隐私保护的推理机策略加密及应用

张明武[1),4)] 杨 波[2)] 王春枝[1)] TAKAGI Tsuyoshi[3)]

[1)](湖北工业大学计算机学院 武汉 430068)

[2)](陕西师范大学计算机学院 西安 710072)

[3)](九州大学工业数学研究所 福冈 819-0392 日本)

[4)](中国科学院信息工程研究所信息安全国家重点实验室 北京 100093)

摘 要 确定性有限自动状态机是能表示有限个状态以及在这些状态之间转移和动作等行为的数学模型.本文提出两种基于有限状态自动机策略的加密方案:第 1 种方案称为无消息负载方案,方案中密文关联到一有限自动机 M 而令牌关联到一个任意长的输入串 $\boldsymbol{w}$,系统能在密文空间和密钥空间测试是否令牌关联的输入串可以被密文中的自动机接受.同时给出了转换到素数阶群构建的方法.第 2 种方案以前种方案为原语,扩展到支持消息负载保密的方案,当自动机接受输入串时,可以成功从密文中提取明文.采用双系统加密技术,在静态安全假设下证明了该方案达到标准模型下自适应语义安全性.同时给出了两种方案的性能评测.所提出的方案可应用于隐私保护的安全外包计算、网络防火墙内容过滤、模板隐私保护的 DNA 比对等领域,文中给出了实际应用中的具体案例.

关键词 有限自动机;双系统加密;隐私保护;语义安全

中图法分类号 TP309 **DOI 号** 10.3724/SP.J.1016.2015.00897

Privacy-Preserving and Adaptively-Secure Encryptions with Deterministic Finite Automata Policy and Their Applications

ZHANG Ming-Wu[1),4)] YANG Bo[2)] WANG Chun-Zhi[1)] TAKAGI Tsuyoshi[3)]

[1)](*School of Computer Sciences, Hubei University of Technology, Wuhan* 430068)

[2)](*School of Computers, Shaanxi Normal University, Xi'an* 710072)

[3)](*Institute of Mathematics for Industry, Kyushu University, Fukuoka* 819-0392, *Japan*)

[4)](*State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing* 100093)

Abstract Deterministic Finite Automata (DFA) is a useful mathematical tool in defining the finite states and describing the transition among these states. In this paper, we propose two adaptively secure functional encryptions in the standard model that are based on DFA policies. In the first scheme, the ciphertext is associated with a DFA M and the token is associated with an arbitrary length string $\boldsymbol{w}$, and there is a check algorithm to test whether the string $\boldsymbol{w}$ is accepted by the automata M in the key/ciphertext spaces. In the second scheme, we extend the first scheme to support payload confidentiality, in which the decryption can extract the encrypted message if the associated automata accepts the string. Using the technique of dual system encryption, we prove the schemes can achieve adaptive security under the static assumptions, and we then give the performance evaluation. We also provide the deployments in privacy-preserving outsource computation in cloud, text filtering in firewall, and privacy-carrying DNA match in decentralized network etc.

收稿日期:2013-08-12;最终修改稿收到日期:2014-11-28.本课题得到国家自然科学基金(61370224,61272436)、湖北省自然科学基金重点项目(2013CFA046)、中国科学院信息工程研究所信息安全国家重点实验室开放课题、湖北工业大学高层次人才项目计划资助.**张明武**,男,博士,教授,中国计算机学会(CCF)高级会员,主要研究领域为网络与信息安全、隐私保护技术.E-mail:csmwzhang@gmail.com.**杨 波**,男,1963 年生,博士,教授,主要研究领域为密码学与信息安全.**王春枝**,女,1963 年生,博士,教授,中国计算机学会(CCF)会员,主要研究领域为网络协议与安全.**TAKAGI Tsuyoshi**,男,博士,教授,主要研究领域为密码学.

Keywords deterministic finite automata; dual system encryption; privacy preservation; semantic security

1 引 言

有限状态机是自动机理论和计算理论中重要工具,主要功能是描述对象在其生命周期内所经历的状态序列,以及如何响应来自外界的各种事件,它在硬件系统设计、软件工程、编译器、网络协议、计算语言、软件集成分析、网络漏洞分析等领域得到广泛的应用[1-2].一个有限状态自动机中存在有限数量的状态,每个状态可以迁移到零个或多个状态,输入字串决定执行哪个状态的迁移.有限状态自动机可以表示为一个有向图结构,可以分成确定型有限状态自动机(Deterministic Finite Automata,DFA)[1]与非确定型有限状态自动机(Non-deterministic Finite Automata,NFA)[2]两种.在确定型自动机模型中,每个状态对每个可能输入只有精确的一个转移.在非确定型自动机中,给定状态对可能的输入可以没有或存在多个转移.一个非确定有限状态自动机可以转化为确定有限状态自动机,有限状态自动机识别的语言称为正则语言.

状态机的本质是对具有逻辑顺序或时序规律事件进行描述,可以用作软件程序的控制结构.但在一些应用中,采用明文有限自动机的证明,会泄露程序的流程结构,对软件保护产生不利的影响.例如,传统的编译代码中,程序在得到输入串后是以编译后的可执行代码运行,最后输出运行结果.但由于编译后的可执行程序没有加密,因此很容易被反编译并破解程序流程.

函数加密[3-5]是一种灵活的加密技术,可实现细粒度的密文控制,克服了传统的基于公钥和身份加密方案的单一性,它采用灵活的策略函数定义接收者的解密能力.加密策略定义为策略函数 F,而解密角色定义为属性向量 $\boldsymbol{w}$.当接收者所持有的属性 $\boldsymbol{w}$ 满足密文所关联的函数 F(即 $F(\boldsymbol{w})=1$)时,使用其持有的密钥可以成功解密密文.事实上,公钥加密[4-6]、(层次)基于身份加密[7]、广播加密[8]、属性基加密[9-10]、隐藏向量加密[11-12]、内积加密[13-14]、空间加密[15-16]和代理加密[17]等均是函数加密的一个特例.例如,基于身份的加密方案中,策略函数表达为相等测试函数 $F(I,I'):I'=I$,即加密策略与解密角色定义为简单的单射函数.广播加密方案中策略函数定义为元素与集合的包含关系,即当解密角色属于发送者所确定的广播用户集成员时可以利用其所拥有的密钥解密.在属性基加密方案中,加密策略定义为访问结构 Γ 而解密角色定义为属性集合 S,策略函数定义为属性集满足访问结构,即 $\Gamma(S)$.事实上,由于任意层次深度的通用访问结构在有限域上密码技术较难构建,因此研究者通常将访问结构转化为访问树(Access Tree,AT)[9]或线性秘密共享(Linear Secret Sharing Scheme,LSSS)[10]等形式来描述,接收者角色描述为访问结构上一定长向量.

文献[13]定义了基于通用谓词作为加密策略的方案,通过实例化内积谓词实现负载隐藏和属性隐藏的方案,内积加密可以实现包括相等测试、合取、析取、多项式比较、隐藏向量等功能.文献[16]中采用仿射空间上的仿射函数作为加密策略和解密角色,实现抗密钥加密,其中函数匹配定义为两个仿射空间存在最小不动点.文献[18]构造出在正则语言上的选择性安全函数加密,但其安全性归约基于非标准假设.文献[19]中首次基于合数阶群上的子群判定问题,构建了密文上不同功能的函数查询,如子集查询、范围查询、合取和析取查询等.文献[6]中设计一种基于有限自动机可逆理论的公钥方案,安全性基于求非线性有限自动机的弱逆的困难性和矩阵多项式的因式分解困难性基础上.与文献[6]不同之处在于,本文主要思路在于以有限自动机作为函数策略生成密文,并以满足自动机的串作为解密角色,安全性基于数论上的子群判定问题.文献[3]给出了一般意义上的函数加密模型,指出通用意义上的函数加密很难构建和证明,只能在具体可操作的函数上针对其特点进行设计.

有限状态自动机可以看作一类特殊的函数,即自动机定义为一函数而作用在自动机上的输入串作为函数的输入,自动机是否接受输入串作为函数的输出,即 $f_M(\boldsymbol{w})=\mathrm{Accept}(M,\boldsymbol{w})$.与传统的函数加密相比,基于有限自动状态机策略的加密方案有如下优点:(1)有限状态自动机是随逻辑顺序或时序规律事件而运行的,这种动态的函数结构可以实现在密文空间状态的动态控制功能;(2)输入串是任意长的.传统的函数加密中,都假定输入属性串长度是固定的,因此限制这些方案在实际应用中的灵活性和可扩展性.

基于有限状态自动机策略的加密可对密文解密能力实现细粒度控制，保护密文中自动机的机密性，而解密属性串长度不受限制，极大地改进加密系统控制的灵活性、降低系统的通信代价和提高系统的计算效率，在外包计算、访问控制、隐私保护、防火墙内容过滤、数据搜索与挖掘、生物信息保护等方面有良好的应用前景[20-24]. 文献[25]提出适应于云数据共享环境下以有限状态推理机为代理策略的代理加密机制. 文献[26]提出一种基于密钥泄漏条件下的推理机控制策略加密机制.

