基于安全多方计算的可信防共谋协议模型

程柏良[1]，曾国荪[1]，揭安全[2]

(1. 同济大学 电子与息信工程学院，上海 201804；2. 江西师范大学 计算机信息工程学院，江西 南昌 330022)

摘 要：*n* 个互不信任的参与方共同计算一个函数，其中部分参与方形成联盟，通过共谋而破坏其他参与方的安全性，利用安全多方计算技术和通信通道，针对可嵌套共谋，提出可信防共谋协议模型。将模型运用到博弈论中，借助相关均衡的概念取代了可信第三方。

关键词：安全多方；可信计算；防共谋；相关均衡

中图分类号：TN918.1 **文献标识码**：A **文章编号**：1000-436X(2011)08-0023-08

Trusted coalition-proof protocol model based on secure multi-part computing

CHENG Bai-liang[1], ZENG Guo-sun[1], JIE An-quan[2]

(1. College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China;

2. College of Computer Information and Engineering, Jiangxi Normal University, Nanchang 330022, China)

Abstract: *N* parties were involved in the computation of a function, which did not trust each other. Some of them found a coalition and destroy the security of others through collusion. Based on secure multi-party computation and communications channel, to the collusion that could be nested, a trusted coalition-proof protocol model was proposed. The concept of correlated equilibrium was used by the model in game theory for talking place the trusted third party.

Key words: secure multi-part; trust computing; coalition-proof; correlated equilibrium

1 引言

互不信任的多个参与方如何协同合作完成一项任务是可信计算的研究方向之一，涉及信任的建立，隐私数据的保护，防止部分参与方私下形成共谋联盟以及确保各方获得正确结果等诸多问题。现有的技术从 2 个方面研究该问题，一是假定有可信方，通过可信方来协调多个参与方之间的合作。其优点是简单有效，不足是可信方的效率和安全性成为瓶颈，一旦可信方负载过大，或者受到攻击，整个系统将面临巨大的风险。另一种方法是通过安全多方技术设计有效的安全协议以替代可信方。姚期智[1]首先研究了安全两方计算，随后 Goldreich[2]从理论上建立安全多方计算的基本模型。

在安全多方计算中，*n* 个参与方提供各自的隐私数据给计算任务，计算完成之后，*n* 个参与方获得计算结果。设计一个安全多方计算协议最基本的要求是保证计算的安全性(包括各方的输入隐私性

收稿日期：2010-04-16；**修回日期**：2010-11-02

基金项目：国家高技术研究发展计划（“863”计划）基金资助项目（2007AA01Z425, 2009AA012201）；国家重点基础研究发展计划（“973”计划）基金资助项目（2007CB316502）；国家自然科学基金资助项目（90718015）；NSFC-微软亚洲研究院联合资助项目(60970155)

Foundation Items: The National High Technology Research and Development Program of China (863 Program) (2007AA01Z425, 2009AA012201); The National Basic Research Program of China (973 Program)(2007CB316502); The National Natural Science Foundation of China (90718015); The Joint of NSFC and Microsoft Asia Research (60970155)

和输出正确性)。n 方中的部分参与方可能形成联盟，通过共谋而破坏其他参与方的安全性，联盟的共谋用一个攻击者来表示，它协调该联盟中成员的活动，攻击者对一个参与方的攻击通过截获、修改该参与方所发送和接收的消息进行。针对攻击者的能力，通常把参与方分成3种类型：①诚实参与方；②半诚实参与方；③恶意参与方。在协议的执行过程中，诚实参与方完全按照协议的要求完成协议的各个步骤，同时保密自己的所有输入、输出及中间结果，即攻击者对它们的攻击是无效的。半诚实参与方完全按照协议的要求完成协议的各个步骤，同时可能将自己的所有输入、输出及中间结果泄露给攻击者。即攻击者可以窃取它们的信息，称之为被动攻击。恶意参与方完全按照攻击者的指令执行协议的各个步骤，它不但将自己的所有输入、输出及中间结果泄露给攻击者，还可以根据攻击者的意图改变输入信息、中间结果等信息，甚至终止协议。即攻击者可以修改它们的信息，称之为主动攻击。在安全多方计算中，除了基本的安全性要求以外，还扩展到公平性、移动攻击、动态攻击等更复杂的要求。所谓公平性是指：如果任何一方获得它的计算结果，所有其他参与方也同时获得它们各自的计算结果。文献[3]指出，公平性需要诚实方占多数，否则除非拥有物理通信通道，安全多方协议才可能获得公平性。

当前的防共谋研究中，共同点均指出在一般安全多方计算中，攻击者可以完全控制所有的不诚实(恶意)方这一问题，即所有共谋成员彼此完全信任对方。为了防止共谋，文献[4]将每个不诚实(恶意)方视为独立的参与者，不能与其他不诚实方进行秘密通信，同时对广播的消息进行可验证确定性来防止不诚实方之间的隐秘通信，从而获得所有不诚实方都是彼此独立、无法共谋的安全协议。文献[5]则从另一个方面来处理防共谋，通过在所有参与方之间设置一个消息转发参与者，以形如路由器的方式，作为所有参与者之间传递消息的转发者，并且在转发消息时对消息进行过滤和重随机处理以屏蔽不诚实方之间的隐秘通信。而所有参与者在消息发送时对消息进行签名，以防止消息转发参与者对消息进行修改。文献[6]进一步完善了文献[5]的思想，并指出了其中的缺陷（文献[5]中的理想模型中的不诚实方，不仅依赖于现实模型中的不诚实方，而且还依赖于后者的执行策略），并通过建立执行策略的模拟器来解决该缺陷。

