中国科学 : 信息科学 2025 年 第 55 卷 第 8 期: 2002–2017
SCIENTIA SINICA Informationis
论文

前向安全的 SM9 密文等值测试

张健龙[1], 赖建昌[1*], 黄欣沂[2], 陈立全[1], 韩金广[1]

1. 东南大学网络空间安全学院, 南京 211189
2. 暨南大学网络空间安全学院, 广州 510632
* 通信作者. E-mail: jclai@seu.edu.cn

收稿日期: 2025–01–04; 修回日期: 2025–03–30; 接受日期: 2025–05–07; 网络出版日期: 2025–08–11

国家自然科学基金 (批准号: 62032005, U21A20466) 和东南大学科研启动经费 (批准号: RF1028623200) 资助项目

摘要 密文等值测试能有效解决密态数据无法进行底层明文比较的问题, 在数据库分区、加密数据统计、去重存储等方面发挥着重大作用. SM9 标识密码现已成为我国商用密码行业标准和国家标准, 广泛应用在金融、政务等领域. 然而现有基于 SM9 的密文等值测试方案无法实现前向安全性, 即前期生成的陷门可用于测试任意时间点生成的密文, 不满足特殊应用对前向安全性的需求, 阻碍了 SM9 密码算法的推广. 前向安全性指只有在指定时间之前生成的密文才能进行有效的等值测试, 如果密文中的时间大于等于指定时间, 则无法测试成功. 本文基于 SM9 标识加密算法提出了具有前向安全性的密文等值测试方案 (SM9-FSET), 只有满足时间条件的密文才能进行等值测试, 进一步保障了数据隐私. 方案的安全性基于 q-BDHI 困难问题, 在随机谕言机模型中可证明方案满足 IND-ID-CCA 和 OW-ID-CCA 的安全性. 最后, 分析比较新方案与现有相关方案, 结果表明 SM9-FSET 在计算开销、通信代价和安全性方面与相关方案是可比的且更具实用性.

关键词 SM9, 密文等值测试, 前向安全, 选择密文攻击 (CCA)

1 引言

在传统公钥密码系统中, 用户公钥为 “无意义” 的随机字符串, 需借助证书进行认证, 确保用户公钥的有效性和正确性, 但证书的管理是一项繁琐的工作. 为解决公钥需认证的问题, 1984 年, Shamir [1] 首次提出了标识密码的概念. 它采用具有唯一标识身份作用的任意字符串作为用户公钥, 如用户的手机号、邮箱等, 解决了证书管理问题. 之后 Boneh 与 Franklin [2] 定义了标识加密的安全模型, 并基于双线性配对技术设计了首个实用且可证明安全的标识加密方案. 之后关于标识加密方案的研究受到了学术界的广泛关注, 一系列工作 [3,4] 从安全性与性能等方面对标识加密进行了提升.

云存储迅速发展的今天, 其巨大的存储容量成为用户解决终端存储容量有限问题的有效途径. 为保护云端数据安全, 数据通常以加密形式存储在云端, 但这会给数据的使用带来不便. 例如, 当用户检

引用格式: 张健龙, 赖建昌, 黄欣沂, 等. 前向安全的 SM9 密文等值测试. 中国科学: 信息科学, 2025, 55: 2002–2017, doi: 10.1360/SSI-2025-0003
Zhang J L, Lai J C, Huang X Y, et al. Forward secure equality test based on SM9. Sci Sin Inform, 2025, 55: 2002–2017, doi: 10.1360/SSI-2025-0003

索数据时, 需将全部密文下载到本地, 在本地解密之后再执行检索操作, 不仅效率低下, 且对用户造成了极大不便. 在此环境下, 可搜索加密 (searchable encryption, SE) [5] 技术应运而生. SE 技术允许云服务器在不知道密文底层明文的情况下对其进行安全检索, 并将满足检索条件的密文数据返回给用户. 最后由用户在本地将检索结果解密, 从而获得自己需要的明文数据.

与其功能类似的有密文等值测试 (equality test, ET) [6], 主要包括基于公钥的密文等值测试 (public key encryption with equality test, PKEET) 和基于标识的密文等值测试 (identity-based encryption with equality test, IBEET). ET 技术实现了在未知明文的情况下, 判断不同公钥加密的密文是否对应同一明文的功能, 通常用于数据统计、分区领域. 比如邮件系统存储了海量数据, 用户希望按优先级分区以提高检索效率. 同等优先级的邮件在同一个区, ET 技术能够有效解决密态邮件分区的问题. 在医疗系统中, 医院通常将大量病例加密存储于云端以保护病人的隐私. 当医生需要对病例进行统计分析时, 可利用 ET 技术, 委托云对密态数据进行等值测试, 并将结果告知于医生. 然而, 在某些特殊场景下, 更关注某一时间点之前数据的分析, 比如医生需要分析某个流行病随时间变化的信息以对未来此疾病的流行度进行预测. 不幸的是传统的密文等值测试方案无法很好地满足这一需求. 且在传统的密文等值测试方案中, 云能够对用户任意时间生成的密文进行等值测试. 由于云是不完全可信的, 因此可能威胁用户的隐私安全. Ma 等 [7] 利用 0/1 编码技术 [8] 给出了前向安全的标识密文等值测试概念, 然而其方案构造并未实现密文等值测试的前向安全性.

标识密码因解决了公钥认证问题而成为学术界的研究热点. 然而现有研究大多围绕国外设计的算法展开. 为实现密码算法的自主可控, 保障国家网络与信息安全, 我国自主研发了一个商用标识密码算法标准 [9], 涵盖数字签名、密钥交换和密钥封装等内容. SM9 标识密码于 2016 年正式发布成为行业标准, 2020 年确立为国家标准, 2021 年被纳入 ISO/IEC 国际标准, 其影响力在国内外逐步提升. 近年来, 国内学者们设计了一系列基于 SM9 标识加密算法的扩展方案, 丰富了算法的功能, 包括支持密文检索 [10] 和密文等值测试 [11] 等功能. 但目前尚未在密码学国际主流期刊和学术会议上发现对前向安全的 SM9 密文等值测试算法的研究.

因此, 本文基于商用密码 SM9 标识加密算法的结构特征, 提出一个实现前向安全的密文等值测试方案 (forward secure SM9 with equality test, SM9-FSET). 在 SM9-FSET 中, 测试者只能对指定时间之前的密文进行等值测试. 这一设计不仅提高了密文等值测试的灵活性, 同时限制了测试者的能力, 进一步保障了数据隐私. 具体而言, 加密者在生成的密文中嵌入指定时间, 并将嵌入时间的密文发给测试者. 解密者利用私钥生成包含指定时间的陷门, 并发给测试者. 测试者获取对应的陷门和密文后, 当且仅当密文中嵌入的时间小于陷门中嵌入的时间时, 密文才可有效进行等值测试, 反之密文无法进行等值测试. 方案的安全性基于非对称群上计算性 q-BDHI 问题的困难性假设. 在随机谕言机模型中证明提出的 SM9-FSET 方案满足选择密文攻击下的不可区分 (indistinguishability identity against chosen ciphertext attacks, IND-ID-CCA) 和单向 (one-wayness identity against chosen ciphertext attacks, OW-ID-CCA) 安全性, 并对比分析了 SM9-FSET 与相关文献的性能. 结果表明 SM9-FSET 方案在计算效率、通信效率和安全性方面与现有相关方案 [7,11] 是可比的.

2　相关工作

由于公钥加密中的公钥通常是 “无意义” 的随机字符串, 因此需要权威的证书授权中心 (certificate authority, CA) 将用户身份与公钥进行绑定. 然而, 随着证书增多, CA 面临的负载问题已成为制约公钥密码学发展的一个重要因素. 为解决此问题, 1984 年 Shamir [1] 提出了标识密码的概念. 在该概念下, 用户的公钥不再是 “无意义” 的随机字符串, 而是任意一个具有唯一标识用户身份的字符串. 随后, 2001 年 Boneh 和 Franklin [2] 使用双线性配对技术给出了首个实用且可证明安全的标识加密方案, 并给出了在随机谕言机模型下的安全性分析. 为提高方案的安全性, Boneh 和 Boyen [12] 给出了在标准模

型下可证明安全的标识加密方案, 但方案只满足静态选择明文攻击的安全性. 为进一步提高标识密码的安全性, Gentry [13] 在 2006 年给出了在标准模型下可证明选择密文攻击安全的标识加密方案. 之后, 许多具有提高安全性和效率的标识加密方案相继被提出 [14,15].

为实现不同公钥加密的密文间明文对比, Yang 等 [6] 首次提出了基于公钥的密文等值测试概念, 并给出了可证明 OW-CCA 安全的具体构造. 密文等值测试实现了在不解密时对比两密文对应明文是否相同的功能. 由于 Yang 等所提方案在外包数据库管理场景中安全性不足等问题, Lu 等 [16] 对 Yang 等的 OW-CCA 安全模型进行分析, 给出了强于 OW-CCA 安全模型的新定义, 推动了追求更高安全级别的密文等值测试算法的研究. 传统密文等值测试方案中测试者无需用户同意即可对其任意密文进行等值测试, 一定程度上侵犯了用户隐私. 为限制测试者的测试能力, 2011 年 Tang [17] 首次利用授权令牌机制实现对测试者密文等值测试能力的细粒度控制. 为进一步优化隐私控制效率, Ma 等 [18] 研究了 PKEET 的授权机制, 提出了 4 种授权策略以提高用户隐私安全, 并给出了一个同时支持这 4 种授权机制的高效密文等值测试方案. 在随后的 2016 年, Ma [19] 首次给出了基于标识密文等值测试概念, 并给出了一个可证明安全的具体构造. 近年来针对等值测试的研究主要关注安全性的提高 [20] 和具体场景的应用 [21].