本文首先设计一种采用有限状态自动机作为密文策略而输入串作为令牌角色的无消息负载加密方案. 方案中密文关联到一个有限自动机 M，令牌关联到一个任意长的输入串 $\boldsymbol{w}$，系统能在密文空间和秘密令牌空间通过配对运算测试是否令牌关联的输入串可以被密文中的自动机接受. 本文给出了该方案详细构造和安全性证明，并且给出方案在相关领域中的具体应用. 同时，本文扩展第一种方案达到有消息负载的方案，当密文关联的自动机接受密钥关联的串时，可以成功从密文中提取消息明文.

与其它加密方案相比，本文方案设计的困难在于：(1) 有限自动机不是通用意义上的函数结构，它定义为五元组的数据结构以及在这个结构上的随输入串而变化的自动机模型，首要的关键问题是把其理论数学上的结构模型转换为密码技术可表达和运算的有限域上的模型；(2) 自动机要根据不同输入串而动态的运行，这在明文空间比较容易实现，但转化在密文和秘密令牌空间来实现这种动态变迁，存在一定难度；(3) 为达到自适应性安全性，方案设计和安全证明技术必须采用与之相适应的安全假设和证明框架. 相对文献[26]方案，虽然本方案不能抗密钥的泄漏攻击，但本文方案效率更高. 与文献[18]的非标准假设下选择性安全相比，本文方案达到自适应安全性下的自动机隐私，且密钥和密文结构更紧凑.

2　基础知识

状态(State)：指对象在其生命周期中的状况，处于某个特定状态中的对象必然会满足某些条件、执行某些动作或者是等待某些事件.

事件(Event)：指在时间和空间上占有一定位置，事件通常会引起自动机状态的变迁，促使状态机从一种状态切换到另一种状态.

转移(Transition)：指两个状态之间的一种关系，表示对象将在一个状态中执行一定的动作，并将在某个事件发生同时某个特定条件满足时进入另外一个状态.

动作(Action)：指状态机中可以执行的原子操作，一个动作在运行过程中不能被其他消息所中断.

定义 1(确定有限状态自动机，DFA).　确定有限状态自动机 $M=(Q,\Sigma,\delta,s_0,F)$ 定义为五元组，其中：

(1) Q 是非空有限的状态集合；

(2) Σ 是非空有限的字符集合的输入字母表；

(3) $\delta: Q\times\Sigma\to Q$ 是转移函数；

(4) $s_0(\in Q)$ 是开始状态；

(5) $F\subseteq Q$ 是接受状态的集合.

确定有限状态自动机从起始状态 s_0 开始，依次读入字符串，并根据给定的转移函数 δ 依次地转移至下一个状态. 在读完该字符串后，如果该自动机停在一个属于 F 的接受状态，就接受该字符串，反之则拒绝该字符串. 若自动机 M 接受串 $\boldsymbol{w}$，记为 $\mathrm{Accept}(M,\boldsymbol{w})$，否则记为 $\mathrm{Reject}(M,\boldsymbol{w})$. 本文为简化表达，记 $\mathrm{Accept}(M,\boldsymbol{w})=1$ 表示自动机 M 接受串 $\boldsymbol{w}$，$\mathrm{Accept}(M,\boldsymbol{w})=0$ 表示 M 拒绝串 $\boldsymbol{w}$.

定义 2(扩展转移函数).　设 $\delta^*: Q\times\Sigma^*\to Q$，这里 $\delta^*(q,\boldsymbol{w})$ 是自动机从状态 q 顺序读入串 $\boldsymbol{w}$ 后到达的状态. 扩展转移函数的形式化递归定义满足如下两个性质：

(1) $\delta^*(q,\varepsilon)=q$；

(2) $\delta^*(q,u\sigma)=\delta(\delta^*(q,u),\sigma),\ \forall\delta^*\in\Sigma^*\ \&.\ \sigma\in\Sigma$

定义 3(转移系统).　没有接受状态列表和指定开始状态的确定有限状态机叫做转移系统或半自动机.

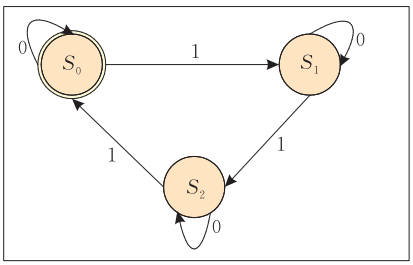
对于一个确定有限状态自动机 $M=(Q,\Sigma,\delta,s_0,F)$，如果 $\delta^*(s,\boldsymbol{w})\in F$，称该自动机 M 接受字符串 $\boldsymbol{w}$，否则表明该自动机拒绝字符串 $\boldsymbol{w}$. 被一个确定有限自动机 M 接受的语言定义为

$$L(M)=\{\boldsymbol{w}\in\Sigma^*\mid M\ \text{Accept string}\ \boldsymbol{w}\}$$

该语言定义为所有被接受的字符串组成的集合. 为方便表达，本文使用符号 Γ 作为 δ 函数的转移集合，即 $t=(x,y,\sigma)\in\Gamma\Leftrightarrow\delta(x,\sigma)=y$. 这里 x 是当前状态，y 是下一个状态，δ 是当前输入的字符.

图 1 给出一个简单的转移函数实例，设接受状态集合 $F=\{S_0,S_2\}$. 该自动机的初始状态是 S_0，转移函数是 $t_i(i=1,\cdots,6)$. 若输入串 $\boldsymbol{w}$ 是 01010，则该自动机所确定的转移状态序列是 $S_0\xrightarrow{0}S_0\xrightarrow{1}$

$S_1 \xrightarrow{0} S_1 \xrightarrow{1} S_2 \xrightarrow{0} S_2$，所对应的转移函数序列是 $t_1 \to t_2 \to t_3 \to t_4 \to t_5$. 经过输入串的状态转移，最后自动机停在 S_2 状态，而 $S_2 \in F$，因此该自动机接受串 $\boldsymbol{w}$，即 Accept$(M,\boldsymbol{w})=1$.



t_i	x	y	δ
t_1	S_0	S_0	0
t_2	S_0	S_1	1
t_3	S_1	S_1	0
t_4	S_1	S_2	1
t_5	S_2	S_2	0
t_6	S_2	S_0	1

图 1 确定有限自动状态机转移函数实例

接受函数 Accept$(M,\boldsymbol{w})$用于表达基于确定有限自动机 $M=(Q,\Sigma,\delta,q_0,F)$对串 $\boldsymbol{w}=(w_1,w_2,\cdots,w_n)$的运行及输出：对于任何 M，设初值 $r_0=q_0$；对给定的输入序列 $w_i(i=0,\cdots,n-1)$，$r_{i+1}=\delta(r_i,w_{i+1})$；若 $r_n\in F$，输出 1，否则输出 0. 即

$$\text{Accept}(M,\boldsymbol{w}) = \left[r_n \in F \,\middle|\, \begin{array}{c} M=(Q,\sum,\delta,q_0,F) \\ w=(w_1,\cdots,w_n)\in\{0,1\}^n \\ r_0 \leftarrow q_0 \\ \text{For } i\in[0,n-1], r_{i+1}\leftarrow\delta(r_i,w_{i+1}) \end{array} \right] \tag{1}$$

显然，若串 $\boldsymbol{w}$ 满足自动机 M，则 Accept$(M,\boldsymbol{w})=1$，否则 Accept$(M,\boldsymbol{w})=0$.

非确定有限自动机是对每个状态和输入符号对可以有多个可能的下一状态的有限状态自动机. 区别于确定有限状态自动机 DFA，NFA 的下一个可能状态是不唯一确定的. 虽然 DFA 与 NFA 有不同的定义，在形式理论中可以证明它们是等价的：对于任何给定 NFA，都可以构造一个与之等价的 DFA，反之亦然. DFA 是非 NFA 的一种极限形式，并且 DFA 在计算能力上等价于 NFA. 本文所设计的加密方案中，只考虑基于确定有限自动机 DFA 策略的方案.

3 DFA 策略加密体制

3.1 确定有限状态自动机加密体制

确定有限自动机策略的加密体制中，密文对应于一个有限自动机 M，而秘密令牌对应一个输入串 $\boldsymbol{w}$，解文/令牌测试能够成功的必要条件是 Accept$(M,\boldsymbol{w})=1$. 本文首先给出无消息负载的方案，它由 4 个概率性随机算法构成：$L=$(Setup, GenTK, Enc, Check). 然后通过将该方案扩展，达到支持负载保密的确定性有限自动机策略加密 $L^*=$(Setup*, GenTK*, Enc*, Dec*).

定义 4(确定有限状态自动机策略加密). 1 个确定有限状态自动机策略加密方案 L 包括 4 个概率多项式时间的算法，$L=$(Setup, GenTK, Enc, Check)，对所有 $M\in$DFA 和 $w\in\{0,1\}^*$.