本文并不将防共谋设定为每个不诚实(恶意)方视为彼此独立，互不协调通信的参与者，而是将防共谋设定在不诚实 (恶意) 方并不完全合作的情形下，在最坏的情况下，所有参与方彼此独立，完全不合作，在最好情况下，所有共谋者完全彼此信任。而一般情况下，可以获得部分参与者形成共谋的局面。因此将研究多个攻击者对协议进行攻击。即存在一个攻击者集合，并限定集合中攻击者按照攻击次序进行嵌套攻击。比如：攻击者 a 在攻击了它所控制的参与方集合 s 后，下一个攻击者 b 可以继续攻击 s 中的子集 r，而 a 并不知道 b 在其之后继续对 r 进行攻击。这样联盟在共谋之后，联盟中的部分成员组成新的联盟，而该新联盟对原联盟进行背离，开始新一轮共谋，并且这种分裂背离可以嵌套进行，形成多轮共谋，这样每一个分裂出来的新联盟对应一个攻击者。

同时本文将文献[7]中提出的完全公平性纳入上述嵌套攻击中。完全公平性在公平性的基础上，要求攻击者可以攻击任意数量的参与方，即联盟可以是 n 个参与方的任意子集。依据文献[8]的结论，在只有一个攻击者的安全多方协议中，对于被动攻击，安全协议可以满足所攻击的联盟是 n 个参与方的任意子集，而对于主动攻击，可以满足所攻击的联盟，其规模不超过 $n/2$。本文论述对于攻击者集合中任意一个攻击者而言，它所攻击的联盟可以是 n 个参与方的任意非空子集，借助通信通道，即使在恶意攻击下，上述结论也成立。

需要指出的是，文中所研究的范围限定在非自适应攻击，即攻击者在协议执行开始前就明确了它所要攻击的参与方（在非自适应攻击中，被动攻击下，参与方之间存在秘密信道，在主动攻击下，存在广播信道。所有参与方都是有限计算能力的，即交互式图灵机）。对于自适应攻击而言[9]，即它所攻击的参与方在协议执行过程确定，这将在后续研究给予论述。

2 可信防共谋协议模型

文献[8]给出了一般安全多方协议的基本模型，其方法是通过现实协议环境模型、模拟理想环境模型，将函数的计算转换为电路计算的过程。在电路计算中，通过在电路的每个计算门上，对门的输入和输出数据在所有参与方中共享，每个参与方持有

的共享份额不能产生相应值的任何信息。但当组合所有参与方共享份额时，就可以完全恢复相应值。每个基本门的操作不产生任何附加信息。这样电路上所有的共享份额是在秘密的方式下进行。主要的实现技术是零知识证明[10]，不经意传输[11]，可验证秘密共享[12]等。在通过电路转换获得半诚实协议模型后，在利用多方承诺，多方掷币函数，多方认证计算等技术构成一个编译器，将半诚实协议编译成恶意攻击下的协议。文献[8]给出了详细的论述过程。

本文在其基础上，定义防共谋协议模型，首先，在有可信方的理想环境下定义防共谋模型，之后在没有可信方的现实环境下，定义可信防共谋协议模型，通过密码学中计算不可分辨性的概念来表示现实协议模型对理想模型的模拟程度。希望所建立的可信协议模型可以抵抗 $n-1$ 轮以内的共谋。为方便论述，下面给出相关形式化描述：$N=(1,\cdots,n)$ 是 n 个参与方的集合，$f:(\{0.1\}^*)^n\to(\{0.1\}^*)^n$ 是 n 元可计算函数，N 的输入集合为 $X=(x_1,\cdots,x_n)$，x_i 是 $\forall i\in N$ 的输入，$(f_1(x_1,\cdots,x_n),\cdots,f_n(x_1,\cdots,x_n))$ 是 $f(X)$ 计算后的输出集合，简记为 $Y=(y_1,\cdots,y_n)$，$f_i(x_1,\cdots,x_n)$ 是 i 的输出，简记为 y_i。$I=(I_1,\cdots,I_g)\,|\,g\leqslant 2^n$ 是不诚实(恶意)参与方联盟集合，满足 $\forall i$、$j\in(1,\cdots,g)$, $i<j, I_i\supseteq I_j$，$I_i=(i_1,\cdots,i_t)\subset N$ 是 $t\,|\,t<n$ 个参与方形成的联盟, $X_{I_i}=(x_{i_1},\cdots,x_{i_t})$ 是 I_i 的输入，$Y_{I_i}=(y_{i_1},\cdots,y_{i_t})$ 是 I_i 的输出。

2.1　理想环境模型

在理想环境下，假定存在一个 n 方都完全信任的可信方，该可信方完全保持中立，对于任意一个有效的可计算的函数，每个参与方将自己的隐私输入发送给可信方，由可信方启动函数计算，并将输出计算结果给 n 个参与方。