前向安全概念由 Günther [22] 和 Diffie 等 [23] 首次在密钥交换协议中提出, 前向安全密钥交换协议保证了长期密钥的暴露不会损害先前生成的会话密钥的安全. Katz 等 [24] 使用双线性配对技术首次给出了可证明安全的前向安全公钥加密方案. 为预防量子时代的威胁, Zhang 等 [25] 利用格基委托机制 [26] 给出了首个前向安全的格基可搜索加密方案 (lattice-based forward secure public-key encryption with keyword search, FS-PEKS), 并将其应用到物联网中. Yang 等 [27] 给出了前向安全的格基标识可搜索加密方案, 并证明了方案在选择明文攻击下的安全性, 解决了 FS-PEKS 中公钥认证的问题. 密文等值测试实现了密态数据对比的功能, 但传统的密文等值测试方案无法阻止测试者对任意时间的密文进行等值测试. 在一定程度上威胁着数据安全和用户隐私.

为限制测试者的能力, 实现测试者只能对指定时间之前的密文进行等值测试, Ma 等 [7] 首次提出了前向安全的标识密文等值测试的概念, 并利用 0/1 编码技术 [8] 给出了具体方案构造. 该方案利用时间编码技术在密文和陷门中嵌入时间, 意图阻止测试者对指定时间后的密文进行等值测试. 然而文献 [7] 中所提方案并未实现前向安全性, 即好奇的测试者可用临时陷门对用户所有密文进行等值测试. 攻击者即使针对未来时间 $\eta \geqslant \xi$ 的密文, 仍可从陷门中随机选取 $\delta \in \mathrm{Code}_{\xi}^{1}$, 并计算 $A_i = \frac{e(g^{\alpha}, h_{\mathrm{ID}}^{s_1} g^{H_4(\delta) r_y})}{e(g^{r_y}, (g^{\alpha})^{H_4(\delta)})} = e(g^{s_1}, h_{\mathrm{ID}})^{\alpha}$, 该计算消去了时间相关的随机数 r_y 和哈希项 $H_4(\delta)$, 直接暴露 $H(m) = [C_{i,2} \oplus H_2(A_i)]_1^{l_2}$. 攻击者无需验证 $\eta < \xi$ 的时间约束, 即可通过比较 $H(m)$ 判断明文等值性, 导致前向安全性失效. 同时该方案的 IND-ID-CCA 安全性证明未考虑攻击者利用临时陷门对非授权时间密文发起等值测试. 由于前向安全性失效, 攻击者可通过强制恢复明文哈希值直接比较不同时间段的密文内容, 从而区分挑战密文并攻破方案, 因此方案不满足 IND-ID-CCA 安全性.

SM9 标识密码作为我国自主研发的商用密码, 提出后得到了国内学者的积极研究. Yang 等 [28] 对 SM9 标识加密算法进行改进, 并首次将其应用于区块链的隐私保护领域. 由于 R-ate 是应用于 SM9 中较为重要的一种双线性映射, 为提高 R-ate 的计算效率, Gan 等 [29] 提出了一种快速计算的算法. Lai 等 [30] 给出了 SM9 数字签名算法在随机谕言机模型下的安全性证明, 并采用 Twin-Hash-ElGamal 技术, 基于 SM9 密钥封装机制提出了可证明安全的新型密钥封装机制. 并在文献 [31] 中提出了首个基于 SM9 的高效签密方案. 之后又对基于 SM9 的广播加密方案 [32] 进行了研究. Zhang 等 [33] 给出了首个基于 SM9 的可搜索加密方案, 解决经 SM9 密码算法加密后数据的检索问题. 为解决 SM9 标识加密算法不具备分层加密的功能, 不适合大规模网络的应用场景的问题, Lai 等 [34] 提出一个基于 SM9 的高效分层标识加密方案, 并在随机谕言机模型下给出了方案选择明文攻击下的不可区分性安全性证明. 随后, Li 等 [35] 设计了一个基于 SM9 的分层标识广播内积函数加密方案, 同时实现内积函数、密文广播和密钥授权, 且解密开销仅包含两个双线性配对运算. Xiong 等 [11] 提出了一种支持等值测试且

具备密码逆向防火墙 (cryptographic reverse firewalls, CRF) 机制的 SM9 标识加密方案. 目前尚未在国际主流的密码学期刊和会议上发现对前向安全的 SM9 密文等值测试的研究.

3　算法和安全模型定义

一个前向安全的标识密文等值测试方案通过以下 6 个多项式时间算法描述.

• $(\mathrm{mpk}, \mathrm{msk}) \leftarrow \mathbf{Setup}(\lambda)$. 输入安全参数 λ, 由密钥生成中心 (private key generator, PKG) 运行此算法并输出系统主公钥 mpk 和主私钥 msk, 其中系统主公钥 mpk 是系统的公共参数, msk 由 PKG 秘密保存.

• $\mathrm{sk}_{\mathrm{ID}} \leftarrow \mathbf{KeyGen}(\mathrm{mpk}, \mathrm{msk}, \mathrm{ID})$. 输入系统主公钥 mpk、主私钥 msk 和一个标识 ID, 由 PKG 运行此算法并输出与 ID 对应的私钥 $\mathrm{sk}_{\mathrm{ID}}$.

• $\mathrm{TD} \leftarrow \mathbf{TdGen}(t, \mathrm{sk}_{\mathrm{ID}}, \mathrm{mpk})$. 输入指定时间 t、标识 ID 对应的用户私钥 $\mathrm{sk}_{\mathrm{ID}}$、主公钥 mpk, 由用户执行此算法并输出与时间 t 绑定的陷门 TD.

• $\mathrm{CT} \leftarrow \mathbf{Encrypt}(t', m, \mathrm{ID}, \mathrm{mpk})$. 输入指定时间 t'、明文 m、接收者标识 ID、系统主公钥 mpk, 由加密者执行此算法并输出与时间 t' 绑定的密文 CT.

• $m/\perp \leftarrow \mathbf{Decrypt}(\mathrm{CT}, \mathrm{ID}, \mathrm{sk}_{\mathrm{ID}}, \mathrm{mpk})$. 输入私钥 $\mathrm{sk}_{\mathrm{ID}}$、主公钥 mpk、密文 CT、解密者标识 ID, 由解密者执行此算法并输出 m 表示解密成功, 或者 $\perp$ 表示解密失败.

• $0/1 \leftarrow \mathbf{Test}(\mathrm{CT}_1, \mathrm{CT}_2, \mathrm{ID}_1, \mathrm{ID}_2, \mathrm{TD}_1, \mathrm{TD}_2, \mathrm{mpk})$. 输入主公钥 mpk、接收者标识 $\mathrm{ID}_1, \mathrm{ID}_2$、对应的密文 $\mathrm{CT}_1, \mathrm{CT}_2$ 和陷门 $\mathrm{TD}_1, \mathrm{TD}_2$, 由测试者执行此算法并输出 0/1, 若两陷门中嵌入的时间均大于对应密文中嵌入的时间, 且两密文底层明文相同则输出 1; 否则输出 0.

解密正确性. 对于任意整数 λ、任意用户标识 ID 和消息 m, 有 $\Pr[\mathrm{Decrypt}(\mathrm{CT}, \mathrm{ID}, \mathrm{sk}_{\mathrm{ID}}, \mathrm{mpk}) = m] = 1$, 其中, $(\mathrm{mpk}, \mathrm{msk}) \leftarrow \mathrm{Setup}(\lambda)$, $\mathrm{sk}_{\mathrm{ID}} \leftarrow \mathrm{KeyGen}(\mathrm{mpk}, \mathrm{msk}, \mathrm{ID})$, $\mathrm{CT} \leftarrow \mathrm{Encrypt}(t', m, \mathrm{ID}, \mathrm{mpk})$. 表示在给定主公钥 mpk、用户标识 ID、对应私钥 $\mathrm{sk}_{\mathrm{ID}}$ 和正确密文 CT 的情况下, 解密算法总能正确恢复原始消息 m.

等值测试正确性. 对于任意整数 λ、任意两个用户标识 $\mathrm{ID}_1, \mathrm{ID}_2$ 以及对应用户的密文 $\mathrm{CT}_1, \mathrm{CT}_2$ 和陷门 $\mathrm{TD}_1, \mathrm{TD}_2$, 有如下性质.

• **一致性.** 当 $\mathrm{CT}_1, \mathrm{CT}_2$ 底层明文相同, 且密文中嵌入的时间小于对应陷门中嵌入的时间时, 等值测试算法以概率 1 返回 "1", 即

$$\Pr[\mathrm{Test}(\mathrm{CT}_1, \mathrm{CT}_2, \mathrm{ID}_1, \mathrm{ID}_2, \mathrm{TD}_1, \mathrm{TD}_2, \mathrm{mpk}) = 1] = 1.$$

• **完备性.** 当 $\mathrm{CT}_1, \mathrm{CT}_2$ 底层明文不同, 或任意一个密文中嵌入的时间不小于对应陷门中嵌入的时间时, 等值测试算法返回 "1" 的概率可忽略不计, 即 ($\mathrm{negl}(\lambda)$ 为关于 λ 的可忽略函数)

$$\Pr[\mathrm{Test}(\mathrm{CT}_1, \mathrm{CT}_2, \mathrm{ID}_1, \mathrm{ID}_2, \mathrm{TD}_1, \mathrm{TD}_2, \mathrm{mpk}) = 1] \leqslant \mathrm{negl}(\lambda),$$

其中, $(\mathrm{mpk}, \mathrm{msk}) \leftarrow \mathrm{Setup}(\lambda), \mathrm{TD}_i \leftarrow \mathrm{TdGen}(t_{\mathrm{ID}_i}, \mathrm{sk}_{\mathrm{ID}_i}, \mathrm{mpk}), \mathrm{CT}_i \leftarrow \mathrm{Encrypt}(t'_{\mathrm{ID}_i}, m, \mathrm{ID}_i, \mathrm{mpk})$. 表示在给定主公钥 mpk 和任意两个用户标识以及对应用户正确的陷门和密文的情况下, 测试算法总是能够正确地进行两密文底层明文的比较.

安全模型. 考虑两种类型的概率多项式时间敌手, 分别对应外部攻击者和内部攻击者 (测试者).

外部攻击者 (Type-1 敌手). 通过公开参数、部分密文以及陷门尝试获取密文底层的明文信息. 这类敌手可能利用已知的密文和陷门进行逆向分析, 试图破解密文内容.