(1)(PubKey, MstKey)$\leftarrow$Setup$(1^\lambda,\Sigma)$(生成系统主公/密钥对)

(2) Tok$_w\leftarrow$GenTK(PubKey, MstKey, $\boldsymbol{w}$)(创建串 $\boldsymbol{w}$ 的令牌)

(3) Ctx$_M\leftarrow$Enc(PubKey, M)(生成有限自动机 M 的密文)

(4) 0/1$\leftarrow$Check(PubKey, Ctx$_M$, Tok$_w$)(输出关系 Accept$(M,\boldsymbol{w})$)

对所生成的系统公钥和主密钥，由上述算法生成的任何令牌Tok$_w$和密文Ctx$_M$，满足

(1) 若 Accept$(M,\boldsymbol{w})=1$，则Check(PubKey, Enc(PubKey, M), GenTK(PubKey, MstKey, $\boldsymbol{w}$))$=1$

(2) 若 Accept$(M,\boldsymbol{w})=0$，则 Check(PubKey, Enc(PubKey, M), GenTK(PubKey, MstKey, $\boldsymbol{w}$))$=0$

即 Accept$(M,\boldsymbol{w})$与 Check(PubKey, Enc(PubKey, M), GenTK(PubKey, MstKey, $\boldsymbol{w}$))在计算上是不可区分的.

3.2 安全性模型

定义 5(安全性游戏). 对于 $\beta=0,1$，定义一个敌手 A 在 β 上的游戏 $\text{Exp}^L_{A,\beta}(1^\lambda)$：

(1) 初始化. 运行(PubKey, MstKey)$\leftarrow$Setup$(1^\lambda,\Sigma)$并将 PubKey 发送给 A；

(2) 询问. 对 $i=1,\cdots,q$，A 自适应地提供串 $\boldsymbol{w}$ 并要求获得相应的令牌 Tok$_{wi}$；

(3) 挑战. A 提交两个有限自动机$(M^{(0)},M^{(1)})$并获得对其中随机选择一个自动机的密文 Enc(PubKey, $M^{(\beta)}$)；

(4) 输出. 对 $j=q+1,\cdots,Q$，A 继续执行如前的令牌询问并获得相应令牌. 为避免敌手获得挑战自动机所接受的输入串，此次询问要求 Accept$(M^{(0)},\boldsymbol{w})=$Accept$(M^{(1)},\boldsymbol{w})=0$. 询问结束后 A 输出一个比特作为对 β 的猜测.

对于 $\beta=0,1$，设 W_β 是敌手 A 输出 1 时的事件，定义敌手在 $\mathrm{Exp}_{A,\beta}^{L}(1^\lambda)$ 的优势函数为

$$\mathrm{Adv}_A^L(1^\lambda)=|Pr[W_0]-Pr[W_1]| \tag{2}$$

定义 6（有限自动状态机策略加密语义安全性）. 设敌手 A 在试验 $\mathrm{Exp}_{A,\beta}^{L}(1^\lambda)$ 中执行了 Q 次对有限自动状态机的令牌询问，如果 A 在 $\mathrm{Exp}_{A,\beta}^{L}(1^\lambda)$ 中所获得的优势 $\mathrm{Adv}_A^L(1^\lambda)$ 在安全参数 λ 上是可以忽略的，称有限自动状态机策略加密方案 L 是自适应 Q-语义安全的.

若在 $\mathrm{Exp}_{A,\beta}^{L}(1^\lambda)$ 中，在 Setup 前增加一个预处理 Pre 过程，此过程中敌手 A 提供所需要挑战的两个有限自动机 $(M^{(0)},M^{(1)})$. 若敌手在这个试验游戏中所获得的优势函数输出是可忽略的，称该方案是选择性安全的. 显然，选择性安全要求在系统所有参数生成前敌手就要提供挑战，这样的安全性比定义 6 中的适应安全性要弱.

4　方案构造

4.1　方案设计

(1) Setup$(1^\lambda,\Sigma)$　系统初始化算法以一个安全参数 λ 和一个自动机字母表 Σ 作为输入，为系统生成系统级公钥 PubKey 和主密钥 MstKey，过程如下：以安全参数 λ 为输入，调用合数阶双线性群生成算法生成 $\Phi=(N=p_1p_2p_3,G,G_t,e)\leftarrow Gen(\lambda)$. 这里 p_1,p_2 和 p_3 是长度相等但不相同的素数，即对 $i=1,2,3$，满足 $2^{\lambda-1}\le p_i\le 2^\lambda$；随机选取 $g,z\leftarrow G_{p_1}$，并对任意 $\sigma\in\Sigma$，随机选取 $h_\sigma\leftarrow G_{p_1}$；随机选取 $\alpha\leftarrow Z_N$ 以及 G_{p_3} 的随机生成元 $g_3\leftarrow G_{p_3}$；置系统级公钥 PubKey$=(\Phi,g,z,g_3,Y=e(g,g)^\alpha,\forall\sigma\in\Sigma h_\sigma)$；保存系统主密钥 MstKey$=g^\alpha$.

(2) GenTK (PubKey, MstKey, $\boldsymbol{w}$)　设 $\boldsymbol{w}=\{w_1,\cdots,w_\ell\}$，令牌生成算法首先随机选取 $U,V\leftarrow G_{p_3}$，$s_0,s_1,\cdots,s_\ell\leftarrow Z_N$，$X,Z\leftarrow G_{p_3}^{\ell}$，然后为 $\boldsymbol{w}$ 创建令牌

$$\mathrm{Tok}_w=(A=g^{s_0}U,(K_i,D_i)_{i=1}^{\ell},E=g^{\alpha+s_\ell}V)$$

其中

$$\begin{pmatrix}K_i\\D_i\end{pmatrix}_{i\in[\ell]}=\begin{pmatrix}X\\Z\end{pmatrix}_{i\in[\ell]}\times\begin{pmatrix}g^{s_1} & g^{s_2} & \cdots & g^{s_\ell}\\ (h_{w_1})^{s_1}z^{s_0} & (h_{w_2})^{s_2}z^{s_1} & \cdots & (h_{w_\ell})^{s_\ell}z^{s_{\ell-1}}\end{pmatrix} \tag{3}$$

注意：①U,V,X,Z 在子群 G_{p_3} 中随机选取，它们可以由生成元 g_3 作随机的 Z_N 中的幂指数得到. 由于 p_1,p_2 和 p_3 互素，根据中国剩余定理，$\forall\vartheta\in Z_N$，$g_3^{\vartheta}\bmod N\equiv g_3^{\vartheta}\bmod p_3$；

②秘密令牌空间为 $G_{p_1p_3}^{2\ell+2}$. 根据子群不可区分性，从敌手的角度来看，一个令牌空间为 $G_{p_1p_2p_3}^{2\ell+2}=G^{2\ell+2}$.

(3) Enc(PubKey, M)　设确定有限自动状态机 $M=(Q,\Sigma,\delta,q_0,F)$，$Q$ 个状态为 $Q=\{q_0,q_1,\cdots,q_{\omega-1}\}$，这里 $\omega=|Q|$. 有限自动机 M 的初始状态是 q_0，转换函数 δ 定义为三元组 $\{(x_i,y_i,\sigma_i)\in Q\times Q\times\Sigma\}\in\Gamma$ 的序列. 密文生成算法执行：首先随机选取 ω 个元素 $\gamma_0,\gamma_1,\cdots,\gamma_{\omega-1}\leftarrow Z_N$，对 $i=0,1,\cdots,\omega-1$，计算 $B_i=g^{\gamma_i}$，并用每个 B_i 关联有限自动机的状态 q_i；对每个 $t_i=(x_i,y_i,\sigma_i)\in\Gamma$，随机选择 $r_{t_i}\leftarrow Z_N$，然后计算 $C_{t_i,1}=z^{r_{t_i}}/B_{x_i}$，$C_{t_i,2}=g^{r_{t_i}}$，$C_{t_i,3}=B_{y_i}(h_{\sigma_i})^{r_{t_i}}$；输密文 $\mathrm{Ctx}_M=(B_0,(C_{t_i,1},C_{t_i,2},C_{t_i,3})_{t_i\in\Gamma},\forall x\,B_x\in F,\forall x\,H_x=Y^{\gamma_x})$.

(4) Check(PubKey, Ctx_M, Tok_w)　设 $\mathrm{Ctx}_M=(B_0,(C_{t_i,1},C_{t_i,2},C_{t_i,3})_{t_i\in\Gamma},\forall x\,B_x\in F,\forall x\,H_x=Y^{\gamma_x})$，$\mathrm{Tok}_w=(A,(K_i,D_i)_{i=1}^{\ell},E)$. 若密文 Ctx_M 中关联的有限自动机 $M=(Q,\Sigma,\delta,q_0,F)$ 能接受令牌 Tok_w 中关联串 $\boldsymbol{w}=\{w_1,\cdots,w_\ell\}$，即 Accept$(M,w)=1$，则存在一个 $\ell+1$ 个状态序列 $B_{w_0},B_{w_1},\cdots,B_{w_\ell}$ 和 ℓ 个状态转移 $t_1,\cdots,t_\ell$，满足

$$\begin{cases}B_{w_0}=q_0, & i=0\\ t_i=(B_{w_{i-1}},B_{w_i},w_i)\in\Gamma, & i=1,\cdots,\ell-1\\ B_{w_\ell}\in F, & i=\ell\end{cases} \tag{4}$$

接下来，计算

$$W_0=e(A,B_0)=e(g^{s_0},B_0)=e(g,B_0)^{s_0}.$$

然后由 $i=1,2,\cdots,\ell$，通过递推的方法计算

$$W_i=W_{i-1}\frac{e(C_{t_i,1},K_{i-1})e(C_{t_i,3},K_i)}{e(C_{t_i,2},D_i)}=e(g,B_{w_i})^{s_i} \tag{5}$$

若 Accept$(M,\boldsymbol{w})=1$，在状态转移结束（$\boldsymbol{w}$ 输入完成）后接受状态集 F 中存在 $q_x=B_{w_\ell}$，对应 B_x 满足 $W_\ell=e(g,B_x)^{s_\ell}$. 最后验证等式

$$H_x=e(E,B_x)/W_\ell \tag{6}$$

若式(6)成立，输出 1，否则输出 0.