定义 1（理想防共谋模型）　给定 N、f、I 以及可信方，N 发送输入 X 给可信方，在可信方完成 f 的计算后，获得输出 Y，一个攻击者集合 $S=\{s_1,\cdots,s_g\}$，$g\leqslant 2^n$ 中的任意一个攻击者都是一个交互图灵机(概率多项式时间)，$\forall i$、$j\in[1,\cdots,g]$，$i<j$，$s_i>s_j$，表示 s_i 早于 s_j 开始攻击。即 s_j 在 s_i 攻击 I_i 后，攻击 I_i 的子集 I_j，每个攻击者 s_i 对应一个 I_i。s_i 攻击 I_i 的行为模式以控制 I_i 成员与可信方的通信来体现，首先 s_i 初始化自己的随机带 R_{s_i}，获得辅助输入 z_i，安全参数 k_i，当 $\forall p\in I_i$ 发送 x_p 给可信方时，s_i 截获 x_p，修改 $x_p=\{0.1\}^*\bigcup\perp\,p\in I_i$，其中 $\perp$ 表示中断可信方操作。可信方在收到所有输入后，如果输入有 $\perp$，则发送 $\perp$ 给所有参与方以表示可信方中断执行。否则开始计算 f，当可信方发送 y_p 给 p 时，s_i 截获并修改 $y_p=\{0.1\}^*$ 后发送给 p。在攻击 I_i 中所有参与方后，s_i 获得自己的输出 ω_i。N、S 和可信方对 f 的联合执行记为 $IDEAL_{f,I,S(z)}(X)$，记为以下 3 个序列：①n 方的最终输入 X；②n 方的最终输出 Y；③所有攻击者的输出 $(\omega_1,\cdots,\omega_g)$。其中，$z=(z_1,\cdots,z_g)$ 是攻击者集合的辅助输入。

2.2　现实环境协议模型

在没有可信方的现实模型下，所有参与方通过一个协议来计算 f，一个 n 方的协议 $\Pi=(P_1,\cdots,P_n)$ 是 n 个交互图灵机序列(概率多项式时间)，$\forall P_{i\in(1,\cdots,n)}\in\Pi$ 代表第 i 个可交互式图灵机。协议 Π 是有效的，必须满足 Π 中每一个图灵机是有效的。在 Π 的执行过程中，有一个攻击者集合，集合中每一个攻击者控制了理想模型攻击下相同的联盟。协议执行后每一个攻击者获得自己的观点(攻击者所截获的所有消息)。在有效的输入、输出函数下，攻击者的观点被映射到该攻击者所控制联盟的输入和输出向量上。下面给出形式化的定义。

定义 2（现实协议模型）　给定 N、f、I、Π、N 中成员发送输入 X 给 Π，并且在 Π 完成 f 的计算后，获得输出 Y，$A=\{a_1,\cdots,a_g\}$ 是攻击 Π 的攻击者集合，$\forall a_{i\in(1,\cdots,g)}\in A$ 是一个交互图灵机(概率多项式时间)，且 $\forall i$、$j\in[1,\cdots,g]$, $i<j$，则 $a_i>a_j$，表示 a_i 早于 a_j 开始攻击。即 a_j 在 a_i 攻击联盟 I_i 后，攻击 I_i 的联盟子集 I_j，a_i 攻击 Π 的行为模式以控制 I_i 中成员在执行 Π 的过程中所发送和接收的消息来体现。

a_i 初始化自己的随机带 R_{a_i}，在辅助输入 r_i，安全参数 k_i，截获 I_i 的输入下开始攻击 Π。$\forall p\in I_i$ 以安全参数 k_p，辅助输入 r'_p，当前输入 x_p 作为参数运行 P_p。在 Π 的执行过程中，$\forall j\in\{1,\cdots,n\}$,假定该 $j\in I_i\wedge p\notin I_{i+1}$，则 j 所发送(接收)的消息，由 $\{a_1,\cdots,a_{i\leqslant g}\}$ 按照攻击次序进行截获和修改。即先由 a_1 截获、修改，而后由 a_2 截获、修改，直到 a_i 截获，修改后再发送(接收)。

对于 Π 的某此执行 e,以 $view_{a_i}(e)$ 表示攻击者 a_i 在对 Π 此次执行过程中进行攻击所获得的观点(a_i 的初始随机带，辅助输入，截获接收的所有消息)。如果 Π 是个有效的安全多方协议，则存在输入、输出函数 AI 、 AO ，其中 $AI(view_{a_i}(e))=X_{I_i}$ ，$X_{I_i}=\{x_{i_t}=(0.1)^*\bigcup\perp\}\,|\,i_t\in I_i$ ， $AO(view_{a_i}(e))=Y_{I_i}$ ，$Y_{I_i}=\{y_{i_t}=(0.1)^*\bigcup\perp\}\,|\,i_t\in I_i$ ， X_{I_i} 、Y_{I_i} 是由 a_i 攻击 I_i 时所设置的输入和输出向量。

在现实协议模型下，N、A、Π 对 f 的联合执行记为 $REAL_{f,\Pi,A(r)}(X)$，由以下 3 个序列构成：①n 方的最终输入 X ；②n 方的最终输出 Y ；③所有攻击者的观点 $view_{a_1}(e),\cdots,view_{a_g}(e)$ 。其中 $r=(r_1,\cdots,r_g)$ 是攻击者集合辅助输入。

2.3 可信防共谋协议

在定义理想模型和现实协议模型后，本文利用密码学中的计算不可分辨性的概念[13]来度量 2 种模型的近似程度。如果 2 种模型是计算不可分辨的，就可以认为 Π 是安全计算 f 。即不管攻击者在现实模型中可以获得什么，具有相同攻击能力的另一攻击者在理想模型中也可得到，使得 Π 能够模拟可信方，并且保持安全性。这表明在理想环境中满足的性质可以通过执行 Π，使得这些性质在现实环境中也满足。