内部攻击者 (Type-2 敌手). 通过获取陷门对密文进行等值测试, 试图从中推断出密文底层的明文信息. 这类敌手可能滥用权限, 利用陷门对密文进行非法操作.

针对挑战标识, (1) Type-1 敌手: 敌手没有嵌入时间大于挑战密文嵌入时间的陷门; (2) Type-2 敌手: 敌手拥有嵌入时间大于挑战密文嵌入时间的陷门.

本文给出前向安全标识密文等值测试算法的两个安全模型: 针对 Type-1 敌手的 IND-ID-CCA 和针对 Type-2 敌手的 OW-ID-CCA. 这两个安全模型通过挑战者 $\mathcal{C}$ 和攻击者 (敌手) $\mathcal{A}$ 之间的游戏定义. 首先介绍 IND-ID-CCA 安全模型定义.

• **系统建立.** 已知安全参数 λ, $\mathcal{C}$ 运行系统建立算法 **Setup**, 生成系统主公私钥对 $(\mathrm{mpk}, \mathrm{msk})$, 并将 mpk 发给 $\mathcal{A}$.

• **询问阶段 1.**

(1) 私钥询问. $\mathcal{A}$ 允许自适应地询问标识 ID_i 的解密私钥. $\mathcal{C}$ 运行私钥生成算法 **KeyGen** 生成私钥 $\mathrm{sk}_{\mathrm{ID}_i}$ 发给 $\mathcal{A}$.

(2) 陷门询问. $\mathcal{A}$ 选择时间 t_i, 以及标识 ID_i 发给 $\mathcal{C}$. $\mathcal{C}$ 首先生成指定 ID_i 的私钥, 然后运行陷门生成算法 **TdGen** 生成陷门 TD_i 发给 $\mathcal{A}$.

(3) 解密询问. $\mathcal{A}$ 允许自适应地选择 $(\mathrm{CT}_i, \mathrm{ID}_i)$ 对 $\mathcal{C}$ 发起解密询问. $\mathcal{C}$ 运行私钥生成算法 **KeyGen** 生成私钥 $\mathrm{sk}_{\mathrm{ID}_i}$ 进行解密, 并将解密结果发给 $\mathcal{A}$.

• **挑战.** 询问阶段 1 结束后, $\mathcal{A}$ 选择挑战标识 ID^*、挑战时间 t^* 和两个挑战明文 m_0, m_1 发给 $\mathcal{C}$. 模型要求 $\mathcal{A}$ 未询问关于 ID^* 的嵌入时间大于 t^* 的陷门. $\mathcal{C}$ 随机选择 $c \in \{0,1\}$, 令 $m^* = m_c$, 运行加密算法 $\mathbf{Encrypt}(t^*, m^*, \mathrm{ID}^*, \mathrm{mpk})$ 生成挑战密文 CT^*, 并将其发给 $\mathcal{A}$.

• **询问阶段 2.** $\mathcal{A}$ 继续向 $\mathcal{C}$ 发起私钥询问、陷门询问和解密询问. 模型要求 $\mathcal{A}$ 不能询问 ID^* 的私钥和挑战密文 $(\mathrm{CT}^*, \mathrm{ID}^*)$ 的解密, 同时针对 ID^* 的陷门询问嵌入的时间不能大于 t^*. $\mathcal{C}$ 根据询问阶段 1 回复 $\mathcal{A}$.

• **猜测.** $\mathcal{A}$ 输出对 c 的猜测 $c' \in \{0,1\}$. 如果 $c' = c$, 则 $\mathcal{A}$ 获胜. 定义 $\mathcal{A}$ 获胜的优势为

$$\mathrm{Adv}_{\mathcal{A}}^{\text{IND-ID-CCA}}(\lambda) = \left| \Pr[c' = c] - \frac{1}{2} \right|.$$

定义1　若不存在多项式时间 $\mathcal{A}$ 能以不可忽略的优势 $\mathrm{Adv}_{\mathcal{A}}^{\text{IND-ID-CCA}}(\lambda)$ 赢得上述游戏, 则称方案是 IND-ID-CCA 安全的.

其次, 给出 OW-ID-CCA 安全模型定义.

• **系统建立**和**询问阶段 1**与 IND-ID-CCA 安全模型定义相同.

• **挑战.** 询问阶段 1 结束后, $\mathcal{A}$ 选择挑战标识 ID^* 和挑战时间 t^* 并发给 $\mathcal{C}$. $\mathcal{C}$ 自适应地选择明文 m^*, 并使用挑战标识和时间加密明文得到密文 CT^*, 最后将挑战密文发给 $\mathcal{A}$.

• **询问阶段 2.** $\mathcal{A}$ 继续向 $\mathcal{C}$ 发起私钥询问、陷门询问和解密询问. 模型要求 $\mathcal{A}$ 不能询问 ID^* 的私钥和挑战密文 $(\mathrm{CT}^*, \mathrm{ID}^*)$ 的解密. $\mathcal{C}$ 根据询问阶段 1 回复 $\mathcal{A}$.

• **猜测.** $\mathcal{A}$ 输出猜测 m'. 如果 $m' = m^*$, 则 $\mathcal{A}$ 获胜. 定义 $\mathcal{A}$ 获胜的优势为

$$\mathrm{Adv}_{\mathcal{A}}^{\text{OW-ID-CCA}}(\lambda) = \left| \Pr[m' = m^*] - \frac{1}{2} \right|.$$

定义2　若存在多项式时间 $\mathcal{A}$ 能以不可忽略的优势 $\mathrm{Adv}_{\mathcal{A}}^{\text{OW-ID-CCA}}(\lambda)$ 赢得上述游戏, 则称方案是 OW-ID-CCA 安全的.

4　方案构造

本文所设计方案基于非对称双线性对, 首先给出非对称双线性对的定义. 设 λ 为安全参数, p 为大素数且 $p > 2^{\lambda}$. $\mathbb{G}_1$, $\mathbb{G}_2$ 和 $\mathbb{G}_T$ 都是通过 λ 生成的循环群, 且阶为 p. 映射 $e : \mathbb{G}_1 \times \mathbb{G}_2 \to \mathbb{G}_T$ 称为双线性映射, 满足对任意群元素 $P \in \mathbb{G}_1$, $Q \in \mathbb{G}_2$ 和整数 $a, b \in \mathbb{Z}_p$, $e(P, Q)$ 都能高效的计算, 且等式

表 1　10 和 5 进行 0/1 编码的结果.

Table 1　Result of 0/1-Encoding for 10 and 5.

	0-Encoding	1-Encoding
$x=(1010)_2$	1011，11	101，1
$y=(0101)_2$	011，1	0101，01

$e(aP,bQ)=e(P,Q)^{ab}$ 成立. 此外, 至少存在群元素 $P\in\mathbb{G}_1,Q\in\mathbb{G}_2$, 满足 $e(P,Q)\neq 1$. 记双线性群为 $\mathcal{BP}=(\mathbb{G}_1,\mathbb{G}_2,\mathbb{G}_T,e,p)$.

4.1　0/1 编码技术

0/1 编码技术是由 Lin 等 [8] 提出用来解决百万富翁问题的方法. 用 $k=k_nk_{n-1}\cdots k_1=\{0,1\}^n$ 表示一个 n 位二进制字符串, 其中 n 足够大以涵盖所有可能的值, 在本文所提方案中只考虑整数值. 在 0 编码算法中, k 被编码为一个集合 S_k^0, 该集合最多包含 $\log_2 n$ 个元素: $S_k^0=\{k_nk_{n-1}\cdots k_{i+1}1|k_i=0,i\in[1,n]\}$. 在 1 编码算法中, k 被编码为一个集合 S_k^1, 该集合最多包含 $\log_2 n$ 个元素: $S_k^1=\{k_nk_{n-1}\cdots k_i|k_i=1,i\in[1,n]\}$.

当使用 0/1 编码技术比较整数 x 和 y 的大小时, 可以使用 1 编码算法将整数 x 编码为集合 S_x^1, 使用 0 编码算法将整数 y 编码为集合 S_y^0. 如果 $S_x^1\cap S_y^0$ 不是空集, 那么可以得出 $x>y$, 否则, 可以知道 $x\leqslant y$, 即 $S_x^1\cap S_y^0\neq\varnothing\Longleftrightarrow x>y$.

为详细描写该编码技术, 以比较两个 4 位二进制字符串 $x=(1010)_2=10$ 和 $y=(0101)_2=5$ 为例, 它们的编码如表 1 所示. 从表中可以看出, x 的 1 编码和 y 的 0 编码存在交集, 因此可以判断出 $x>y$. 为便于描述, 文中用右上角为 0 或 1 表示时间经过 0 或 1 编码.

4.2　SM9 密钥封装机制

为便于理解本文, 首先给出 SM9 密钥封装机制的简单描述. PKG 首先根据安全参数 λ 生成系统主公钥 $\mathrm{mpk}=(\mathcal{BP},P_1,P_2,P_{\mathrm{pub}}=\alpha P_1,H_1,\mathrm{KDF},\mathrm{hid},\mathrm{len})$ 和主私钥 $\mathrm{msk}=\alpha$, 其中 P_1,P_2 分别为群 $\mathbb{G}_1$ 和 $\mathbb{G}_2$ 的生成元, $H_1:\{0,1\}^*\to\mathbb{Z}_p^*$ 为密码函数和 KDF: $\{0,1\}^*\to\{0,1\}^{\mathrm{len}}$ 为密钥派生函数, hid 为一个字节表示的私钥生成函数识别符, len 为封装密钥的长度. 给定用户标识 ID, PKG 计算用户私钥为 $\mathrm{sk}_{\mathrm{ID}}=\frac{\alpha}{H_1(\mathrm{ID}||\mathrm{hid},p)+\alpha}P_2$. 为生成封装密文 C 和封装密钥 K, 加密者首先选取随机数 $r\in\mathbb{Z}_p$, 计算 $C=r(H_1(\mathrm{ID}||\mathrm{hid},p)P_1+P_{\mathrm{pub}})$, $w=e(P_{\mathrm{pub}},P_2)^r$, $K=\mathrm{KDF}(C||w||\mathrm{ID},\mathrm{len})$. 为解密封装密文 C 得到封装密钥 K, 解密者 (拥有标识 ID) 首先利用私钥 $\mathrm{sk}_{\mathrm{ID}}$ 计算 $w'=e(C,\mathrm{sk}_{\mathrm{ID}})$, 计算封装密钥 $K'=\mathrm{KDF}(C||w'||\mathrm{ID},\mathrm{len})$. 为描述方便, 后文用 $H_1(\mathrm{ID})$ 代替 $H_1(\mathrm{ID}||\mathrm{hid},p)$.