4.2　正确性与一致性

在式(4)转移函数 $t_i\in\Gamma$ 中，$B_{w_{i-1}}$ 对应于加密时的 x_i，B_{w_i} 则对应于加密时的 y_i，而 w_i 则对应于 σ_i. 在加密过程中，对所有可能的转移函数 t_i 作了盲化处理，而在解密过程中，则通过递推的方法恢复由输入串 $\boldsymbol{w}$ 所决定的转移序列.

解密测试时最开始的初始输入 w_0 对应于 q_0，即

$B_{w_0}=B_0$，计算出对应的配对 $W_0=e(g,B_0)^{s_0}=e(g,B_{w_0})^{s_0}$ 作为递推时的初始值. 接下来我们考查式(5)，由于 Tok_M 中所有元素组件在 $G_{p_1p_3}$ 中而 Ctx_w 中所有组件在 G_{p_1} 中，根据子群的正交性，当令牌组件和密文组件作双线性配对运算时，可以消去 G_{p_3} 部分. 因此

$$W_i=W_{i-1}\frac{e(C_{t_i,1},K_{i-1})e(C_{t_i,3},K_i)}{e(C_{t_i,2},D_i)}$$

$$=\frac{e(g,B_{w_{i-1}})^{s_{i-1}}e(C_{t_i,1},K_{i-1})e(C_{t_i,3},K_i)}{e(C_{t_i,2},D_i)}$$

$$=\frac{e(g,B_{w_{i-1}})^{s_{i-1}}e(z^{r_{t_i}}/B_{w_{i-1}},g^{s_{i-1}})e(B_{w_i}(h_{w_i})^{r_{t_i}},g^{s_i})}{e(g^{r_{t_i}},(h_{w_i})^{s_i}z^{s_{i-1}})}$$

$$=\frac{e(g,B_{w_{i-1}})^{s_{i-1}}e(z^{r_{t_i}},g^{s_{i-1}})e(B_{w_i},g^{s_i})e((h_{w_i})^{r_{t_i}},g^{s_i})}{e(B_{w_{i-1}},g^{s_{i-1}})e(g^{r_{t_i}},(h_{w_i})^{s_i})e(g^{r_{t_i}},z^{s_{i-1}})}$$

$$=e(B_{w_i},g^{s_i})$$

$$=e(g,B_{w_i})^{s_i} \tag{7}$$

考查式(6)，$W_\ell=e(g,B_{w_\ell})^{s_\ell}$，$H_x=Y^{\gamma_x}$，$B_x=g^{\gamma_x}$，所以

$$e(E,B_x)/W_\ell=\frac{e(g^{\alpha+s_\ell},g^{\gamma_x})}{e(g,B_{w_\ell})^{s_\ell}}=\frac{e(g^\alpha,g^{\gamma_x})e(g^{s_\ell},g^{\gamma_x})}{e(g,B_{w_\ell})^{s_\ell}}$$

$$=\frac{e(g^\alpha,g^{\gamma_x})e(g,B_x)^{s_\ell}}{e(g,B_{w_\ell})^{s_\ell}}$$

$$=e(g^\alpha,g^{\gamma_x})=e(g,g)^{\alpha\gamma_x}=Y^{\gamma_x} \tag{8}$$

5 安全性

定理 1[13]. 设 $N=\prod_{i=1}^{m}p_i$ 为两两不同的素数之积且 N 在多项式时间内不能被分解出 p_i，且对 $i\in[m]$，有 $p_i>2^\lambda$. $\Omega=(N,G,G_t,e)$ 是一满足双线性映射的阶为 N 的有限群描述，$\{A_i\}$ 是群 G 上一组随机变量，$\{B_i\}$ 是群 G_t 上一组随机变量，T_0，T_1 是群 G 或 G_t 上的随机变量，且所涉及随机变量阶不高于 t. 考虑如下试验：算法 A 获得实例 $(\Omega,\{A_i\},\{B_i\})$ 以及 T_0 和 T_1 上随机选择一元素 $T_\beta(\beta\in\{0,1\})$，A 在已有知识情况下输出对 β 的猜测 β'. 算法 A 猜测成功的优势定义为 $Adv=|\Pr[\beta'=0]-\Pr[\beta'=1]|$.

若算法 A 在上述试验中最多进行了 Q 次操作并获得不可忽略 ε 大小的优势，在 T_0 和 T_1 独立于 $\{B_i\}\bigcup\{e(A_i,A_j)\}$ 的情况下，可以构造另一算法以 $\varepsilon-O(Q^2t/2^\lambda)$ 优势分解合数 N.

根据定理 1，在 $N=p_1p_2p_3$ 为三素数积阶的双线性群描述 $(\Phi=(N,G,G_t,e))$ 中，可以推导出如下推论.

推论 1. 设 $g,X_1\in G_{p_1}$，$X_2\in G_{p_2}$，$X_3\in G_{p_3}$，$T_0\in G_{p_1p_2}$，$T_1\in G_{p_1}$，$\beta\in\{0,1\}$，$T=\beta T_0+(1-\beta)T_1$，区分 $(\Phi,g,X_1X_2,X_3,T=T_0)$ 和 $(\Phi,g,X_1X_2,X_3,T=T_1)$ 在计算上是不可行的.

推论 2. 设 $g,X_1,Z_1\in G_{p_1}$，$X_2,Y_2,Z_2\in G_{p_2}$，$X_3,Y_3,Z_3\in G_{p_3}$，$T_0=Z_1Z_3$，$T_1=Z_1Z_2Z_3$，$\beta\in\{0,1\}$，$T=\beta T_0+(1-\beta)T_1$，区分 $(\Omega,g,X_1X_2,X_3,Y_2Y_3,T=T_0)$ 和 $(\Omega,g,X_1X_2,X_3,Y_2Y_3,T=T_1)$ 在计算上是不可行的.

双系统加密. 双系统加密技术是 Waters 最先提出的实现适应性安全的基于身份的安全性证明方法[7]，后来 Lewko 和 Waters[27] 将其扩展到合数阶群以实现简单的假设和固定长度密文. 双系统加密核心思想是通过一系列在计算上不可区分的转换方法，将正常的挑战密文和询问秘密钥转换到半功能化状态(semi-functional)，而根据双系统加密的性质，正常密钥可以解密正常或半功能化形式的密文，半功能化的密钥只能对正常密文形式的密文解密，而半功能化的秘密钥在解密半功能化密文时无法成功，其可解密优势是可忽略的.

双系统加密能达到上述功能的主要原因是合数阶群存在阶是 N 因子的子群，且这多个子群是正交的. 例如，在我们的设计的阶为 $N=p_1p_2p_3$ 的双线性群中，由于 $p_1\neq p_2\neq p_3$，对 $i\neq j$，$gcd(p_i,p_j)=1$，对 G 的生成元 g，子群 G_{p_1} 的生成元是 $g^{p_2p_3}$，G_{p_2} 的生成元是 $g^{p_1p_3}$. 对任意 $h_1\in G_{p_1}$，$h_2\in G_{p_2}$，设 h_1 和 h_2 分别是 G_{p_1} 和 G_{p_2} 生成元作指数 a_1 和 a_2 得到，即 $h_1=(g^{p_2p_3})^{a_1}$，$h_2=(g^{p_1p_3})^{a_2}$，则

$$e(h_1,h_2)=e((g^{p_2p_3})^{a_1},(g^{p_1p_3})^{a_2})$$

$$=(e(g,g)^{p_1p_2p_3})^{a_1a_2p_3}$$

$$=(e(g,g)^N)^{a_1a_2p_3}=1 \tag{9}$$

系统构建中，如图 2(b)，令牌在 $G_{p_1p_3}$ 子群，而密文在 G_{p_1} 子群. 为达到自适应性安全(不需要敌手在系统参数生成前提交挑战的有限自动机)，本文通过 G_{p_2} 子群来隐藏敌手可能存在的匹配询问. 事实上，系统只公开了合数阶 N 及对应的群 (G,G_t,e) 和子群 G_{p_1} 和 G_{p_3} 的群元素，对子群 G_{p_1} 和 G_{p_3} 元素的选取可以利用其生成元作 Z_N 上的指数运算得到. 得到如下定理.

定理 2. $g\in G_{p_1}$，$r\in Z_N$，则 $g^r\in G_{p_1}$.

证明. $r\in Z_N$，$g^r=g^{r\bmod N}=g^{r\bmod p_1p_2p_3}$. 而 p_1，p_2，p_3 互素，即对 $i\neq j$，$gcd(p_i,p_j)=1$，根据中国剩余定理，$g^{r\bmod p_1p_2p_3}=g^{r\bmod p_1}\in G_{p_1}$. 证毕.