定义 3 2 个由字符串标记的概率总体 $X\overset{\text{def}}{=}\{X_w\}_{w\in s}$，$Y\overset{\text{def}}{=}\{Y_w\}_{w\in s}$ 是计算不可分辨的 $X\overset{c}{\equiv}Y$，则满足如果对每个多项式长的电路族 $\{C_n\}_{n\in N}$ ，每个正多项式 $P(*)$ ，每个足够大的 n，每个 $w\in S\bigcap(0.1)^*$ ，有 $|\Pr[C_n(X_w)=1]-\Pr[C_n(Y_w)=1]|<\dfrac{1}{P(n)}$ 成立。

在此定义基础上，把计算不可分辨性的概念扩展到攻击者集合上，通过集合空间不可分辨性，以建立定义 2 和定义 3 的计算不分辨性。为了获得完全公平性，期望现实模型中的攻击者可以控制任何数量的参与方。为此给出如下定义。

定义 4（可信防共谋协议） 给定任意一个有效，可计算函数 f ，对于每一个攻击者集合 A，存在一个模拟攻击者集合 S ，使得对于 I ，满足 $REAL_{f,\Pi,A(r)}(X)\overset{c}{\equiv}IDEAL_{f,I,S(z)}(X)$ 。则 Π 安全计算函数 f ，称 Π 是可信防共谋协议，即协议是完全公平防共谋安全多方的。

为了方便后文博弈论中应用，在理想模型中，可信方收到一个参与方 p 的中断请求 $x_p=\perp$，它发送 $(p,\perp)$ 给所有的参与方。

3 可信防共谋协议模型的构建

本节通过文献[2]的基本协议和通信通道来构建符合第 2 节中定义的协议模型。首先考虑将定义 1 和定义 2 中的模型退化成只有一个攻击者的情况，在这种情况下，上述定义退化成一般意义下的安全多方计算协议。为获得一般意义下安全多方计算协议的完全公平性，文献[7]通过信封这一物理通信通道来获得完全公平性：其基本思路在文献[2]的基础上，在电路计算协议输出阶段的共享信息交换时，将所有交换的信息放入信封中，利用信封的保密通信来确保数据的隐私性，并通过非交互式零知识证明技术[14]确保接收方在打开信封之前，发送方向接收方证明所交换的共享数据是文献[2]中所发送的真实数据，从而获得数据的正确性，以防止发送方将虚假信息发送给接收方。在具体的信封设计上，文献[7]以公钥加密技术来表示信封，将物理信封以电子信封来表示。接下来将证明，对只有一个攻击者的完全公平安全多方协议模型[2]进行扩展，可以获得满足上文中所定义的可信防共谋协议模型。

定理 1 假定 Π 是在攻击者集合只有一个攻击者的情况下的安全多方计算协议(记为：防 1 轮共谋的安全多方计算协议)， Π' 是在攻击者集合由 2 个攻击者组成情况下的计算协议，其第一攻击者是 Π 中的攻击者，如果 Π' 满足定义 2，则 Π' 是一个具有防 2 轮共谋的安全多方计算协议。

证明 标记攻击者集合为 $A_1=(a_1)$ ， $A_2=(a_1,a_2)$ 分别对应协议 Π 和 Π'，a_1,a_2 控制的参与方联盟记为 I_1,I_2，由 Π 是一般安全多方计算协议，有 $I_1\subset N$，由定义 2 有 $I_1\supseteq I_2,a_1>a_2$ 。在 Π' 的运行中，依据定义 2 中攻击次序， a_1 先攻击 Π'，控制 I_1 中的成员，$\forall p\in I_1$ 所发送给 $\forall q\in N$ 的消息都由 a_1 先截获，在修改后由 a_1 发送给 q 。 p 所接收来自 q 的消息都由 a_1 先截获，在修改后由 a_1 发送给该 p 。

1) 如果 $p\notin I_2$,则 a_2 不能攻击 p ，所以 p 发送和接收的消息不受 a_2 的影响，仅有 a_1 决定。

2) 如果 $p\in I_2$，则 p 发送的消息先由 a_1 截获，修改后发送，并再次由 a_2 截获，修改后再发送。同

时， p 接收的消息先由 a_1 截获，修改后向 p 发送，并再次由 a_2 截获，修改向 p 发送。对 a_1 来说，其先于 a_2 对 p 进行攻击，a_1 并不知道 a_2 修改了它对 p 的攻击行为，因此， a_1 在攻击 Π' 的执行过程中所获得的观点同 a_1 在攻击 Π 的执行过程中所获得的观点是一致的。

由上述 2 点可得：a_1 对于 p 的观点在 Π' 与 Π 中是保持不变的。而对 a_2 而言，对 $p\in I_2$，它对 p 的观点由截获 a_1 对 p 的攻击而获得。直观的理解可以认为当 $p\notin I_2$ 时，相对于 a_1 来说， a_2 可以把它看成一个诚实参与方(实际上 p 不是诚实参与方，这里仅仅是从 a_2 的视角来看，因为 a_2 看不见 a_1 对 p 的攻击)，而 $p\in I_2$ 时，相对于 a_2 来说， a_1 以为它完全控制了 p（实际上 p 不是完全由 a_1 所控制，这里仅仅是从 a_1 的视角来看，因为它不知道在其之后， a_2 又对 p 进行了攻击）。证毕。

定理 2　假定 Π 是在攻击者集合只有一个攻击者情况下的安全多方计算协议，并且该攻击者控制了 $t\,|\,t\leqslant n-1$ 个参与方（n 个参与方最多只有 n–1 个参与方共谋，n 个参与方同时共谋是无意义的[5]，所以攻击者最多控制 n–1 个参与方）（记为：防 1 轮共谋的安全多方计算协议），则存在满足防 $k\,|\,k\leqslant t$ 轮共谋的安全多方计算协议 Π'，只要 Π' 的攻击者集合中的第一个攻击者是 Π 中的攻击者，且攻击者集合的攻击行为满足定义 2。