4.3　前向安全的 SM9 密文等值测试方案

• **Setup.** 已知安全参数 λ, PKG 使用安全参数 λ 生成非对成群 $\mathcal{BP}=(\mathbb{G}_1,\mathbb{G}_2,\mathbb{G}_T,e,p)$, 其中, p 为大素数且 $p>2^\lambda$, 并随机选择生成元 $P_1\in\mathbb{G}_1$, $P_2\in\mathbb{G}_2$. 选择随机数 $\alpha,\beta\in\mathbb{Z}_p^*$, 并计算 $P_{\mathrm{pub}}=\alpha P_1$ 和 $\mathbb{G}_T$ 中元素 $g_1=e(P_1,P_2)^\alpha$, $g_2=e(P_1,P_2)^\beta$. 选择密码函数 $H_1:\{0,1\}^*\to\mathbb{Z}_p^*$, $H_2:\{0,1\}^*\to\{0,1\}^{n+2\log_2 p}$, $H_3:\{0,1\}^*\to\mathbb{G}_2$, $H_4:\{0,1\}^n\to\mathbb{Z}_p^*$, $H_5:\mathbb{G}_T\to\mathbb{Z}_p^*$, 其中 n 为明文的长度. 系统的主公钥 mpk 和主私钥 msk 为

$$\mathrm{mpk}=(\mathcal{BP},P_1,P_2,P_{\mathrm{pub}},g_1,g_2,H_1,H_2,H_3,H_4,H_5,n)\,,\ \mathrm{msk}=(\alpha,\beta)\,.$$

• **KeyGen.** 已知解密者标识 $\mathrm{ID}\in\{0,1\}^*$, PKG 计算 $d_1=\frac{\alpha}{H_1(\mathrm{ID})+\alpha}P_2$, $d_2=\frac{\beta}{H_1(\mathrm{ID})+\alpha}P_2$, 将 $\mathrm{sk}_{\mathrm{ID}}=(d_1,d_2)$ 作为用户私钥.

• **TdGen.** 对应标识 ID 的用户首先选择一个时间 t, 并对 t 进行 1 编码技术得到集合 S_t^1, 使用私钥 $\mathrm{sk_{ID}}$ 计算 $\{\mathrm{td}_{2,y}\}=\{r_y\left(H_1\left(\mathrm{ID}\right)+\alpha\right)P_1\}_{y\in S_t^1}$ 以及

$$\{\mathrm{td}_{1,y}\}=\left\{\frac{\beta}{H_1\left(\mathrm{ID}\right)+\alpha}P_2+r_yH_3\left(y\right)\right\}_{y\in S_t^1},$$

其中 r_y 表示为 S_t^1 中每个元素 y 随机选择的 $\mathbb{Z}_p$ 中元素. 最终得到陷门 $\mathrm{TD}=(\{\mathrm{td}_{1,y}\},\{\mathrm{td}_{2,y}\},\mathrm{td}_3)$ 其中 $y\in S_t^1$, $\mathrm{td}_3=S_t^1$.

• **Encrypt.** 已知用户标识 $\mathrm{ID}\in\{0,1\}^*$, 将要加密的明文 m, 系统主公钥 mpk, 加密者选择指定时间 t', 并使用 0 技术对其编码得到集合 $S_{t'}^0$, 选择两个随机数 $r_1,r_2\in\mathbb{Z}_p^*$, 计算 $w_1=g_1^{r_1}$, $w_2=g_2^{r_1}$, 计算 $C_1=r_1(H_1(\mathrm{ID})+\alpha)P_1$, $C_2=r_2(H_1(\mathrm{ID})+\alpha)P_1$, $C_3=(r_1||r_2||m)\oplus H_2(C_1||C_2||w_1||w_2||\mathrm{ID})$, $C_4=r_2\cdot H_4(m)\oplus H_5(g_2^{r_2})$, $\{C_{5,y}\}=\{r_2H_3(y)\}_{y\in S_{t'}^0}$. 最后输出密文 $\mathrm{CT}=(C_1,C_2,C_3,C_4,\{C_{5,y}\}_{y\in S_{t'}^0},S_{t'}^0)$.

• **Decrypt.** 已知密文 $\mathrm{CT}=(C_1',C_2',C_3',C_4',\{C_{5,y}'\}_{y\in S_{t'}^0},S_{t'}^0)$, 解密者标识 ID 及其私钥 $\mathrm{sk_{ID}}$, 解密者首先计算 $w_1'=e(C_1',d_1)$, $w_2'=e(C_1',d_2)$, 计算 $C_3'\oplus H_2(C_1'||C_2||w_1'||w_2'||\mathrm{ID})$ 得到 $r_1'||r_2'||m'$, 校验 $C_1'\stackrel{?}{=}r_1'(H_1(\mathrm{ID})+\alpha)P_1$, $C_2'\stackrel{?}{=}r_2'(H_1(\mathrm{ID})+\alpha)P_1$, $r_2'\cdot H_4(m')\oplus H_5(g_2^{r_2'})\stackrel{?}{=}C_4'$. 如果等式都校验通过, 那么输出 m', 否则输出 $\perp$.

• **Test.** 已知使用标识 ID_i 加密的两个密文 $\mathrm{CT}_i=(C_1^i,C_2^i,C_3^i,C_4^i,\{C_{5,y}^i\}_{y\in S_{t_i}^0},S_{t_i}^0)$, 对应的陷门 $\mathrm{TD}_i=(\{\mathrm{td}_{1,y}^i\}_{y\in\mathrm{td}_3^i},\{\mathrm{td}_{2,y}^i\}_{y\in\mathrm{td}_3^i},\mathrm{td}_3^i),i\in[1,2]$. 首先测试者比较 CT_i 中的 $S_{t_i}^0$ 与 TD_i 中的 td_3^i 是否有交集 ($i\in[1,2]$). 如果任意一组没有交集, 那么就无法有效进行密文等值测试, 返回 0. 否则, 两组都有交集, 令 ξ_i 表示交集中的任一元素, 计算

$$E_i=\frac{e(C_2^i,\mathrm{td}_{1,\xi_i}^i)}{e(\mathrm{td}_{2,\xi_i}^i,C_{5,\xi_i}^i)},\ X_i=C_4^i\oplus H_5\left(E_i\right).$$

校验 $E_1^{X_2}\stackrel{?}{=}E_2^{X_1}$, 如果校验成功就返回 1, 表示两个密文中嵌入的时间分别小于对应陷门中嵌入的时间, 并且底层明文相同, 否则返回 0 表示等值测试失败.

4.4　正确性分析

解密正确性分析. 假设 $\mathrm{CT}=(C_1',C_2',C_3',C_4',\{C_{5,y}'\}_{y\in S_{t'}^0},S_{t'}^0)$ 为正确封装的密文, 对应解密者标识为 ID 且其私钥为 $d_1=\frac{\alpha}{H_1(\mathrm{ID})+\alpha}P_2$, $d_2=\frac{\beta}{H_1(\mathrm{ID})+\alpha}P_2$, 则有

$$w_1'=e\left(C_1',d_1\right)=e\left(r_1\left(H_1(\mathrm{ID})+\alpha\right)P_1,\frac{\alpha}{H_1(\mathrm{ID})+\alpha}P_2\right)=e\left(P_1,\alpha P_2\right)^{r_1}=g_1^{r_1}=w_1,$$

$$w_2'=e\left(C_1',d_2\right)=e\left(r_1\left(H_1(\mathrm{ID})+\alpha\right)P_1,\frac{\beta}{H_1(\mathrm{ID})+\alpha}P_2\right)=e\left(P_1,\beta P_2\right)^{r_1}=g_2^{r_1}=w_2.$$

因此, 若 $\mathrm{CT}=(C_1',C_2',C_3',C_4',\{C_{5,y}'\}_{y\in S_{t'}^0},S_{t'}^0)$ 为正确封装的密文, 则 $w_1'=w_1,w_2'=w_2$, 可从 $C_3'\oplus H_2(C_1'||C_2'||w_1'||w_2'||\mathrm{ID})$ 中得到明文.

等值测试正确性分析.

• 假设测试者收到的密文、陷门都没有经过篡改, 并且密文中嵌入的时间分别小于对应陷门中嵌入的时间, 那么对 E_i 和 X_i 展开计算可得

$$\begin{aligned}E_i&=\frac{e(C_2^i,\mathrm{td}_{1,\xi_i}^i)}{e(\mathrm{td}_{2,\xi_i}^i,C_{5,\xi_i}^i)}=\frac{e(r_2^i(H_1(\mathrm{ID})+\alpha)P_1,\frac{\beta}{H_1(\mathrm{ID})+\alpha}P_2+r_{\xi_i}\cdot H_3(\xi_i))}{e(r_{\xi_i}(H_1(\mathrm{ID})+\alpha)P_1,r_2^i\cdot H_3(\xi_i))}\\&=\frac{e(r_2^i(H_1(\mathrm{ID})+\alpha)P_1,\frac{\beta}{H_1(\mathrm{ID})+\alpha}P_2)\cdot e(r_2^i(H_1(\mathrm{ID})+\alpha)P_1,r_{\xi_i}\cdot H_3(\xi_i))}{e(r_{\xi_i}(H_1(\mathrm{ID})+\alpha)P_1,r_2^i\cdot H_3(\xi_i))}\\&=e(P_1,P_2)^{\beta r_2^i}\end{aligned}$$

$$= g_2^{r_2^i},$$

$$X_i = C_4^i \oplus H_5(E_i) = r_2^i \cdot H_4(m_i).$$

得 $E_1^{X_2} = g_2^{r_2^1 \cdot r_2^2 \cdot H_4(m_1)}$, $E_2^{X_1} = g_2^{r_2^2 \cdot r_2^1 \cdot H_4(m_2)}$. 若 $E_1^{X_2} = E_2^{X_1}$, 则 $H_4(m_1) = H_4(m_2)$, 即 $m_1 = m_2$.