定理 3.　设 $N=\prod_{i=1}^{L} p_i$ 是不可因式分解的，π_0，$\pi_1 \subseteq \{p_1, p_2, \cdots, p_L\}$ 满足 $gcd(\pi_1, \pi_2)=1$，对 (N, G, G_t) 的子群元素 $g_{\pi_1} \in G_{\pi_1}$，$g_{\pi_1\pi_2} \in G_{\pi_1\pi_2}$，则 g_{π_1} 和 $g_{\pi_1\pi_2}$ 在计算上是不可区分的.

证明.　由于双线性群只公开阶 N 而未公开 p_i，区分 g_{π_1} 和 $g_{\pi_1} g_{\pi_2}$ 等价于区分 π_1 和 $\pi_1\pi_2$. 这 π_1 和 π_2 都是 N 的因子，要区分 N 的因子等价于 N 可以因式分解.　　证毕.

根据定理 3，可以得到如下两个推论.

推论 3.　对确定性有状态机加密密文 $\mathrm{Ctx}_M=(B_0, (C_{t_i,1}, C_{t_i,2}, C_{t_i,3})_{t_i\in\Gamma}, \forall x\ B_x\in F, \forall x\ H_x=Y^{\gamma_x})$，对 $B_0, (C_{t_i,1}, C_{t_i,2}, C_{t_i,3})_{t_i\in\Gamma}, \forall x\ B_x\in F$ 每个组件元素对应乘以 G_{p_2} 中的随机元素得到 $\overrightarrow{\mathrm{Ctx}}_M$，则 $\overrightarrow{\mathrm{Ctx}}_M$ 与 Ctx_M 是计算上不可区分的.

推论 4.　对确定性有状态状态机加密令牌 $\mathrm{Tok}_w=(A=g^{s_0}U, (K_i, D_i)_{i=1}^{\ell}, E=g^{\alpha+s_\ell}V)$，每个令牌组件元素乘以 G_{p_2} 中的随机元素得到 $\overrightarrow{\mathrm{Tok}}_w$，则 $\overrightarrow{\mathrm{Tok}}_w$ 与 Tok_w 是计算上不可区分的.

所构建方案 L 中，秘密令牌在 $G_{p_1p_3}$ 子群中而密文在 G_{p_1} 子群中（H_x 除外），而 G_{p_2} 子群在实际方案中没有使用，主要用于安全性证明. 借助于双系统加密的证明思想[7,27]，半功能化的令牌和密文生成方法是在原来令牌或密文上附加随机的 G_{p_2} 部分. 设 $g_2\in G_{p_2}$ 为 G_{p_2} 的生成元.

(1) 半功能密文. 设 $\mathrm{Ctx}_M=(B_0, (C_{t_i,1}, C_{t_i,2}, C_{t_i,3})_{t_i\in\Gamma}, \forall x\ B_x\in F, \forall x\ H_x=Y^{\gamma_x})$ 为 Enc 算法生成的正常密文，一个半功能密文的构造：随机选取 $b_0\in Z_N$，对 $t_i\in\Gamma$，随机选取 $c_{t_i,1}, c_{t_i,2}, c_{t_i,3}\in Z_N$，对 $B_x\in F$，选取 $b_x\in Z_N$，输出半功能密文 $\overrightarrow{\mathrm{Ctx}}_M=(B_0 g_2^{b_0}, (C_{t_i,1} g_2^{c_{t_i,1}}, C_{t_i,2} g_2^{c_{t_i,2}}, C_{t_i,3} g_2^{c_{t_i,3}})_{t_i\in\Gamma}, \forall x\ B_x g_2^{b_x}\in F, \forall x\ H_x=Y^{\gamma_x})$.

(2) 半功能令牌. 设 $\mathrm{Tok}_w=(A=g^{s_0}U, (K_i, D_i)_{i=1}^{\ell}, E=g^{\alpha+s_\ell}V)$ 是 GenTK 算法生成的令牌，一个半功能令牌构造：随机选择 $a, e\in Z_N$，对 $i\in[\ell]$，选取 $k_i, d_i\in Z_N$，输出半功能令牌 $\overrightarrow{\mathrm{Tok}}_w=(Ag_2^a, (K_i g_2^{k_i}, D_i g_2^{d_i})_{i=1}^{\ell}, Eg_2^e)$.

显然，使用 Check 算法，一个半功能令牌 $\overrightarrow{\mathrm{Tok}}_w$ 检测一半功能密文 $\overrightarrow{\mathrm{Ctx}}_M$ 时，除了 G_{p_1} 部分要匹配外，在测试过程中，由于半功能令牌和密文中都有 G_{p_2} 部分，在进行配对运算时，将会产生额外的 $e(g_2, g_2)$ 部分，而 G_{p_3} 部分只在令牌中有，而在密文中没有，配对运算中可约去. 在 W_i 迭代运算时，测试产生的 G_{p_2} 部分是

$$
\begin{aligned}
W_0:\quad & e(g_2,g_2)^{ab_0}\\
& \theta_0 = ab_0\\
W_1:\quad & e(g_2,g_2)^{\theta_0+c_{t_1,1}k_0+c_{t_1,3}k_1-c_{t_1,2}d_1}\\
& \theta_1 = \theta_0 + c_{t_1,1}k_0 + c_{t_1,3}k_1 - c_{t_1,2}d_1\\
\vdots\quad & \qquad\vdots\\
W_\ell:\quad & e(g_2,g_2)^{\theta_{\ell-1}+c_{t_\ell,1}k_{\ell-1}+c_{t_\ell,3}k_\ell-c_{t_\ell,2}d_\ell}\\
& \theta_\ell = \theta_{\ell-1} + c_{t_\ell,1}k_{\ell-1} + c_{t_\ell,3}k_\ell - c_{t_\ell,2}d_\ell\\
& \quad = e(g_2,g_2)^{ab_0+\sum_{i=1}^{\ell}(c_{t_i,1}k_{i-1}+c_{t_i,3}k_i-c_{t_i,2}d_i)}\\
& \quad = ab_0 + \sum_{i=1}^{\ell}(c_{t_i,1}k_{i-1} + c_{t_i,3}k_i - c_{t_i,2}d_i)
\end{aligned}
\tag{10}
$$

在计算 H_x 过程中，产生 G_{p_2} 部分是 $e(g_2, g_2)^{eb_x-\theta_\ell} = e(g_2,g_2)^{eb_x-ab_0-\sum_{i=1}^{\ell}(c_{t_i,1}k_{i-1}+c_{t_i,3}k_i-c_{t_i,2}d_i)}$. 若半功能密钥能正确测试半功能密文，要求 $eb_x - ab_0 - \sum_{i=1}^{\ell}(c_{t_i,1}k_{i-1} + c_{t_i,3}k_i - c_{t_i,2}d_i) = 0 \bmod p_2$. 而该式中所有变量都是随机选取的，使得结果为 0 的可能性是 $1/p_2$. 事实上，p_2 是 G_{p_2} 的阶，对安全参数 λ 来说，$2^{\lambda-1}\le p_2\le 2^{\lambda}$ 是足够安全的，即半功能令牌无法解密测试半功能密文.

证明中首先把挑战密文 Ctx_M^* 转换为半功能化的形式 $\overrightarrow{\mathrm{Ctx}}_M^*$，根据推论 1，这种转换在计算上是不可区分的，敌手无法感知这种转换. 即若攻击者可以以不可忽略的优势 ε 区分这两种分布，则可以构造一算法，以攻击者计算能力作为子算法，以 $O(\varepsilon)$ 的优势区分定理 3 的实例，其中，$\pi_1=p_1$，$\pi_2=p_2$. 给出如下的定理.

定理 4.　任何多项式时间的敌手无法区分正常形态的密文和半功能形态的密文.

证明.　采用反证法，假定一敌手 A 以不可忽略的概率优势区分正常形态的密文和半功能形态的密文，我们将利用 A 作为子算法，构建算法 D 作为模拟器解决定理 3 所陈述的对 $\pi_1=p_1$ 和 $\pi_2=p_2$ 下的不可区分性.

D 扮演系统模拟器并回答敌手 A 的询问：

初始化. 随机选择 $\alpha, a, b_i\in Z_N^{2+|\Sigma|}$，计算 $z=g^a$，$\forall\sigma\in\Sigma h_\sigma=g^{b_i}$，$Y=e(g,g)^{\alpha}$. D 公开双线性群描述以及公钥 $\mathrm{PubKey}=(g, z, g_3, Y, \forall\sigma\in\Sigma h_\sigma)$. 同时模拟器 D 保存系统主密钥 $\mathrm{MstKey}=g^{\alpha}$.

询问. 基于 D 自己选取的随机数，包括主密钥 α，它可以为任何输入串 $w=\{w_1, \cdots, w_\ell\}$ 生成令牌，

直接调用令牌生成算法即可.同时由于它拥有主密钥,可以回答任何敌手在定义 5 中安全游戏所设定的令牌询问.

挑战.本阶段 A 两个自动机($M^{(0)}$,$M^{(1)}$),以便模拟器 D 生成挑战 Enc(PubKey,$M^{(0|1)}$).模拟器 D 随机选择随机选择 $\beta\in\{0,1\}$,设自动状态机表示为 $M^{(\beta)}=(Q,\Sigma,\delta,q_0,F)$,其中 Q 个状态集合是 $Q=\{q_0,q_1,\cdots,q_{\omega-1}\}$(初始状态是 q_0),自动机 $M^{(\beta)}$ 的转换函数 δ 定义为$\{(x_i,y_i,\sigma_i)\in Q\times Q\times\Sigma\}\in\Gamma$ 的序列.D 随机选取 $\gamma_0,\gamma_1,\cdots,\gamma_{\omega-1}\leftarrow Z_N$,计算 $B_i=T^{\gamma_i}$,这里 T 是定理 3 中判定挑战($T\in G_{\pi_1\pi_2}$ 或 $T\in G_{\pi_1}$).