证明　下面采用归纳法证明，由定理 1 可知，当攻击者集合 A 中仅有 2 个成员，并且攻击者集合的攻击行为满足定义 2 时，则存在防 2 轮共谋的安全多方计算协议 Π'。即 $|A|=2$ 时，定理 2 成立。假定 $|A|=k, k\leqslant t-1$ 时，定理 2 也成立，即存在防 k 轮共谋协议 $\Pi'(k)$，那么对于 $|A|=k+1$ 时，可以将 A 中的成员看成 2 部分，一个前 k 个攻击者 $\{a_1,\cdots,a_k\}$ 集合，另一个是攻击者 a_{k+1}，由定义 2 可知，对于协议 $\Pi'(k+1)$ 而言， $\{a_1,\cdots,a_k\}$ 的攻击行为与 $\Pi'(k)$ 中保持一致，即攻击 $\{I_1,\cdots,I_k\}$，而对于 a_{k+1} 而言，将在 $\{a_1,\cdots,a_k\}$ 攻击之后，对 I_k 的子集 I_{k+1} 进行攻击。这样 $\forall p\in I_j, p\notin I_{j+1}\,|\,j\leqslant k$ 所发送给 $\forall q\in N$ 的消息都由 a_j 最终截获，修改后发送给 q。p 所接收的消息都由 a_j 最终截获，修改后发送给该 p。

1) 当 $j<k$，则有 $p\notin I_{j+1}$，进而 $p\notin I_{k+1}$，则 a_{k+1} 不能攻击 p，所以 p 发送和接收的消息不受 a_{k+1} 的影响，仅由 a_j 决定。

2) 当 $j=k$，有 $p\in I_k$，则 p 发送的消息最终由 a_k 截获，修改后发送。同时， p 接收的消息最终由 a_k 截获，修改后发送给 p。如果 $p\notin I_{k+1}$，情形如 1)所述，如果 $p\in I_{k+1}$，对 a_k 来说，其先于 a_{k+1} 对 p 进行攻击， a_k 并不知道 a_{k+1} 修改了它对 p 的攻击行为，因此 a_k 在攻击 $\Pi'(k+1)$ 的执行过程中所获得的观点同 a_k 在攻击 $\Pi'(k)$ 的过程中所获得的观点是一致的。

从上可知，$\{a_1,\cdots,a_k\}$ 对于 p 的观点在 $\Pi'(k+1)$ 与 $\Pi'(k)$ 中是保持不变的。而对 a_{k+1} 而言，但 $p\in I_{k+1}$，它对 p 的观点由截获 a_k 对 p 的攻击而获得。直观的理解可以认为当 $p\notin I_{k+1}$ 时，相对于 $\{a_1,\cdots,a_k\}$ 来说，a_{k+1} 可以把它看成一个诚实参与方(实际上 p 不是诚实参与方，这里仅仅是从 a_{k+1} 的视角来看，因为 a_{k+1} 看不见 $\{a_1,\cdots,a_k\}$ 对 p 的攻击)，而 $p\in I_{k+1}$ 时，相对于 a_{k+1} 来说，$\{a_1,\cdots,a_k\}$ 以为它完全控制了 p (实际上 p 不是完全由 $\{a_1,\cdots,a_k\}$ 所控制，这里仅仅是从 $\{a_1,\cdots,a_k\}$ 的视角来看，因为它不知道在其之后， a_{k+1} 又对 p 进行了攻击)。

最极端的情况下，第一个攻击者控制了 $t=n-1$ 个参与方，而后每个攻击者依次比上一个攻击者少控制一个参与方，这样攻击者总数不多于 $n-1$，则可以获得防 $n-1$ 轮共谋的安全多方协议，而 $k\leqslant t$，所以存在防 $k\leqslant n-1$ 轮共谋的安全多方计算协议 Π'。证毕。

为获得一般意义下安全多方计算协议的完全公平性，文献[7]通过信封这一通信通道来获得完全公平性：但是该方法假定任意一个信封在打开之前，所有的信封都已发送到相应的接收方。本文利用可验证秘密共享机制[12]来解决这一问题。

一个 $(t,n)\,|\,t\leqslant n$ 可验证秘密共享机制是秘密共享函数，直观的理解是包括秘密分发和秘密恢复 2 个算法，一个秘密通过秘密分发算法被分成 n 份，在 n 个参与方之间共享，即每个参与方获得一份共享数据值。使得其中任意 t 或 t 个以上参与方的共享份额合在一起，通过秘密恢复算法可以恢复这个秘密，而对任意少于 t 数量的参与方而言，秘密恢复算法获得它们的共享份额都不能恢复这个秘密。文献[8]给出了一般可验证秘密共享的方法。

这里利用 (n,n) 可验证秘密共享技术，可以对文献[7]进行改进，以获得任意一个信封在打开之

前，所有的信封都已发送到相应的接收方。即所有参与方在电路计算的输出阶段，每个参与方发送给其他参与方的信封不是直接发送给接收方，而是发送到一个黑盒中，并用非交互式零知识证明技术向接收方证明信封中所发送数据的正确性后，该黑盒的开锁密钥被分成 n 个子密钥，只有 n 个子密钥合在一起，才可以打开黑盒。对每一个参与方，直到它发送了 n–1 个信封给黑盒后（以确保其他 n–1 方获得共享信息交换中的数据），黑盒才分发一个子密钥给该参与方。因此任何一个参与方要从黑盒中取出信封，必须要 n 个参与方的子密钥合在一起才能共同打开黑盒，任何少于数量 n 的参与方集合都不能打开黑盒。这样任意一个信封在从黑盒中被取出处理之前，所有的信封都已发送到黑盒中。所有参与方同时从黑盒中取出其他参与方发送给自己的信封，从而克服文献[7]的不足，从而获得完全公平性。