• 否则, 若任意一个密文中嵌入的时间不小于对应陷门中嵌入的时间, 由于陷门与密文的时间编码集合无交集, 导致无法正确计算 E_i; 或者当 $m_1 \neq m_2$ 时, 由于哈希函数 H_4 的抗碰撞性, 使得 $H_4(m_1) \neq H_4(m_2)$, 因此等值测试算法将返回结果 0.

4.5　安全性分析

SM9-FSET 方案的安全性基于 q-BDHI 问题的困难性假设. 在随机谕言机模型中给出了方案的安全性证明. 首先给出非对称群上的 q-BDHI 问题的定义.

定义3 (q-BDHI 问题)　已知双线性群 $\mathcal{BP} = (\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, e, p)$, q+2 个元素 $(P, Q, aQ, a^2Q, \ldots, a^qQ) \in \mathbb{G}_1 \times \mathbb{G}_2^{q+1}$, 其中 a 未知, P, Q 分别为群 $\mathbb{G}_1, \mathbb{G}_2$ 生成元, 计算 $e(P, Q)^{1/a}$.

由于 SM9 标识密码基于第 2 类双线性群, 因此存在群 $\mathbb{G}_2$ 到群 $\mathbb{G}_1$ 的同构映射 $\psi : \mathbb{G}_2 \to \mathbb{G}_1$, 满足 $\psi(\mu Q) = \mu P, \forall \mu \in \mathbb{Z}_p^*$.

定理1　令方案中的密码函数 $H_i, i \in \{1, 2, 3, 5\}$ 为随机谕言机. 如果 q-BDHI 问题是难解的, 则本文提出的 SM9-FSET 方案满足 IND-ID-CCA 的安全性.

证明　假定存在攻击算法 $\mathcal{A}$ 能以不可忽略的优势 (概率) ε 攻破方案, 则可构造一个模拟算法 $\mathcal{B}$ 通过与 $\mathcal{A}$ 交互, 以不可忽略的优势求解 q-BDHI 问题. 在证明中 $\mathcal{B}$ 扮演 IND-ID-CCA 安全模型中 $\mathcal{C}$ 的角色. $\mathcal{B}$ 以 q-BDHI 问题实例 $(P, Q, aQ, a^2Q, \ldots, a^qQ)$ 为输入, 其目标是求 $e(P, Q)^{1/a}$, 其中 a 未知, $q = q_{H_1}$ 为询问随机谕言机 H_1 的次数.

系统建立阶段. 令 $H_i (i \in \{1, 2, 3, 5\})$ 为随机谕言机由 $\mathcal{B}$ 控制. $\mathcal{B}$ 随机选取 $k^*, x_1, x_2 \in \mathbb{Z}_p^*$, a 未知的前提下隐含的令 $\alpha = a - k^*$, $\beta = x_1 + x_2\alpha$. 接着, 随机选择 $i^* \in [1, q]$, 并从 $\mathbb{Z}_p^*$ 中选取 $q-1$ 个彼此不同且与 k^* 不同的随机数 $k_1, k_2, \ldots, k_{i^*-1}, k_{i^*+1}, \ldots, k_q$. 定义

$$f(x) = \prod_{i=1, i \neq i^*}^{q} (x - k^* + k_i) = \sum_{i=0}^{q-1} c_i x^i \mod p,\ f_i(x) = \frac{f(x)}{x - k^* + k_i} = \sum_{i=0}^{q-2} b_i x^i \mod p,$$

那么

$$f(a)Q = \sum_{i=0}^{q-1} c_i \left(a^i Q\right), af(a)Q = \sum_{i=0}^{q-1} c_i \left(a^{i+1} Q\right), f_i(a)Q = \sum_{i=0}^{q-2} b_i \left(a^i Q\right), af_i(a)Q = \sum_{i=0}^{q-2} b_i \left(a^{i+1} Q\right),$$

可以通过已知问题实例计算得到. 接着计算

$$P_2 = \sum_{i=0}^{q} c_i \left(a^i Q\right) \mod p = f(a)Q, P_1 = \psi(P_2) = f(a)P,\ P_{\text{pub}} = \psi(af(a)Q) - k^*\psi(P_2) = \alpha P_1,$$

$$g_1 = e(P_{\text{pub}}, P_2) = e(P_1, P_2)^{\alpha},\ g_2 = e(P_1, P_2)^{x_1} \cdot e(P_{\text{pub}}, P_2)^{x_2} = e(P_1, P_2)^{x_1 + x_2\alpha} = e(P_1, P_2)^{\beta}.$$

最后 $\mathcal{B}$ 发送 $\text{mpk} = (P_1, P_2, P_{\text{pub}}, g_1, g_2)$ 给 $\mathcal{A}$. 从系统公钥的设置可知, mpk 中的元素都可以通过已知问题实例计算得到.

哈希询问阶段. $\mathcal{A}$ 被允许询问随机谕言机 $H_i, i \in \{1, 2, 3, 5\}$.

(1) H_1-询问. $\mathcal{B}$ 首先建立列表 $\mathcal{L}_1$ 用于记录 H_1 询问的输入和输出, 表中元素以 (ID, k) 形式存储. 令 ID_i 为第 i 个 H_1 询问, 若 ID_i 在 $\mathcal{L}_1$ 中, 则返回相应的 k_i. 否则设

$$H_1(\text{ID}_i) = \begin{cases} k^*, & \text{if } i = i^*, \\ k_i, & \text{otherwise.} \end{cases}$$

最后, $\mathcal{B}$ 发送 k^* 或者 k_i 给 $\mathcal{A}$ 并以 (ID_i, k_i) 或者 (ID_i, k^*) 更新 $\mathcal{L}_1$.

(2) H_2-询问. $\mathcal{B}$ 首先建立列表 $\mathcal{L}_2$ 用于记录 H_2 询问的输入和输出, 表中元素以 $(C_1, C_2, w_1, w_2, \mathrm{ID}, \gamma)$ 形式存储. 令 $(C_{1,i}, C_{2,i}, w_{1,i}, w_{2,i}, \mathrm{ID}_i, \gamma_i)$ 为第 i 个 H_2 询问, 如果 $(C_{1,i}, C_{2,i}, w_{1,i}, w_{2,i}, \mathrm{ID}_i)$ 在列表 $\mathcal{L}_2$ 中, 则返回相应的 γ_i. 否则以 ID_i 发起 H_1 询问得到响应 k', $\mathcal{B}$ 根据以下步骤回复 $\mathcal{A}$.

• $\mathrm{ID}_i = \mathrm{ID}_{i^*}$. 即表示 $k' = k^*$, $\mathcal{B}$ 判断以下等式是否成立:

$$w_{2,i} = \left(\frac{e\left(C_{1,i}, P_2\right)}{w_{1,i}}\right)^{\frac{x_1}{k^*}} \cdot w_{1,i}^{x_2}. \tag{1}$$

(i) 若等式 (1) 成立, 则检查列表 $\mathcal{L}_D$ (解密询问所建列表) 中是否存在元素包含 $(C_{1,i}, C_{2,i}, \mathrm{ID}_i)$. 若存在, 则获取列表 $\mathcal{L}_D$ 中对应元素在面临这 3 个元素相同的解密询问时处理过程中获取的 H_2 的值 γ_i. 若不存在, 则 $\mathcal{B}$ 从 $\{0,1\}^{n+2\log_2 p}$ 中随机选取元素 γ_i, 设 $H_2(C_{1,i}, C_{2,i}, w_{1,i}, w_{2,i}, \mathrm{ID}_i) = \gamma_i$.

(ii) 否则, $\mathcal{B}$ 从 $\{0,1\}^{n+2\log_2 p}$ 中随机选取元素 γ_i, 设 $H_2(C_{1,i}, C_{2,i}, w_{1,i}, w_{2,i}, \mathrm{ID}_i) = \gamma_i$.

• $\mathrm{ID}_i \neq \mathrm{ID}_{i^*}$. $\mathcal{B}$ 从 $\{0,1\}^{n+2\log_2 p}$ 中随机选取元素 γ_i, 设 $H_2(C_{1,i}, C_{2,i}, w_{1,i}, w_{2,i}, \mathrm{ID}_i) = \gamma_i$.

最后 $\mathcal{B}$ 发送 γ_i 给 $\mathcal{A}$ 并以 $(C_{1,i}, C_{2,i}, w_{1,i}, w_{2,i}, \mathrm{ID}_i, \gamma_i)$ 更新 $\mathcal{L}_2$.

(3) H_3-询问. $\mathcal{B}$ 首先建立列表 $\mathcal{L}_3$ 用于记录 H_3 询问的输入和输出, 表中元素以 (y, μ, v, s) 形式存储, 其中 $s = 0$ 表示 $v = \mu P_2$, $s = 1$ 表示 $v = \mu a P_2$. 令 y_i 为第 i 个 H_3 询问, 若 y_i 在 $\mathcal{L}_3$ 中则返回相应的 v_i, 否则 $\mathcal{B}$ 随机选择 $\mu_i \in \mathbb{Z}_p^*$ 计算 $v_i = \mu_i P_2$ 发给 $\mathcal{A}$, 并以 $(y_i, \mu_i, v_i, 0)$ 更新 $\mathcal{L}_3$.

