一个 p_1 次根, 因为 $q_{1,j^*}=p_1$;另一方面 $q_{2,0}, \cdots, q_{2,l-1}$都是随机的素数, r_2 是随机的, 因此 h_2 含有 p_2 次根的概率是可忽略的.

综上所述, 可知 A 检验 $h_1^{\phi(m_1)/p_1}\equiv 1 \pmod{m_1}$成立, 而 B 检验 $h_2^{\phi(m_2)/p_2}\equiv 1 \pmod{m_2}$不成立.

对于 $a<b$ 的情况, 协议正确性的证明类似.

2 Cachin 协议隐私性证明的错误和修改

2.1 证明中的错误

在证明 A 或 B 不能破坏对方的隐私性时, Cachin 使用的是归约到矛盾的方法. 以 B 为例, 假定 B 可以得到除了比较结果以外的关于 A 的输入 a 的消息, 则可以构造一个区分器 D, D 可以攻破加密方案的语义安全性, 这就与加密方案的安全性假设相矛盾, 也就证明了 B 不能获得除比较结果以外的关于 A 的秘密输入的信息. 为简单起见, 下文将 $E_{pk}(r, m)$简写为 $E(m)$.

Cachin 的证明思路如下: 以敌手 B 为例. B 在协议中收集到的消息包括: K, x_l, $x_{l-1}, \cdots, x_0$, $s_{l-1}, \cdots, s_0$, $\delta_{1,l-1}, \cdots, \delta_{1,0}$, $\delta_{2,l-1}, \cdots, \delta_{2,0}$, $y_{1,l-1}, \cdots, y_{1,0}$, $y_{2,l-1}, \cdots, y_{2,0}$, m_1, m_2, h_2. 如果 B 能获得除了比较结果以外的关于 a 的信息, 则不失一般性, 当 A 的输入是 a', a''之一时, B 应该能够区分 A 的输入, 这里 $a'\neq a''$, $f(a', b, \varepsilon)=f(a'', b, \varepsilon)$, a', a''长度均为 l bit. 根据 Cachin 的推理, 唯一可能泄漏 A 的秘密输入有关信息的是 $y_{1,l-1}, \cdots, y_{1,0}$.

令 $J\subseteq[0, l-1]$代表 a', a''不同比特位的下标组成的集合, 即 $\boldsymbol{J}=\{i \mid a_i'\neq a_i'', 0\leqslant i\leqslant l-1\}$.

D 的输入为 $a\in[0,2^{l-1}]$以及加密方案明文空间的一个消息 w. D 随机选择一个 $j^*\in\boldsymbol{J}$, 并仿真协议的运行, 仿真中 A 的输入为 a, B 的输入为 b, T 的输入为 ε. 在对 A 的仿真过程中, 如果$a_{j^*}=0$则用 $w+x_{j^*+1}$代替 x_{j^*}, 如果$a_{j^*}=1$则用 $w+x_{j^*+1}-s_{j^*}$ 代替 x_{j^*}, 其他部分的仿真与协议一致. 仿真的输出就是敌手 B 的输出. 如果 D 发起一个对协议的仿真, 参数为 a', b,ε, w', 记仿真的输出为 B′. 如果 D 发起另一个仿真, 参数为 a'', b,ε, w'', $w''\neq w'$, 记仿真的输出为 B″. 由于敌手 B 可以区分 A 的输入 a', a'', 因此 B′和 B″是可以区分的. 也即 D 可以区分 $E(w')$, $E(w'')$, 这与加密方案具有语义安全性的假设相矛盾, 因此 B 不能从协议中获得任何除比较结果以外的关于 A 的秘密输入的信息.

为了方便理解上面的仿真过程, 举例说明. 假设 $a'=12$, 其二进制表示为 $a'=(1100)_2$, $a''=11$, 其二进制表示为 $a''=(1011)_2$, $b=9$, 其二进制表示为 $b=(1001)_2$, 则 $\boldsymbol{J}=\{0, 1, 2\}$, 选择 $j^*=2$. 由于$a_2'=1$, 于是在第 1 个仿真中用 $w'+x_3-s_2$ 代替 x_2, 则得到的密文为 $y_{13}'=E(x_3-x_4+s_3)$, $y_{12}'=E(w')$, $y_{11}'=E(x_1-w'-x_3+s_2)$, $y_{10}'=E(x_0-x_1)$. 由于 $a_2''=0$, 于是在第 2 个仿真中用 $w''+x_3$ 代替 x_2, 则得到的密文为 $y_{13}''=E(x_3-x_4+s_3)$, $y_{12}''=E(w'')$, $y_{11}''=E(x_1-w''-x_3)$, $y_{10}''=E(x_0-x_1+s_0)$.

D 如何根据 B 的输出来攻破加密方案的语义安全性呢? 攻击加密方案的游戏主要是 D 选择 2 个等长的消息 m_0, m_1, 把它们发送给加密预言, 加密预言随机选择 $t\in\{0,1\}$, 对 m_t 加密得到一个密文 c 并返回给 D, D 回答 c 是对明文 m_0, m_1 中哪一个的加密. 上面的仿真证明并没有清晰地反映出 D 如何完成攻击加密方案的游戏, 即如何使得仿真中出现 $E(w_t)$, 当 $t=0$ 时 $w_t=w'$, 当 $t=1$ 时 $w_t=w''$, t 对于 D 来说是未知的随机数. 从而 D 根据 B 的输出来回答加密预言的挑战, 即回答 t', $t'\in\{0,1\}$.