由定理 1 和定理 2，以及通信通道和可验证秘密共享技术，可以获得防 n–1 轮共谋的完全公平性安全多方协议。

4 防共谋协议模型在博弈论中的应用

这一节将论述上述所定义的协议，应用到非合作博弈中以获得防共谋相关均衡[15]。一个博弈活动 $\Gamma=(N,(G_i)_{i\in I},(u_i)_{i\in I})$，符号涵义如下。

参与方集合：$N=(1,\cdots,n)$，N 是所有参与方构成的集合，$\forall i\in N$ 是第 i 个参与方。

纯策略集合：$G_i,\cdots,G_n$，$G=\prod_{i\in N}G_i$ 是所有参与方的策略组合所构成的集合，G_i 表示参与方 i 的策略集合(有限集)。

收益函数：$u_i,\cdots,u_n,u_i:G\to R$ 是参与方 i 的收益。

N 中的一个子集 $s\in 2^N\wedge s\neq\Phi$ 是一个联盟，s 的补集记为 $-s$，以 G_s 表示集合 $\prod_{i\in s}G_i$，给定一个策略组合 $a\in G$，记 $a=(a_s,a_{-s})$，其中 $a_s\in G_s$，$a_{-s}\in G_{-s}$。如果 $s=N$，则 $(a_s,a_{-s})=a_s=a$。

4.1 防共谋相关均衡

在博弈活动中，每个参与方都希望使自己的收益最大化，假定在博弈活动开始之前，n 个参与方达成一致，同意以 $a\in G$ 开始活动，而部分参与方结成联盟 s，背离 a，通过共谋而希望获得更高的收益。即该联盟在其他参与方服从 a_{-s} 的情况下整体背离 a。以 $\eta_s:G_s\to G_s$ 表示共谋函数。联盟 s 共谋后，形成一个新的策略组合 $a'=(a_s',a_{-s})$。

定义 5（可行共谋） 设联盟 $s\in 2^N,s\neq\Phi$，$a=(a_s,a_{-s})\in G$，可以说 a' 是一个背离 a 的可行共谋，如果 N/s 的参与方在保持 a_{-s} 的条件下，存在一个函数 η_s 将 a_s 映射到 a_s'，有 $a'=(a_s',a_{-s})$ 成立。

如上文所述，可以把共谋看成一个攻击者对联盟 s 的攻击，在可信方按照策略组合 a 发送 a 中相应的分量策略信息给每个参与方作为推荐活动，攻击者截获发送给 s 中所有成员的消息得到 a_s，而后以 η_s 重新确定 s 的推荐信息 $\eta_s(a_s)=a_s'$，并将 a_s' 中相应的分量策略信息发送给 s 中每个成员，从而形成新的策略组合 a'，这样所有成员以新的组合策略 a' 开始活动。以 $D(a,s)$ 记为联盟 s 背离 a 的可行共谋集合。特别地，对任意 $a\in G$，有 $D(a,s)=G$。而当 s 中只有一个参与方时，即 $|s|=1$，若满足单个参与方没有可行的背离共谋，则可获得相关均衡。文献[16]给出其形式化定义。

定义 6（相关均衡） 一个组合策略 a 是一个相关均衡，是指没有任何一个参与方 $i\in N$ 有可行共谋 $a'\in D(a,i)$，使得 $u_i(a')>u_i(a)$ 成立。

正如前文所述，一个联盟不断分裂形成多轮共谋时，利用博弈论中的防共谋相关均衡的概念描述从第一次共谋到最后一轮共谋的背离过程。首先给出关于联盟 s 的自我强化背离的概念[15]。

定义 7 （自我强化背离） 设 $a\in G$，联盟 $s\in 2^N$ 且 $s\neq\Phi$，联盟 s 背离 a 的自我强化背离记为 $SED(a,s)$，递归定义如下：

1) 如果 $|s|=1$，$SED(a,s)=D(a,s)$。

2) 如果 $|s|>1$，$SED(a,s)=\{a'=D(a,s)\}$，使得在 a' 下，没有任何子集 $R\in 2^s\wedge R\neq\Phi$，有 $a''\in SED(a',R)$，使得 $\forall i\in R$，有 $u_i(a'')>u_i(a')$ 成立。

直观上可以理解为，在给定组合策略 a，联盟 s 的条件下，联盟 s 背离 a 的自我强化，背离是指 s 获得一个共谋 a'，s 中任意子集都无法进一步形成新的共谋而获得比 a' 中有更高的收益。在自我强化背离的基础上，下面给出防共谋相关均衡的概念。

定义 8 （防共谋相关均衡） 一个组合策略 a 是一个防共谋相关均衡，如果没有任何联盟 $s\in 2^N,s\neq\Phi$，有 $a'=SED(a,s)$，使得 $u_i(a')>u_i(a)|\forall i\in s$ 成立。

直观上可以理解为在给定组合策略 a 下，N 中的任意子集都无法进一步形成新的共谋而获得更高收益。即组合策略 a 是任何一个联盟都不会在背离的，因此没有联盟会共谋。