(4) H_5-询问. $\mathcal{B}$ 首先建立列表 $\mathcal{L}_5$ 用于记录 H_5 询问的输入和输出, 表中元素以 (ς, ϑ) 形式存储. 令 ς_i 为第 i 个 H_5 询问, 若 ς_i 在 $\mathcal{L}_5$ 中则返回相应的 ϑ_i, 否则 $\mathcal{B}$ 随机选择 $\vartheta_i \in \mathbb{Z}_p^*$, 发送 ϑ_i 给 $\mathcal{A}$ 并以 $(\varsigma_i, \vartheta_i)$ 更新 $\mathcal{L}_5$.

询问阶段 1. 在这一阶段, $\mathcal{B}$ 允许私钥询问、陷门询问和密文解密询问.

(1) 私钥询问. 已知标识 ID_i, 如果 $\mathrm{ID}_i = \mathrm{ID}_{i^*}$, $\mathcal{B}$ 停止模拟并输出失败. 否则计算

$$d_{1,i} = \left(a - k^*\right) f_i\left(a\right) Q = \frac{\left(a - k^*\right) f\left(a\right)}{a - k^* + k_i} Q = \frac{\alpha}{H_1\left(\mathrm{ID}_i\right) + \alpha} P_2,$$

$$d_{2,i} = \left(x_1 + x_2\left(a - k^*\right)\right) f_i\left(a\right) Q = \frac{\left(x_1 + x_2\left(a - k^*\right)\right) f\left(a\right)}{a - k^* + k_i} Q = \frac{\beta}{H_1\left(\mathrm{ID}_i\right) + \alpha} P_2.$$

结合系统建立阶段可知, 通过上述等式可正确计算私钥. 最后 $\mathcal{B}$ 将 $\mathrm{sk}_{\mathrm{ID}} = (d_{1,i}, d_{2,i})$ 发给 $\mathcal{A}$.

(2) 陷门询问. 已知标识 ID_i 时间 t_i, 模拟算法 $\mathcal{B}$ 建立列表 $\mathcal{L}_{\mathrm{td}}$ 用于记录陷门询问的输入和输出, 列表元素以 $(\mathrm{ID}, \mathrm{TD}, t)$ 的形式存储. 若 (ID_i, t_i) 在列表 $\mathcal{L}_{\mathrm{td}}$ 中, 则返回相应的 TD_i, 否则根据以下步骤回复.

• $\mathrm{ID}_i \neq \mathrm{ID}_{i^*}$. $\mathcal{B}$ 首先以 ID_i 发起私钥询问得到 $\mathrm{sk}_{\mathrm{ID}_i} = (d_{i,1}, d_{i,2})$, 然后以 $\mathrm{sk}_{\mathrm{ID}_i}$ 和 t_i 为输入运行 TdGen 算法获得 TD_i 并发给 $\mathcal{A}$.

• $\mathrm{ID}_i = \mathrm{ID}_{i^*}$. $\mathcal{B}$ 首先获取时间 t_i 对应的 1 编码集合 $S_{t_i}^1$, 对于集合中每个元素 y 选择对应的随机数 $\eta_{y,1}$. $\mathcal{B}$ 检查 $\mathcal{L}_3$ 中是否存在元素包含 y, 如果不存在则以 y 发起 H_3 询问. 最终 $\mathcal{B}$ 得到 (y, μ, v, s), $\mathcal{B}$ 根据 s 的值决定以何种方式计算陷门

(i) $s = 0$. 表示 $v = \mu P_2$, 即 $H_3(y) = \mu P_2$, 令 $r_y = \eta_{y,1} - \frac{\beta}{\mu(H_1(\mathrm{ID})+\alpha)}$, 令 $\mathrm{td}_3 = S_{t_i}^1$, 计算

$$\begin{aligned}
\mathrm{td}_{1,y} &= \frac{\beta}{H_1(\mathrm{ID}) + \alpha} P_2 + r_y H_3\left(y\right) = \frac{\beta}{H_1(\mathrm{ID}) + \alpha} P_2 + \left(\eta_{y,1} - \frac{\beta}{t\left(H_1(\mathrm{ID}) + \alpha\right)}\right) \mu \cdot P_2 \\
&= \eta_{y,1} \mu \cdot P_2, \\
\mathrm{td}_{2,y} &= r_y \left(H_1(\mathrm{ID}) + \alpha\right) P_1 = \left(\eta_{y,1} - \frac{\beta}{\mu\left(H_1(\mathrm{ID}) + \alpha\right)}\right) \cdot \left(H_1(\mathrm{ID}) + \alpha\right) P_1 \\
&= \eta_{y,1} k^* P_1 + \eta_{y,1} \alpha P_1 - \frac{x_1 + x_2 \alpha}{\mu} P_1
\end{aligned}$$

$$= \left(\eta_{y,1}k^* - \frac{x_1}{\mu}\right) P_1 + \left(\eta_{y,1} - \frac{x_2}{\mu}\right) \alpha P_1.$$

由于 $\eta_{y,1}, \mu$ 是随机选取的, 因此上式得到的陷门于 $\mathcal{A}$ 而言不可区分. 最后模拟者 $\mathcal{B}$ 将 $\mathrm{TD}_i = (\{\mathrm{td}_{1,y}\}, \{\mathrm{td}_{2,y}\}, \mathrm{td}_3)$ 发给 $\mathcal{A}$, 并以 $(\mathrm{ID}_i, \mathrm{TD}_i, t_i)$ 更新 $\mathcal{L}_{\mathrm{td}}$.

(ii) $s = 1$. 表示 $v = \mu a P_2$, 即 $H_3(y) = \mu a P_2$, 无法计算出陷门, $\mathcal{B}$ 返回失败 $\perp$.

(3) 密文解密询问. 已知密文 $\mathrm{CT}_i = (C_{1,i}, C_{2.i}, C_{3,i}, C_{4,i}, \{C_{5,y}\}_{y \in S^0_{t_i}}, S^0_{t_i})$, $\mathcal{B}$ 首先建立列表 $\mathcal{L}_D$ 用于记录密文解密询问的输入和输出, 列表元素以 $(\mathrm{CT}, \mathrm{ID}, \tau)$ 的形式存储. 若 $(\mathrm{CT}_i, \mathrm{ID}_i)$ 在列表 $\mathcal{L}_D$ 中, 则返回相应的 τ_i, 否则根据以下步骤回复.

• $\mathrm{ID}_i \neq \mathrm{ID}_{i^*}$. $\mathcal{B}$ 首先计算标识 ID_i 的私钥 $\mathrm{sk}_{\mathrm{ID}_i}$. 以 $\mathrm{sk}_{\mathrm{ID}_i}$ 和 CT_i 为输入运行解密算法获得 $w_{1,i}, w_{2,i}$, 以 $(C_{1,i}, C_{2.i}, w_{1,i}, w_{2,i}, \mathrm{ID}_i)$ 为输入询问 H_2 谕言机得到 γ_i, 计算 $C_{3,i} \oplus \gamma_i$ 并截取后 n 比特长度作为结果 τ_i 发给 $\mathcal{A}$.

• $\mathrm{ID}_i = \mathrm{ID}_{i^*}$. $\mathcal{A}$ 在生成密文之前必经过 H_2 询问, 故可从 H_2 询问记录获取在此次解密询问前 $\mathcal{A}$ 进行 H_2 询问的参数及结果 γ_i. 若 $w_{1,i}$, $w_{2,i}$ 满足等式 (1), 则计算 $C_{i,3} \oplus \gamma_i = r'_{1,i} || r'_{2,i} || m'_i$, 并校验 $w_{1,i} \stackrel{?}{=} g_1^{r'_{1,i}}$, $w_{2,i} \stackrel{?}{=} g_2^{r'_{1,i}}$, $C_{1,i} \stackrel{?}{=} r'_{1,i}(H_1(\mathrm{ID}_i) + \alpha) P_1$, $C_{2,i} \stackrel{?}{=} r'_{2,i}(H_1(\mathrm{ID}_i) + \alpha) P_1$. 若均校验通过则令 $\tau_i = m'_i$, 否则令 $\tau_i = \perp$, 并将 τ_i 发给 $\mathcal{A}$.

最后, $\mathcal{B}$ 以 $(\mathrm{CT}_i, \mathrm{ID}_i, \tau_i)$ 更新 $\mathcal{L}_D$.

挑战阶段. $\mathcal{A}$ 输出挑战标识 ID^* 和挑战时间 t^* 和两个明文 $m_0, m_1 \in \{0,1\}^n$, 若 $\mathrm{ID}^* \neq \mathrm{ID}_{i^*}$ 或 $\mathcal{A}$ 已获取 ID^* 的嵌入时间大于 t^* 的陷门, $\mathcal{B}$ 停止模拟并输出失败, 否则有 $H_1(\mathrm{ID}^*) = k^*$. $\mathcal{B}$ 随机选取 $c \in \{0,1\}$, 令 $m^* = m_c$, 选取随机数 $r_1, r_2 \in \mathbb{Z}_p^*$, $R_1 \in \{0,1\}^{n+2\log_2 p}$, $R_2 \in \{0,1\}^{2\log_2 p}$. 获取 t^* 的 0 编码集合 $S^0_{t^*}$, $S^0_{t^*}$ 中元素需均未出现在 $\mathcal{L}_3$ 中. 对每个 $y \in S^0_{t^*}$, 选择随机数 η_y, 令 $H_3(y) = a\eta_y P_2$, 并以 $(y, \eta_y, \eta_y a P_2, 1)$ 更新 $\mathcal{L}_3$. 计算挑战密文 $C_1^* = r_1 P_1, C_2^* = r_2 P_1, C_3^* = R_1, C_4^* = R_2, \{C_{5,y}^*\} = \{r_2 \eta_y P_2\}_{y \in S^0_{t^*}}$, 令 $\mathrm{CT}^* = (C_1^*, C_2^*, C_3^*, C_4^*, \{C_{5,y}^*\}_{y \in S^0_{t^*}}, S^0_{t^*})$, 若 $(\mathrm{CT}^*, \mathrm{ID}^*)$ 被解密询问过, 则重新计算 CT^*. 最后发送 $(\mathrm{CT}^*, \mathrm{ID}^*)$ 给 $\mathcal{A}$. 设 CT^* 中的随机数 $r_1^* = \frac{r_1}{a}, r_2^* = \frac{r_2}{a}$, 则有

$$C_1^* = r_1^* \left(H_1\left(\mathrm{ID}^*\right) P_1 + P_{\mathrm{pub}}\right) = \frac{r_1}{a}\left(k^* P_1 + \left(a - k^*\right) P_1\right) = r_1 P_1,$$

$$C_2^* = r_2^* \left(H_1\left(\mathrm{ID}^*\right) P_1 + P_{\mathrm{pub}}\right) = \frac{r_2}{a}\left(k^* P_1 + \left(a - k^*\right) P_1\right) = r_2 P_1,$$

$$C_{5,y}^* = r_2^* H_3\left(y\right) = \frac{r_2}{a} \cdot a \eta_y P_2 = r_2 \eta_y P_2.$$

因此, C_1^*, C_2^*, $C_{5,y}^*$ 是由随机数 r_1, r_2, η_y 生成的有效的密文组成部分. 密文除 C_3^*, C_4^* 是 $\mathcal{A}$ 不可区分之外, 其他部分都能根据已知信息得到.