对自动机的转换函数 $t_i=(x_i,y_i,\sigma_i)$,D 采用递推方式执行:随机选取 $r_{t_i}\leftarrow Z_N$,计算 $C_{t_i,1}=T^{ar_{t_i}}/B_{x_i}$,$C_{t_i,2}=T^{r_{t_i}}$,$C_{t_i,3}=B_{y_i}T^{r_{t_i}b_{\sigma_i}}$.

然后 D 对 $M^{(\beta)}$ 的结束状态集 F 中每个元素 $B_x\in F$,计算 $H_x=e(g^{\alpha},T)^{\gamma_x}$.最后输出挑战密文 $\mathrm{Ctx}_{M^{(\beta)}}=(B_0,(C_{t_i,1},C_{t_i,2},C_{t_i,3})_{t_i\in\Gamma},\forall x\ B_x\in FH_x)$

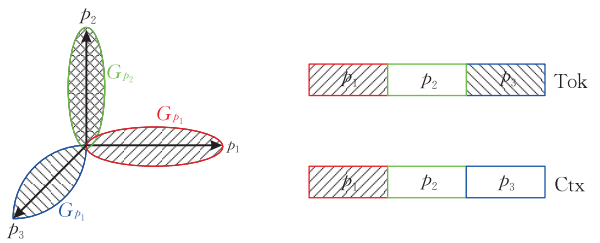
在输出阶段,敌手 A 以 β' 作为对挑战中的随机抛币 β 的猜测.D 以相同的输出 β' 作为对定理 3 中输出元素判定的猜测:$\beta'=0$ 则 $T\in G_{\pi_1}$,$\beta'=1$ 则 $T\in G_{\pi_1\pi_2}$.

上述输出中,若 $T\in G_{p_1}$,此时挑战密文 $\mathrm{Ctx}_{M^{(\beta)}}$ 中没有子群 G_{p_2} 的任何元素,因此是正常形态的密文.若 $T=g^u g_2^v\in G_{p_1p_2}$,这里 g_2 是子群 G_{p_2} 的生成元,$v\neq 0 \bmod p_2$(否则 $T\in G_{p_1}$),根据中国剩余定理,$\mathrm{Ctx}_{M^{(\beta)}}$ 中所有元素模 p_1 不为 0,即 $\mathrm{Ctx}_{M^{(\beta)}}$ 是半功能化密文.因此若敌手可以区分一个正常形态的密文和半功能化形态的密文,则我们可以以相同的优势攻击定理 3 所证明计算不可区分的子群判定难.因此任何多项式时间的敌手是无法区分正常形态密文和半功能形态密文的. 证毕.

接下来把询问令牌转换为半功能化形式.设敌手在安全性游戏中共进行了 Q 次令牌询问.对 $1\leq j\leq Q$,应答秘密令牌询问如下:对前 $j-1$ 个令牌用半功能形式的令牌 $\overrightarrow{\mathrm{Tok}_w}$ 作应答,对 $j+1$ 以后的令牌用正常形式令牌 Tok_w 回答,对第 j 个秘密令牌,用安全推论 4 的挑战作应答,即把 Tok_w 中所有的 g 用定理 4 中的挑战输出 T 来代替.$\mathrm{Tok}_w=(A=T^{s_0}U,(K_i,D_i)_{i=1}^{\ell},E=T^{\alpha+s_\ell}V)$,其中

$$\begin{pmatrix}K_i\\D_i\end{pmatrix}_{i\in[\ell]}=\begin{pmatrix}X\\Z\end{pmatrix}_{i\in[\ell]}\times\begin{pmatrix}T^{s_1} & T^{s_2} & \cdots & T^{s_\ell}\\(h_{w_1})^{s_1}z^{s_0} & (h_{w_2})^{s_2}z^{s_1} & \cdots & (h_{w_\ell})^{s_\ell}z^{s_{\ell-1}}\end{pmatrix} \tag{11}$$

显然,若 $T=Z_1Z_3$,则第 j 个令牌是正常形态的,若 $T=Z_1Z_2Z_3$ 则是半功能化的.采用定理 4 的证明,易证令牌由正常形态转换为半功能化形态时是不经意的,敌手无法在多项式时间的概率上获得这种转换的优势.图 2(a)描述本方案中 3 个子群的正交性,图 2(b)显示构造方案中密文和密钥所利用的子群.



(a) 子群正交性 (b)方案所在子群

图 2 半功能密文/令牌

在定理 4 中的挑战密文生成过程中,用 X_1X_2 代替原来加密算法中的 g,即 $B_i=(X_1X_2)^{\gamma_i}$,$C_{t_i,2}=(X_1X_2)^{r_{t_i}}$.显然挑战密文是半功能化的.若敌手可以通过已知询问的令牌来成功测试挑战密文,借助定理 4 的方法,我们可以利用敌手的输出来解决难题推论 2.因此对所有询问的 j 个令牌 $1\leq j\leq Q$,都无法成功测试挑战密文.

在令牌进行半功能变化过程中,当 $j=0$ 时,所有询问的令牌是正常形态的,当 $j=Q$ 时,所有的询问秘密令牌都是半功能的,此时挑战密文也是半功能的.对于 $j=1,\cdots,Q$ 的每一次令牌提取,根据推论 4,把一个正常令牌改变成半功能化形态时敌手在计算上是不可区分的.根据推论 4,即使敌手询问了 Q 个令牌,仍不能成功测试一个挑战密文.而挑战密文和询问令牌都已是半功能化,根据双系统加密性质,半功能化令牌无法获得半功能密文任何信息,因此敌手询问的令牌无法获得正常形态的挑战密文信息.方案得证.

6 应用实例

6.1 安全外包计算

现有的外包计算假定所要计算的目标或算法是不需要保密的,外包服务器通过其强大的计算能力为请求服务提供计算能力.如果需要计算的代码或算法具有隐私性,如生物模板匹配涉及用户的指纹、DNA 或其它私密信息,信用卡支付涉及用户的帐号及相关隐私信息等,此时外包计算面临很大的隐私泄露风险.

将设计的方案稍作修改,采用加密代码的方法

可以有效地应用于安全外包计算中. 实现如下：首先用户把外包计算程序转换为等价表达的有限自动机 M，然后对该自动机进行加密生成 Ctx_M 并提供给外包服务器. 此时由于程序代码是经过加密处理的，即使是外包服务器也无法从密文中获知有关外包程序 M 的信息. 然后用户将程序所需的输入 $\boldsymbol{w}$ 生成一个令牌 Tok_w 并提供给外包服务器. 外包服务器通过调用测试算法计算和输出运行结果. 系统运行如图 3 (b)所示. 事实上，Check 算法可以作适当改进，达到输出运行结果(自动机运行结束后停止在集合 F 中的状态)而非简单的布尔结果，在第 7.1 节中讨论这种扩展.

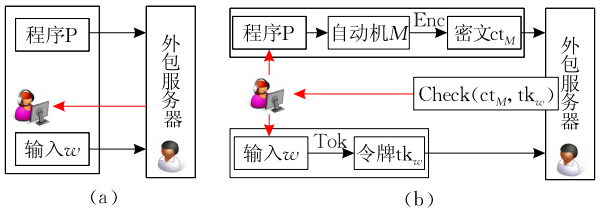


图 3　安全外包服务计算

6.2　防火墙内容过滤

现有的防火墙过滤技术引入有限自动机来进行内容过滤，但由于过滤规则是以明文公式规则形式保存在防火墙服务器中，容易被攻击者识别后绕开过滤规则从而达到穿透防火墙的目的. 采用本文提出的有限自动机策略加密技术，自动机过滤规则通过加密以密文的形式保存，即便攻击者获得这个密文形式的过滤规则，仍无法从中获得有关自动机的有用信息.

6.3　模板隐私保护的 DNA 比对

DNA 是个人敏感信息，需要采用一定的技术手段提供隐私保障. 在基于网络的 DNA 比对时，一方面我们不能直接把 DNA 串以明文的形式在网络上提供给 DNA 匹配模板，同时要求匹配模板也需要保密，这对开放网络平台的 DNA 比对提出挑战. 采用有限自动机加密技术，可以把 DNA 匹配模板加密生成密文并公开给用户在用户端进行比对，从而实现具有隐私的 DNA 比对.