4.2 防共谋安全

文献[7]指出，即使是求解相关均衡也需要有可信方的参与，采用廉价磋商通信机制来取代可信方，但不足是无法防止联盟的共谋，而文献[7]用安全多方协议实现廉价磋商获得防一轮共谋(联盟内部不会再形成共谋)，即要求联盟内部所有成员是完全互信的，不会出现多次背离的多轮共谋，这个条件显然过于严格，为此，利用上文中所定义的可信防共谋协议来实现廉价磋商。

定义 9（防共谋廉价磋商安全协议） 设 a 是博弈 Γ 的防共谋相关均衡，T 和 Π 是实现 a 的可信方和廉价磋商协议，若 $B=(b_1,\cdots,b_g)\,|\,b_1\supseteq,\cdots,\supseteq b_g$ 是任意联盟集合，攻击者集合 A，在协议 Π 中对 B 的任意攻击后所获得收益向量，记为 $U_{A,\Pi,a}=(u_{i\in N})$，都存在一个攻击者集合 S，在可信方的计算下对 B 的任意攻击后所获得收益向量，记为 $U_{S,T,a}=(u_{i\in N})$，使得 $U_{A,\Pi,a}$ 同 $U_{S,T,a}=(u_{i\in N})$ 是计算不可分辨的，则 Π 是防共谋廉价磋商安全协议。

4.3 防共谋安全协议的实现

命题 1 给定用于求解任意有效，可计算函数 f 的完全公平，可中断的可信防共谋协议，则对任意具有防共谋相关均衡 a 的博弈 Γ，存在一个防共谋廉价磋商协议 Π，使得在协议 Π 的执行下，参与方达成一个策略组合，在博弈 Γ 执行这一策略所获得的收益集合与执行 a 后获得的收益集合是不可分辨的。

证明 以如下方式建立一个防共谋廉价磋商协议安全协议 Π：首先把计算 Γ 中的防共谋相关均衡看成一个函数 f，其次所有的参与方共同执行一个协议。该协议是求解 f 的完全公平，可中断识别的可信防共谋协议。在该协议中任意一个攻击者集合 A，控制了联盟集合 B，a_i 控制 b_i。首先在没有攻击者发出中断攻击的情形下，任意一个攻击者 a_i 截获发送给 b_i 中成员的消息，在修改之后重新发送给 b_i 中成员。因为 Π 是可信防共谋协议，所以不论 A 如何攻击，所有参与方所达成的组合策略是同 a 不可分辨的。在这一前提下，A 的任意攻击抽象为一个攻击算法 $\wp$，执行 $\wp$ 后获得最终策略组合 a'。可以在攻击可信方下得到。在可信方下，攻击者集合 S，控制了博弈活动中同样的参与方 B，满足 s_i 控制 b_i，在可信方按照组合策略 a 发送相应的策略分量给所有参与方作为推荐策略时，任意一个攻击者 s_i 截获发送给 b_i 中成员的推荐策略，而后 S 调用算法 $\wp$。获得策略组合 a'。所以 2 种情况下的收益向量不可分辨的。在有中断的情况下，因为 Π 是中断可识别的，文献[17]在相关均衡中通过惩罚中断成员而阻止攻击者对某个成员发送中断指令，在本文中，当攻击者发出中断指令时，由于攻击者对应的是一个联盟，所以在惩罚中断成员不能防止其他参与方与该中断成员在惩罚后继续共谋，所以文献[17]的方法只能适用于单个参与方的背离形成的共谋。对于联盟共谋，文献[7]在防一轮共谋(基于安全多方的廉价磋商协议)中提供方法是每次对攻击者所确定的联盟成员进行排除法，即每次删除中断参与方，重新执行 n–1 个参与方的博弈。并以同样的协议来执行的 n–1 方的博弈以获得防一轮共谋相关均衡，如果在此过程中，再次出现中断，则重新执行 n–2 个参与方的博弈。并以此类推，直到所有可能中断的参与方都被排除，于是剩余的参与方获得防一轮共谋的相关均衡。为了获得防共谋相关均衡，本文在排除法的基础上进行改进，一旦协议中一个参与方发出中断信息，则所有可以控制该参与方的攻击者构成一个集合 $A'=(a_1,\cdots,a_k)$，$k\leqslant g$，在 A' 中，首先从 a_k 所控制的联盟 b_k 开始，进行文献[4]中的操作。如果在排除 b_k 所有成员后，中断继续出现，则从 a_{k-1} 所控制的联盟 b_{k-1} 中的 $b_{k-1}-b_k$ 继续文献[4]的操作。直到没有中断出现，则所有剩余参与方获得的防共谋相关均衡。证毕

需要指出的是，为了方便论述在博弈论中的应用，这一节的工作都是基于纯策略的，对于将防共谋安全多方协议应用到博弈论中的概率策略将在以后的研究中论述，文献[15,18]给出了概率策略下的防共谋相关均衡的定义。将上述论述推广到概率策略下可以得到更为复杂的防共谋廉价磋商协议。

5 结束语

安全多方、分布式计算、博弈论涉及多个参与方之间的交互活动，如何确保这些交互活动的可信是研究的热点。本文在安全多方的技术上，开展了防共谋安全多方研究，给出了防多轮共谋的形式化定义，并证明在完全公平的基础上，通过一般安全多方协议和通信通道的基础上可以建立可信防共谋协议。最后将协议应用到博弈论中，获得防共谋廉价磋商协议，以克服博弈论在求解防共谋相关均衡时需要可信方这一不足。笔者的工作是建立在安全多方的非自适应攻击模型下，接下来的研究将扩充到自适应攻击模型，以获得