询问阶段 2. $\mathcal{A}$ 允许继续询问私钥、陷门和密文解密, 但不能询问挑战标识 ID^* 的私钥和挑战密文 $(\mathrm{CT}^*, \mathrm{ID}^*)$ 的解密, 且询问关于 ID^* 的陷门时, 指定的时间不能大于挑战密文中嵌入的时间. $\mathcal{B}$ 根据询问阶段 1 回复 $\mathcal{A}$.

猜测阶段. $\mathcal{A}$ 输出它的猜测 c'. $\mathcal{B}$ 忽略 $\mathcal{A}$ 的猜测, $\mathcal{B}$ 分两种情况得到 q-BDHI 问题实例的解.

• $\mathcal{A}$ 通过解密攻破方案. 定义 H_2 中的挑战询问为 $(C_1^*, C_2^*, w_1^*, w_2^*, \mathrm{ID}^*)$, 其中 $w_1^* = e(P_1, P_2)^{r_1^* \alpha}$, $w_2^* = e(P_1, P_2)^{r_1^* \beta}$. 接着, $\mathcal{B}$ 从列表 $\mathcal{L}_2$ 中找到满足等式 (2) 的 w_1^*, w_2^*,

$$w_2^* = \left(\frac{e\left(r_1 P_1, P_2\right)}{w_1^*}\right)^{\frac{x_1}{k^*}} \cdot \left(w_1^*\right)^{x_2}. \tag{2}$$

则有 $w_1^* = e\left(P_1, P_2\right)^{r_1^* \alpha} = e\left(f\left(a\right) P, f\left(a\right) Q\right)^{r_1 (a - k^*)/a} = e\left(P_1, P_2\right)^{r_1} \cdot e\left(P, Q\right)^{-r_1 k^* f^2(a)/a}$. 又由于 $x \nmid f^2\left(x\right)$, 因此 $-r_1 k^* f^2\left(x\right)/x = F\left(x\right) + d/x$, 其中 $F\left(x\right)$ 是一个 $2q-3$ 次多项式, d 是一个可求的非零整

数. 则 $e\left(P,Q\right)^{F(a)}$ 是可求的. 最后, $\mathcal{B}$ 通过以下计算求解 q-BDHI 问题实例:

$$\left(\frac{w_1^*}{e\left(P_1,P_2\right)^{r_1}\cdot e\left(P,Q\right)^{F(a)}}\right)^{\frac{1}{d}}=\left(\frac{e\left(P_1,P_2\right)^{r_1}\cdot e\left(P,Q\right)^{F(a)+\frac{d}{a}}}{e\left(P_1,P_2\right)^{r_1}\cdot e\left(P,Q\right)^{F(a)}}\right)^{\frac{1}{d}}=e\left(P,Q\right)^{\frac{1}{a}}.$$

• $\mathcal{A}$ 通过等值测试攻破方案. 计算挑战密文相关的 E^* (对挑战密文进行密文等值测试过程中的临时变量) 由 $\mathcal{B}$ 通过已知信息正确计算. 若 $\mathcal{A}$ 通过等值测试攻破方案, 必定进行 $H_5(E^*)$ 询问. 因此 $\mathcal{B}$ 随机从 $\mathcal{L}_5$ 中选取 $(\varsigma^*,\vartheta^*)$, 若 ς^* 恰好是 $\mathcal{A}$ 对 E^* 的 H_5 询问, 那么

$$\varsigma^*=g_2^{r_2^*}=e\left(f\left(a\right)P,f\left(a\right)Q\right)^{\frac{\beta r_2}{a}}=e\left(P_1,P_2\right)^{x_2r_2}\cdot e\left(P,Q\right)^{\frac{\left(x_1r_2-x_2k^*r_2\right)f^2(a)}{a}}.$$

又由于 $x\nmid f^2\left(x\right)$, 因此 $\left(x_1r_2-x_2k^*r_2\right)f^2\left(x\right)/x=F\left(x\right)+d'/x$, 其中 $F(x)$ 是一个 $2q-3$ 次多项式, d' 是一个可求的非零整数, 注意到 $\psi(Q)=P,\ f(a)P=\psi(f(a)Q)$, 则 $e(P,Q)^{F(a)}$ 是可求的. 最后, $\mathcal{B}$ 通过以下计算求解 q-BDHI 问题实例:

$$\left(\frac{\varsigma^*}{e\left(P_1,P_2\right)^{x_2r_2}\cdot e\left(P,Q\right)^{F(a)}}\right)^{\frac{1}{d'}}=\left(\frac{e\left(P_1,P_2\right)^{x_2r_2}\cdot e\left(P,Q\right)^{F(a)+\frac{d'}{a}}}{e\left(P_1,P_2\right)^{x_2r_2}\cdot e\left(P,Q\right)^{F(a)}}\right)^{\frac{1}{d'}}=e\left(P,Q\right)^{\frac{1}{a}}.$$

证明中选取的元素都是随机的, 模拟环境和真实攻击具有不可区分性. 若模拟者正确地猜到了挑战标识 ID* 是第 i^* 个询问 H_1 的标识, $\mathcal{A}$ 询问的关于 ID* 的陷门嵌入时间均小于等于挑战密文中嵌入的时间 t^*, 并且 $\mathcal{A}$ 的 H_3 询问不包含 $S_{t^*}^0$ 中任何元素. 满足以上条件才能模拟成功.

模拟者正确猜到挑战标识 ID* 是第 i^* 个询问 H_1 的标识的概率为 $\frac{1}{q_{H_1}}$. 令 $\mathcal{A}$ 发起的 H_3 询问不包含 $S_{t^*}^0$ 中任何元素的概率为 ε_3. 假设时间对应的二进制比特长度固定为 ℓ, 则时间进行 0/1 编码后的集合大小一定不大于 ℓ. ℓ 比特长度的二进制数进行 1 编码后的集合元素最少有 2^ℓ 种可能, 易得 $\varepsilon_3>(1-\frac{\ell}{2^\ell})^{q_{H_3}}$, q_{H_3} 表示 $\mathcal{A}$ 发起 H_3 询问的次数. 因此模拟成功的概率为 $\frac{1}{q_{H_1}}(1-\frac{\ell}{2^\ell})^{q_{H_3}}$.

此外, 若攻击算法 $\mathcal{A}$ 既没有对 $(C_1^*,C_2^*,w_1^*,w_2^*,\mathrm{ID}^*)$ 执行 H_2 询问, 也没有对 E^* 执行 H_5 询问, 则没有优势攻破 SM9-FSET 算法. 根据假设, $\mathcal{A}$ 能以不可忽略的优势 ε 攻破算法, 则 $\mathcal{A}$ 将以 ε 概率询问 $(C_1^*,C_2^*,w_1^*,w_2^*,\mathrm{ID}^*)$ 对应 H_2 的值或者询问 E^* 对应 H_5 的值. 综上所述, $\mathcal{B}$ 可以成功求解 q-BDHI 困难问题的概率大于 $\frac{\varepsilon}{q_{H_1}}(1-\frac{\ell}{2^\ell})^{q_{H_3}}$.

定理2　令方案中的密码函数 $H_i,i\in\{1,2,3\}$ 为随机谕言机. 如果 q-BDHI 问题是难解的, 则本文提出的 SM9-FSET 方案满足 OW-ID-CCA 的安全性.

证明　假定存在攻击算法 $\mathcal{A}$ 能以不可忽略的优势 ε 攻破方案, 则可构造一个模拟算法 $\mathcal{B}$ 通过与 $\mathcal{A}$ 交互, 以不可忽略的优势求解 q-BDHI 问题. 在证明中 $\mathcal{B}$ 扮演 OW-ID-CCA 安全模型中 $\mathcal{C}$ 的角色. $\mathcal{B}$ 以 q-BDHI 问题实例 $(P,Q,aQ,a^2Q,\ldots,a^qQ)$ 为输入, 其目标是求 $e(P,Q)^{1/a}$, 其中 a 未知, $q=q_{H_1}$ 为询问随机谕言机 H_1 的次数.

系统建立阶段. 令 $H_i(i\in\{1,2,3\})$ 为随机谕言机由 $\mathcal{B}$ 控制. 其他设置与定理 1 证明中相同.

哈希询问阶段. $\mathcal{A}$ 被允许询问随机谕言机 $H_i,i\in\{1,2,3\}$.

(1) H_1-询问与 H_2-询问. 与定理 1 证明中相同.

(2) H_3-询问. $\mathcal{B}$ 首先建立列表 $\mathcal{L}_3$ 用于记录 H_3 询问的输入和输出, 表中元素以 (y,μ,v) 形式存储. 令 y_i 为第 i 个 H_3 询问, 若 y_i 在 $\mathcal{L}_3$ 中则返回相应的 v_i, 否则 $\mathcal{B}$ 随机选择 $\mu_i\in\mathbb{Z}_p^*$ 计算 $v_i=\mu_iP_2$ 发给 $\mathcal{A}$, 并以 (y_i,μ_i,v_i) 更新 $\mathcal{L}_3$.