7　扩展与讨论

7.1　带负载的确定有限状态机策略加密

本节我们将讨论如何将无负载消息的方案 $L=(\mathrm{Setup},\mathrm{GenTK},\mathrm{Enc},\mathrm{Check})$ 扩展到可以保护消息机密性的基于确定性有限状态机加密 $L^*=(\mathrm{Setup}^*,\mathrm{GenKey}^*,\mathrm{Enc}^*,\mathrm{Dec}^*)$. L^* 方案要求加密密文中隐藏负载消息 m，当 $\mathrm{Accept}(M,w)=1$ 时用秘密令牌 Tok_w 能解密密文 Ctx_M 并恢复消息 m. 扩展方案采用方案 L 作为核心模块并作适当扩展，具体如下：

(1) $L^*.\mathrm{Setup}^*(1^\lambda,\Sigma)$ 该算法与 $L.\mathrm{Setup}$ 基本相同，唯一区别在于在此算法中要增加定义消息明文空间 $m\in G_t$，并将其作为系统主公钥的一部分公开；

(2) $L^*.\mathrm{GenKey}^*(\mathrm{PubKey},\mathrm{MstKey},w)$ 该算法与 $L.\mathrm{GenTK}$ 相同；

(3) $L^*.\mathrm{Enc}^*(\mathrm{PubKey},M,m)$ 与 $L.\mathrm{Enc}$ 相比较，本算法输入参数增加消息明文 $m\in G_t$. 加密过程中，该算法调用 $L.\mathrm{Enc}$ 生成有限机密文 $\mathrm{Ctx}_M=(B_0,(C_{t_i,1},C_{t_i,2},C_{t_i,3})_{t_i\in\Gamma},\forall x\ B_x\in F,\forall x\ H_x=Y^{\gamma_x})$，然后把消息 m 附在 H_x 上，即置 $H_x=m\cdot Y^{\gamma_x}$，其它密文组件保持不变.

(4) $L^*.\mathrm{Dec}^*(\mathrm{PubKey},\mathrm{Ctx}_M,\mathrm{Tok}_w)$ 解密算法与 $L.\mathrm{Check}$ 基本相同，只是将 $L.\mathrm{Check}$ 中的测试等式(6)修改为

$$m\leftarrow H_x\frac{W_\ell}{e(E,B_x)} \tag{12}$$

事实上，算法 L 可以看作是 L^* 的一个特例：L 中的消息空间定义为 G_t 上的单位元，即 $m=1_{G_t}$.

7.2　进一步讨论

一般来讲，相同安全级别的双线性群构造中，合数阶群中双线性对的运算代价要比在素数阶群中大. 为提高系统效率，一是可以采用文献[28]中提高双线性对运算效率的方法，另外也可以采用 Freeman[29] 和 Lewko[30] 所提出的转换方法，将基于合数阶群的构建方案转换为素数阶的构造. 事实上，安全性证明中利用的是合数阶群的子群正交性来达到自适应安全. 这里我们给出在多个素数阶群中构造出正交子群的方法，利用该性质可将合数阶的方案转换为素数阶方案 L^{**}.

定义 7(可消去双线性映射).　设 G,H 是两个阶为素数 p 的有限乘法群，$\bar{e}$ 是双线性映射满足 $\bar{e}:G\times H\to G_t$. 对 G 的子群 G_1,G_2,G_3 和 H 的子群 H_1,H_2,H_3，即 $G=G_1\times G_2\times G_3$，$H=H_1\times H_2\times H_3$. 对任意 $g_i\in G_i,h_j\in H_j$，若下面性质满足，则称 $\bar{e}$ 是具有可消去的

$$\bar{e}(g_i,h_j)=\begin{cases}\neq 1, & i=j\ (\text{非退化性})\\ =1, & i\neq j\ (\text{正交性})\end{cases} \tag{13}$$

说明.　G 可以看成三个子群构成. 若 $G=H$，称该双线性映射为对称群的可消去映射. 显然，可消去双线性群具有合数阶双线性群的性质.

给出该可消去双线性群的具体构造：

(1)以安全参数 λ 作为输入，生成阶是素数 p 的双线性群 (p,G,H,G_t,e) 满足 $e:G\times H\rightarrow G_t$；

(2)定义 $G=\overline{G}^3$，$\overline{H}=H^3$ 以及 $\overline{G}_t=G_t$；

(3)随机选取生成元 $g_1,g_3,g_3\in G,h_1,h_2,h_3\in H$；

(4)随机选择 $a,b,c,x,y,z\in Z_p$，满足

$$\begin{vmatrix}1 & a & x\\ 1 & b & y\\ 1 & c & z\end{vmatrix}\neq 0,\quad \begin{vmatrix}cy-bz & z-y & b-c\\ cx-az & z-x & a-c\\ bx-ay & y-x & a-b\end{vmatrix}\neq 0 \tag{14}$$

(5)定义子群

$$\overline{G}_1=\langle(g_1,g_1^a,g_1^x)\rangle,\overline{G}_2=\langle(g_1,g_1^b,g_1^y)\rangle,$$

$$\overline{G}_3=\langle(g_1,g_1^c,g_1^z)\rangle,\overline{H}_1=\langle(h_1^{cy-bz},h_1^{z-y},h_1^{b-c})\rangle,$$

$$\overline{H}_2=\langle(h_1^{cx-az},h_1^{z-x},h_1^{a-c})\rangle,$$

$$\overline{H}_3=\langle(h_1^{bx-ay},h_1^{y-x},h_1^{a-b})\rangle;$$

(6) $\forall g_1\in G_1,g_2\in G_2,g_3\in G_3,h_1\in H_1,h_2\in H_2,h_3\in H_3$，定义运算

$$\bar{e}:\overline{G}\times\overline{H}\rightarrow\overline{G}_t$$

$$e((g_1,g_2,g_3),(h_1,h_2,h_3)):=$$

$$e(g_1,h_1)e(g_2,h_2)e(g_3,h_3)$$

(7)输出群描述

$$(p,G,\overline{G}_1,\overline{G}_2,\overline{G}_3,H,\overline{H}_1,\overline{H}_2,\overline{H}_3,\overline{G}_t,\bar{e}).$$

说明. ①式(14)前一行列式用于保证子群 $\overline{G}_i$ $(1\leq i\leq 3)$ 之间不线性相关，后一行列式保证 $\overline{H}_i$ 之间不线性相关.

②利用该群来构造(4.1 节)方案的方法：令 $G=H$，将 $\overline{G}_1$ 对应于合数阶群的子群 G_{p_1}，$\overline{G}_2$ 对应于子群 G_{p_2} 和 $\overline{G}_3$ 对应于子群 G_{p_3}，然后将方案中相应子群中的元素从可取消子群中选取，而方案的思路不变.

7.3 性 能

接下来我们给出本文方案的性能分析. 方案 L，L^* 和 L^{**} 分别对应本文 4.1 节，7.1 节和 7.2 节所提出的方案. 表 1 描述所提出方案的性能参数. L 和 L^* 在合数阶群中构造，L^{**} 在素数阶群中构造，素数阶中的双线性对的计算效率高于合数阶，因此计算性能上 L^{**} 要高于前二者. L^* 扩展 L 达到负载消息隐藏，L 可以看作是一种特殊的 L^* 方案，即其消息 $m\in 1_{G_t}$.

表 1 方案性能

方案	群阶	公钥	令牌	密文	测试/解密	负载消息
L	$p_1p_2p_3$	$(\lvert\Sigma\rvert+3)\lvert G\rvert+\lvert G_t\rvert$	$(2\ell+2)\lvert G\rvert$	$(\lvert\Gamma\rvert+\lvert F\rvert+1)\lvert G\rvert+\lvert\Gamma\rvert\lvert G_t\rvert$	$(3\ell+1)$Pr+1Mul	无
L^*	$p_1p_2p_3$	$(\lvert\Sigma\rvert+3)\lvert G\rvert+\lvert G_t\rvert$	$(2\ell+2)\lvert G\rvert$	$(\lvert\Gamma\rvert+\lvert F\rvert+1)\lvert G\rvert+\lvert\Gamma\rvert\lvert G_t\rvert$	$(3\ell+1)$Pr+2Mul	有
L^{**}	$p^\dagger$	$(6\lvert\Sigma\rvert+18)\lvert G\rvert+\lvert G_t\rvert$	$(6\ell+6)\lvert G\rvert$	$(3\lvert\Gamma\rvert+3\lvert F\rvert+3)\lvert G\rvert+\lvert\Gamma\rvert\lvert G_t\rvert$	$(3\ell+1)$Pr+1Mul	无

注：$\lvert\Sigma\rvert$：DFA 中字母表长度；ℓ：DFA 输入串 w 长度；$\lvert\Gamma\rvert$：DFA 转换函数数量；$\lvert F\rvert$：接受状态集合大小；Pr：一个双线性对运算时间；Mul：一个 G_t 乘法运算时间；$\lvert G\rvert$：一个群 G 元素长度；$\lvert G_t\rvert$：一个群 G_t 元素长度；†：素数阶中的对运算比合数阶的运算速度要快.

实际应用中，L 和 L^{**} 可用于隐私保护的数据库查询、隐私保护的防火墙过滤等应用场合，而含有负载消息的方案 L^* 可应用于关键字可搜索的机密文档传输、隐私保护的垃圾邮件过滤等场合.

8 小 结

本文提出一种以确定有限自动机 DFA 作为加密策略、以 DFA 可接受的串作为解密角色的加密技术，可以有效地应用于隐私保护的推导和计算的应用场合. 在给出系统模型和安全性定义的基础上，本文构建了两个具体的方案：无负载的 DFA 加密方案和有负载的 DFA 加密方案，并给出了转换为素数阶群实现的一般方法. 在双系统加密技术下，证明了所提出的方案达到自适应安全性. 并给出了本文方案在安全外包计算和隐私保护的防火墙内容过滤方面的具体应用. 接下来的工作主要体现在两方面：一是扩展有限自动机的表达能力，如达到非确定性有限自动机或图灵机的加密方案，可以有效地实现对可执行代码的保护，避免被反编译或逆向工程而导致程序及算法思路泄露；二是进一步研究提高系统效率的方法.