动态防共谋安全多方协议。以及将研究拓展到博弈论中的概率策略空间上，以获得更加实际的应用。在本文中，尽管在理论上给出了可信防共谋模型协议的存在性证明，但正如 Goldreich 所言，文献[2]所给出的多方安全计算的通用解决方案效率过低[19]，在实际应用中设计高效可用的安全多方协议仍然十分困难，期望接下来的研究中探讨更加实用的协议模型。另一方面，由于博弈论是一个很好的网络社会学研究工具，随着互联网应用的深入，博弈论与多方安全计算结合起来进行网络安全交互研究，也是目前国际研究的一个热点，而文献[4~6]也明确指出，将共谋方设定相互独立，不能进行隐秘通信这一主要思想也是来源于博弈论。文献[20~23]也给出了密码学与博弈论想结合的一些研究方向，也期望进一步开展这方面的研究工作。

参考文献：

[1] YAO A C. How to generate and exchange secrets[A]. Proc 27th FOCS IEEE Computer Society[C]. 1986.162-167.

[2] GOLDREICH O, MICALI S, WIGDERSON A. How to play any mental game[A]. Proc 19th Symposium on Theory of Computing (STOC)[C]. ACM, 1987. 218-229.

[3] CLEVE R. Limits on the security of coin flips when half the processors are faulty[A]. 18th ACM Symposium on the Theory of Computing[C]. 1986. 364-369.

[4] LEPINSKI M, MICALI S, SHELAT A. Collusion-free protocols[A]. Proc 37th Annual ACM Symposium on Theory of Computing (STOC)[C]. New York, NY, USA, ACM, 2005.543-552.

[5] ALWEN J, SHELAT A, VISCONTI I. Collusion-free protocols in the mediated model[A]. Cryptology-Crypto[C]. Springer, 2008. 497-514.

[6] ALWEN J, KATZ J, LINDELL Y, *et al*. Collusion-free multiparty computation in the mediated model[A]. 29th International Cryptology Conference[C]. Santa Barbara, CA, 2009.

[7] LEPINSKI M, MICALI S, PEIKERT C, *et al*. Completely fair SFE and coalition-safe cheap talk[A]. Proceedings of the 23rd Annual ACM Symposium on Principles of Distributed Computing[C]. 2004. 1-10.

[8] GOLDREICH O. Foundations of Cryptography[M]. (Volume 2 Basic Applications:), Cambridge University Press, Pub, 2004.

[9] CANETTI R, FEIGE U, GOLDREICH O, *et al*. Adaptively secure computation[A]. 28th Symposium on Theory of Computing (STOC), ACM[C]. 1996.

[10] MICALI G S, RACKOFF C. The knowledge complexity of interactive proof system[J]. SIAM Journal on Computing,1989,(18):186-208.

[11] KILIAN J, Basing cryptographt on oblivious transfer[A]. 20th ACM Symposium on the Theory of Computing[C]. 1988. 20-31.

[12] CHOR B, GOLDWASSER S, MICALI S, *et al*. Verifiable secret sharing and achieving simultaneity in the presence of faults[A]. Proceeding 26th Annual Symposium on the Foundations of Computer Science[C]. IEEE, 1985. 383-395.

[13] GOLDREICH O. Foundations of Cryptography[M] (Volume 1 Basic Tools:), Cambridge University Press.

[14] BLUM M, SANTIS A D, MICALI S, *et al*. Noninteractive zero-knowledge[J]. SIAM J Computing, 1991, 20(6): 1084-1118.

[15] MORENO D, WOODERS J. Coalition-proof equilibrium[J]. Games Econ Beha, 1996,(17): 80-112.

[16] AUMANN R. Subjectivity and correlation in randomized strategies[J]. J Math Econ, 1974,(1): 67-96.

[17] Barany. Fair distribution protocols or how the players replace fortune[J]. Mathematics of Operation Research, 1992,(17):327-341.

[18] RAY I. coalition-proof correlated equilibrium: a definition[J]. Games Econ. Beha, 1996, (17):56-79.

[19] 李顺东, 王道顺. 现代密码学:理论,方法与研究前沿[M]. 北京: 科学出版社, 2009.
LI D S, WANG D S. Modern Cryptography: Theory, Method and Cutting-Edge Research[M]. Beijing: Science Press, 2009.

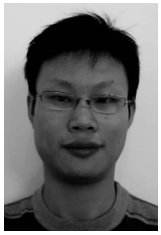
[20] ABRAHAM I, DOLEV D, GONEN R, *et al*. Distributed computing meets game theory: robust mechanisms for rational secret sharing and multiparty computation[A]. Proceedings of the 25th ACM Symposium on Principles of Distributed Computing (PODC)[C]. 2006.53-62.

[21] KOL G, NAOR M. Cryptography and game theory: designing protocols for exchanging information[A]. Proceedings of the 5th Theory of Cryptography Conference (TCC)[C]. 2008.

[22] LYSYANSKAYA A, TRIANDOPOULOS N. Rationality and adversarial behavior in multi-party computation[A]. Cryptology- CRYPTO[C]. 2006.180-197.

[23] KATZ J. Bridging cryptography and game theory. recent results and future directions(invited paper)[A]. 5th Theory of Cryptography Conference, TCC[C]. 2008.

作者简介：



程柏良（1978-），男，江西波阳人，同济大学博士生，主要研究方向为可信计算、智能 agent。

曾国荪（1964-），男，江西新干人，博士，同济大学教授、博士生导师，主要研究方向为网格计算、可信软件。

揭安全（1975-），男，江西广昌人，江西师范大学副教授，主要研究方向为并行计算、信息检索与智能处理。