更重要的是, 在 2 种情况下, 构造的密文 y_{13}', y_{12}', y_{11}', y_{10}'和 y_{13}'', y_{12}'', y_{11}'', y_{10}'', 有 3 对密文是对不同消息的加密. 假设 D 发起的仿真是对应于 a', w'的, 没有理由断定 B 是根据 y_{12}'得到了关于 a'的信息(A 的输入是 a', a''中的 a'). 若 2 组密文为 $E(m_1)$, $E(m'_2)$, $E(m_3)$, $E(m_4)$和 $E(m_1)$, $E(m_2'')$, $E(m_3)$, $E(m_4)$, 而 B 得到其中的一组, 并从中获知 A 的秘密输入的信息, 那么就可以断定 B 一定是从第 2 个密文获得关键信息的. 因此证明不具有足够的说服力.

2.2 修改

下面给出合理的符合加密方案攻击游戏的证明.

假设 B 可以获得关于 A 的秘密输入的有用信息，则不失一般性，B 应该能区分 A 的不同输入：a'，a''，满足 $f(a', b, \varepsilon) = f(a'', b, \varepsilon)$. 以 $a' > b$，$a'' > b$ 的情况为例. 不失一般性，可以这样选择 a'，a''：首先选定 a'，并令 $j^* = \min\{j \mid a'_j = 0, 0 \leqslant j \leqslant l-1\}$，然后令 $a''_j = a'_j$，$j = 0, \cdots, l-1$ 并且 $j \neq j^*$，令 $a''_{j^*} = 1$. 令 $w' = x_{j^*} - x_{j^*+1}$，$w'' = x_{j^*} - x_{j^*+1} + s_{j^*}$. D 把 w'，w''发给加密预言，得到一个密文 c. D 如下进行仿真：让 B 知道 A 的输入范围是$\{a', a''\}$，D 随机地从 a'，a''中选择一个作为仿真中 A 的输入，除了在仿真中令 A 的第 j^* 个密文为 c，其他部分的仿真与协议相同. 在协议结束时，如果 B 区分出 A 的输入了，即 B 推断 A 的输入为 a'，a''中的 a'，那么 D 就可以断定 c 是对 w'的加密；如果 B 推断 A 的输入为 a'，a''中的 a''，那么 D 就可以断定 c 是对 w''的加密，从而 D 可以攻破加密方案的语义安全性. 这与协议的假设矛盾，因此 B 不能从协议的运行中获得除比较结果以外的关于 A 的秘密输入的信息.

举例说明. $a' = (1100)_2$，$b = (1001)_2$，$j^* = 0$，则 $a'' = (1101)_2$，令 $w' = x_0 - x_1$，$w'' = x_0 - x_1 + s_0$. D 把 w'，w''发送给加密预言，假设加密预言返回的密文为 c，则在仿真中 A 发给 B 的密文为 $E(x_3 - x_4 + s_3)$，$E(x_2 - x_3 + s_2)$，$E(x_1 - x_2)$，c. 如果 B 认为 A 的秘密输入为 a'，那么 D 就回答加密预言 c 是对 w'的加密；如果 B 认为 A 的输入为 a''，那么 D 就回答加密预言 c 是对 w''的加密.

3 结束语

协议的安全性证明是协议设计中的重要环节，也是容易出错的地方，因此需要严格论证. 错误的安全性证明不但不能为安全性提供理论依据，反而成为安全性的误导. 本文所分析的 Cachin 协议虽然具有所声明的安全性，但原文给出的证明存在错误，且错误被一些文献继承(如文献[4-5]). 希望本文所给出的改正的安全性证明能为密码协议安全性证明的严格性起到借鉴作用.

参考文献

[1] Yao A C. Protocls for secure computation[C]// Proceedings of 23rd IEEE Symposium on Foundations of Computer Science. 1982:160-164.

[2] Cachin C. Efficient private bidding and auctions with an oblivious third party[C]//6th ACM Conference on Computer and Communications Security. ACM Press, 1990:120-127.

[3] Blake I F, Kolesnikov V. Strong conditional oblivious transfer and computing on intervals[C]// 10th International Conference on the Theory and Application of Cryptology and Information Security, Asiacrypt'04, Jeju Island, Korea. Berlin: Springer, LNCS 3329, 2004: 515-529.

[4] Qin J, Zhang Z F, Feng D G, 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.

[5] Qin J, Zhang Z F, Feng D G, et al. A protocol of specific secure two-party computation[J]. Journal of China Institute of Communications, 2004, 25(11): 35-42(in Chinese).
秦静，张振峰，冯登国，等. 一个特殊的安全双方计算协议[J]. 通信学报，2004，25(11)：35-42.

On the security proof of a protocol for private integer comparison

SHAO Xiu-Feng[1], LI Rong-Hua[2]

(1 *Artificial Intelligence Institute, Beijing City University, Beijing* 100083, *China*;
2 *State Key Lab of Information Security, Graduate University, Chinese Academy of Sciences, Beijing* 100049, *China*)

Abstract In 1990 ACM Conference on Computer and Communications Security, Cachin proposed a protocol for private bidding and auctions and gave the security proof. We show that there is a mistake in Cachin's security proof, and we correct the mistake.

Key words information security, secure protocol, provable security, secure multiparty computation, millionaires' problem