参 考 文 献

[1] Chu Dian-Hui, Meng Fan-Chao, Zhan De-Chen, Xu Xiao-Fei. Multi-level component behavior matching model based on finite automata. Journal of Software, 2011, 22(11): 2668-2683(in Chinese)
(初佃辉，孟凡超，战德臣，徐晓飞. 基于有限自动机的多层次构件行为匹配模型，软件学报，2011，22(11)：2668-2683)

[2] Hu Y, Gao Q, Guo L, Wang P. Giant complete automaton for uncertain multiple string matching and its high speed construction algorithm. Science China Information Sciences, 2011, 54(8): 1562-1571

[3] Boneh D, Sahai A, Waters B. Functional encryption: Definitions and challenges//Proceedings of the 8th Theory of Cryptography Conference (TCC'11). Rhode Island, USA, 2011: 253-273

[4] Boneh D, Sahai A, Waters B. Functional encryption: A new vision for public-key cryptography. Communications of the ACM, 2012, 55(11): 56-64

[5] Okamoto T, Takashima K. Fully secure functional encryption with general relations from the decisional linear assumption// Proceedings of the 30th International Cryptology Conference (CRYPTO'10). Santa Barbara, USA, 2010: 191-208

[6] Tao Ren-Ji, Chen Shi-Hua. A finite automation public cryptogsystem and digital signatures. Chinese Journal of Compters, 1985, 8(6): 401-409(in Chinese)
(陶仁骥，陈世华. 一种有限自动机公开钥密码体制和数字签名. 计算机学报，1985，8(6)：401-409)

[7] Waters B. Dual system encryption: Realizing fully secure ibe and hibe under simple assumptions//Proceedings of the 29th International Cryptology Conference (CRYPTO'09). Santa Barbara, USA, 2009: 619-636

[8] Fiat A, Naor M. Broadcast encryption//Proceedings of the 13th International Cryptology Conference (CRYPTO'93). Santa Barbara, USA, 1994: 56-59

[9] Goyal V, Pandey O, Sahai A, Waters B. Attribute-based encryption for fine-grained access control of encrypted data// Proceedings of the 13th ACM Conference on Computer and Communications Security(ACM-CCS'06). Alexandria, USA, 2006: 89-98

[10] Lewko A, Okamoto T, Sahai A, et al. Fully secure functional encryption: Attribute-based encryption and (hierarchical) inner product encryption//Proceedings of the 29th Annual Conference on Europe Cryptology (EUROCRYPT'10). Nice, French, 2010: 62-91

[11] Lee K, Lee D H. Improved hidden vector encryption with short ciphertexts and tokens. Designs, Codes and Cryptography, 2011, 58(3): 297-319

[12] Zhang M, Yang B, Takagi T. Bounded leakage-resilient functional encryption with hidden vector predicate. The Computer Journal, 2013, 56(4): 464-477

[13] Katz J, Sahai A, Waters B. Predicate encryption supporting disjunction, polynomial equations, and inner products// Proceedings of the 27th Annual Conference on Europe Cryptology (EUROCRYPT'08). Istanbul, Turkey, 2008: 146-162

[14] Agrawal S, Freeman D M, Vaikuntanathan V. Functional encryption for inner product predicates from learning with errors//Proceedings of the 17th International Conference on the Theory and Application of Cryptology and Information Security (ASIACRYPT'11). Seoul, South Korea, 2011: 21-40

[15] Hamburg M. Spatial Encryption [Ph. D. dissertation]. Stanford University, Stanford, California, USA, 2011

[16] Zhang Ming-Wu, Yang Bo, Takagi T. Master-key leakage-resilient and continue leakage-resilient functional encryption in dual affine space. Chinese Journal of Computers, 2012, 35(9): 1856-1867(in Chinese)
(张明武，杨波，Takagi Tsuyoshi. 抗主密钥泄露和连续泄露的双态仿射函数加密. 计算机学报，2012，35(9)：1856-1867)

[17] Weng J, Chen M R, Yang Y, et al. CCA-secure unidirectional proxy re-encryption in the adaptive corruption model without random oracles. Science China Information Sciences, 2010, 53(3): 593-606

[18] Waters B. Functional encryption for regular languages// Proceedings of the 32nd International Cryptology Conference (CRYPTO'12). Santa Barbara, USA, 2012: 218-235

[19] Boneh D, Waters B. Conjunctive, subset, and range queries on encrypted data//Proceedings of the 4th Theory of Cryptography Conference (TCC'07). Amsterdam, Netherlands, 2007: 535-554

[20] Yang X, Wang B, Yu G. Efficient secure data publishing algorithms for supporting information sharing. Science in China Series F: Information Sciences, 2009, 52(4): 627-644

[21] Xu Ming-Di, Zhang Huan-Guo, Yan Fei. Testing on trust chain of trusted computing platform based on labeled transition system. Chinese Journal of Computers, 2009, 32(4): 635-645(in Chinese)
(徐明迪，张焕国，严飞. 基于标记变迁系统的可信计算平台信任链测试. 计算机学报，2009，32(4)：635-645)

[22] Lindell Y, Pinkas B. Secure multi-party computation for privacy-preserving data mining. The Journal of Privacy and Confidentiality, 2009, 1(1): 59-98

[23] Carbunar B, Tripunitara M V. Payments for outsourced computations. IEEE Transactions on Parallel and Distributed Systems, 2012, 23(2): 313-320

[24] Yang D, Shen D R, Yu G, et al. Query intent disambiguation of keyword-based semantic entity search in data spaces. Journal of Computer Science and Technology, 2013, 28(2): 382-393

[25] Liang K, Au M H, Liu Joseph K, et al. A DFA-based functional proxy re-encryption scheme for secure public cloud data sharing. IEEE Transactions on Information Forensics and Security, 2014, 9(10): 1667-1680

[26] Zhang M. ACP-lrFEM: Functional encryption mechanism with automatic control policy in the presence of key leakage// Proceedings of the 10th International Conference on Information Security Practice and Experience (ISPEC'14). Fuzhou, China, 2014: 481-495

[27] Lewko A, Waters B. New techniques for dual system encryption and fully secure hibe with short ciphertexts// Proceedings of the 7th Theory of Cryptography Conference (TCC'10). Zurich, Switzerland, 2010: 455-479

[28] Zhang Y, Xue C J, Wong D S, et al. Acceleration of composite order bilinear pairing on graphics hardware//Proceedings of the 14th International Conference on Information and Communications Security (ICICS'12). Hong Kong, China, 2012: 341-348

[29] Freeman D M. Converting pairing-based cryptosystems from composite-order groups to prime-order groups//Proceedings of the 29th Annual Conference on Europe Cryptology (EUROCRYPT'10). Nice, French, 2010: 44-61

[30] Lewko A B. Tools for simulating features of composite order bilinear groups in the prime order setting//Proceedings of the 31st Annual Conference on Europe Cryptology (EUROCRYPT'12). Cambridge, United Kingdom, 2012: 318-335

ZHANG Ming-Wu, Ph.D., professor. His research interests include information and network security, and privacy preservation.

YANG Bo, born in 1963, Ph.D., professor. His research interests include cryptography and information security.

WANG Chun-Zhi, born in 1963, Ph.D., professor. Her main research interests include network protocols and security.

TAKAGI Tsuyoshi, Ph.D., professor. His main research interest is cryptography.

Background

In decentralized network environments, data provider should make out a flexible and extensible control policy on the encrypted data to allow the client to access the data in a secure and efficient manner. Functional encryption, as a new primitive of cryptographic system, supports a flexible and effect approach to describe a fine-grained control over the ciphertext. The encryption policies are defined as functions and the decryption roles are specified as vectors. A decryptor can extract the plaintext if the role key satisfies the corresponding policy function. Obviously, functional encryption can be considered as a general expansion of public-key encryption, identity-based encryption, attribute-based encryption, predicate encryption and their hierarchical constructions.

In this work, we study the functional encryption that employs the deterministic finite automata as the encryption policy, which is flexible in the practical application such as biometric template comparison, firewall filtering and outsourcing program etc. We propose two adaptively/fully secure deterministic finite automata (DFA-based) encryptions in the standard model. In the first scheme, the ciphertext is associated with a DFA and the token is associated with an arbitrary length string over the alphabet, and there is a check algorithm to test whether the string is accepted by the automata in the key/ciphertext spaces. We also extend the first scheme to support payload confidentiality, in which the decryption will extract the encrypted message if the automata of the ciphertext accepts the string associated with the key. We give the security proof and provide the performance analysis.

We provide the application scenarios of deploying our scheme as the building blocks to implement the secure program outsourcing, privacy-preserving DNA matching, and secure firewall filtering etc.

This study is supported by the National Natural Science Foundation of China under Grant Nos. 61370224 and 61272436, and the Key Program of Natural Science Foundation of Hubei Province under Grant No. 2013CFA046, the High-Level Talent Plan of Hubei University of Technology, and the Open Fund Program for State Key Laboratory of Information Security.