询问阶段 1. 在这一阶段, $\mathcal{B}$ 允许询问私钥、陷门和密文的解密.

(1) 私钥询问和密文解密询问. 与定理 1 证明中相同.

(2) 陷门询问. 已知标识 ID_i 和时间 t_i, 模拟算法 $\mathcal{B}$ 建立列表 $\mathcal{L}_{\mathrm{td}}$ 用于记录陷门询问的输入和输出, 列表元素以 $(\mathrm{ID},\mathrm{TD},t)$ 的形式存储. 若 (ID_i,t_i) 在列表 $\mathcal{L}_{\mathrm{td}}$ 中, 则返回相应的 TD_i, 否则根据以下步骤回复.

• $\mathrm{ID}_i \neq \mathrm{ID}_{i^*}$. 回复方式与定理 1 证明中相同.

• $\mathrm{ID}_i = \mathrm{ID}_{i^*}$. $\mathcal{B}$ 首先获取时间 t_i 对应的 1 编码集合 $S_{t_i}^1$, 对于集合中每个元素 y 选择对应的随机数 $\eta_{y,1}$. 检查 $\mathcal{L}_3$ 中是否存在元素包含 y, 如果不存在则以 y 发起 H_3 询问. 最终 $\mathcal{B}$ 得到 (y, μ, v). 已知 $v = \mu P_2$, 即 $H_3(y) = \mu P_2$, 令 $r_y = \eta_{y,1} - \frac{\beta}{\mu(H_1(\mathrm{ID})+\alpha)}$, $\mathrm{td}_3 = S_{t_i}^1$. $\mathrm{td}_{1,y}$ 和 $\mathrm{td}_{2,y}$ 的计算方式与定理 1 证明中陷门询问阶段 $\mathrm{ID}_i = \mathrm{ID}_i^*$ 且 $s = 0$ 相同.

最后 $\mathcal{B}$ 以 $(\mathrm{ID}_i, \mathrm{TD}_i, t_i)$ 更新 $\mathcal{L}_{\mathrm{td}}$.

挑战阶段. $\mathcal{A}$ 输出挑战标识 ID^* 和时间 t^*, 如果 $\mathrm{ID}^* \neq \mathrm{ID}_{i^*}$, $\mathcal{B}$ 停止模拟并输出失败, 否则有 $H_1(\mathrm{ID}^*) = k^*$. $\mathcal{B}$ 随机选取 $r_1, r_2 \in \mathbb{Z}_p^*$, $m^* \in \{0,1\}^n$, $R_1 \in \{0,1\}^{n+2\log_2 p}$. 获取 t^* 的 0 编码集合 $S_{t^*}^0$. 令 $H_3(y) = \eta_y P_2$ (集合 $S_{t^*}^0$ 中每个元素 y 对应选取一个随机数 η_y, 并以 $(y, \eta_y, \eta_y P_2)$ 更新 $\mathcal{L}_3$), 计算挑战密文 $C_1^* = r_1 P_1, C_2^* = r_2 P_1, C_3^* = R_1, C_4^* = r_2 H_4(m^*) \oplus H_5(g_2^{r_2}), \{C_{5,y}^*\} = \{r_2 \eta_y P_2\}_{y \in S_{t^*}^0}$, 令 $\mathrm{CT}^* = (C_1^*, C_2^*, C_3^*, C_4^*, \{C_{5,y}^*\}_{y \in S_{t^*}^0}, S_{t^*}^0)$, 要求 $(\mathrm{CT}^*, \mathrm{ID}^*)$ 没有询问过解密操作, 否则重新计算 CT^*. 最后发送 $(\mathrm{CT}^*, \mathrm{ID}^*)$ 给 $\mathcal{A}$. 设挑战密文中的随机数 $r_1^* = \frac{r_1}{a}$, 则有

$$C_1^* = r_1^* \left(H_1\left(\mathrm{ID}^*\right) P_1 + P_{\mathrm{pub}}\right) = \frac{r_1}{a}\left(k^* P_1 + \left(a - k^*\right) P_1\right) = r_1 P_1.$$

因此, C_1^* 是由随机数 r_1 生成的有效密文组成部分. 密文除 C_3^* 是 $\mathcal{A}$ 不可区分之外, 其他部分都能根据已知信息得到.

询问阶段 2. $\mathcal{A}$ 允许继续询问私钥、陷门和密文解密, 但不能询问挑战标识 ID^* 的私钥和挑战密文 $(\mathrm{CT}^*, \mathrm{ID}^*)$ 的解密.

猜测阶段. $\mathcal{A}$ 输出它的猜测 m'. 此时, $\mathcal{B}$ 忽略 $\mathcal{A}$ 的猜测结果, 并通过定理 1 证明中猜测阶段分类 1 相同的方式获取 q-BDHI 问题实例的解.

从以上设置可知, 证明中选取的元素都是随机的, 模拟环境和真实攻击具有不可区分性. 若挑战标识 ID^* 是第 i^* 个询问 H_1 的标识, $\mathcal{A}$ 没有询问过 ID^* 的私钥, 模拟成功的概率为 $\frac{1}{q_{H_1}}$.

此外, 若攻击算法 $\mathcal{A}$ 没有对 $(C_1^*, C_2^*, w_1^*, w_2^*, \mathrm{ID}^*)$ 执行 H_2 询问, 则没有优势攻破 SM9-FSET 算法. 根据假设, $\mathcal{A}$ 能以不可忽略的优势 ε 攻破算法, 则 $\mathcal{A}$ 将以 ε 概率询问 $(C_1^*, C_2^*, w_1^*, w_2^*, \mathrm{ID}^*)$ 对应 H_2 的值. 通过等式 (2), $\mathcal{B}$ 可成功找到 (w_1^*, w_2^*) 并求出给定 q-BDHI 问题的解. 综上所述, $\mathcal{B}$ 可以成功求解 q-BDHI 困难问题的概率为 ε / q_{H_1}.

5　方案性能分析

本节通过理论与实验对比分析SM9-FSET方案与相关方案. 现有 IBEET 普遍缺乏对陷门寿命的有效控制, 2024 年, Ma 等 [7] 首次提出了前向安全性的标识密文等值测试概念, 并给出了具体方案. 由于目前基于 SM9 的密文等值测试方案仅见文献 [11]. 因此, 本文将与文献 [7,11] 进行对比.

5.1　理论分析

本小节将从理论方面分析 SM9-FSET 方案的计算效率和存储效率以及安全性, 并与文献 [7] (支持 Tem-Tem 测试的方案) 和文献 [11] 进行比较. 首先比较私钥生成、陷门生成、加密、解密和测试算法的计算开销, 结果如表 2 所示. 接着比较存储开销和安全性, 结果如表 3 所示. 符号说明: H_2 表示哈希到非对称群 $\mathbb{G}_2$ 中运算, H 表示哈希到对称群 $\mathbb{G}$ 中运算 (只考虑计算消耗大的哈希运算). $\mathcal{P}$ 表示对称双线性对的双线性运算, $\mathcal{P}_1$ 表示非对称双线性对的双线性运算. M_i $(i = 1, 2)$ 表示非对称群 $\mathbb{G}_i$ 中的标量乘运算 (等价乘法群中的指数运算), M 表示对称群 $\mathbb{G}$ 中的标量乘运算, E_t 表示群 $\mathbb{G}_T$ 中的指数运算, $|S_t^i|$ $(i = 0, 1)$ 表示对时间 t 进行 0/1 编码之后的集合元素个数, $\mathrm{sum}_{S_t^i}$ $(i = 0, 1)$ 表示时间 t 进行 0/1 编码得到的集合所有元素的二进制比特长度之和, $|\mathbb{G}_i|$ $(i = 1, 2)$ 和 $|\mathbb{G}|$ 分别表示群 $\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}$ 中元素的二进制比特长度, n 表示明文的二进制比特长度. 算法输入的时间均使用 64 bit 的毫秒级时

表 2　计算开销对比.

Table 2　Comparison of computational costs.

	KegGen	TdGen	Encrypt	Decrypt	Test
[7]	$H+2M$	$\lvert S_t^1\rvert \cdot 3M$	$2\mathcal{P}+2E_t+M$	$2\mathcal{P}+M$	$4\mathcal{P}+2M$
[11]	$2M_2$	0	$4M_1+2\mathcal{P}_1+2E_t+2M_2+2H_2$	$\mathcal{P}_1$	$4\mathcal{P}_1+2H_2$
SM9-FSET	$2M_2$	$\lvert S_t^1\rvert \cdot (H_2+2M_2+M_1)$	$\lvert S_t^0\rvert \cdot (M_2+H_2)+2\mathcal{P}_1+E_t+4M_1$	$2\mathcal{P}_1+4M_1+E_t$	$4\mathcal{P}_1+2E_t$

表 3　存储开销和安全性对比.

Table 3　Comparison of storage costs and security.

	Private key size	Trapdoor size	Ciphertext size	Security
[7]	$2\lvert\mathbb{G}\rvert$	$2\lvert S_t^1\rvert \cdot \lvert\mathbb{G}\rvert+\mathrm{sum}_{S_t^1}$	$\lvert\mathbb{Z}_p\rvert+\lvert\mathbb{G}\rvert+\mathrm{sum}_{S_t^0}+3n+64$	OW-ID-CCA
[11]	$2\lvert\mathbb{G}_2\rvert$	$\lvert\mathbb{G}_2\rvert$	$3\lvert\mathbb{G}_1\rvert+\lvert\mathbb{G}_2\rvert+2n$	IND-ID-CCA & OW-ID-CCA
SM9-FSET	$2\lvert\mathbb{G}_2\rvert$	$\lvert S_t^1\rvert \cdot (\lvert\mathbb{G}_1\rvert+\lvert\mathbb{G}_2\rvert)+\mathrm{sum}_{S_t^1}$	$3\lvert\mathbb{Z}_p\rvert+2\lvert\mathbb{G}_1\rvert+\lvert S_t^0\rvert \cdot \lvert\mathbb{G}_2\rvert$ $+\mathrm{sum}_{S_t^0}+n$	IND-ID-CCA & OW-ID-CCA & forward security

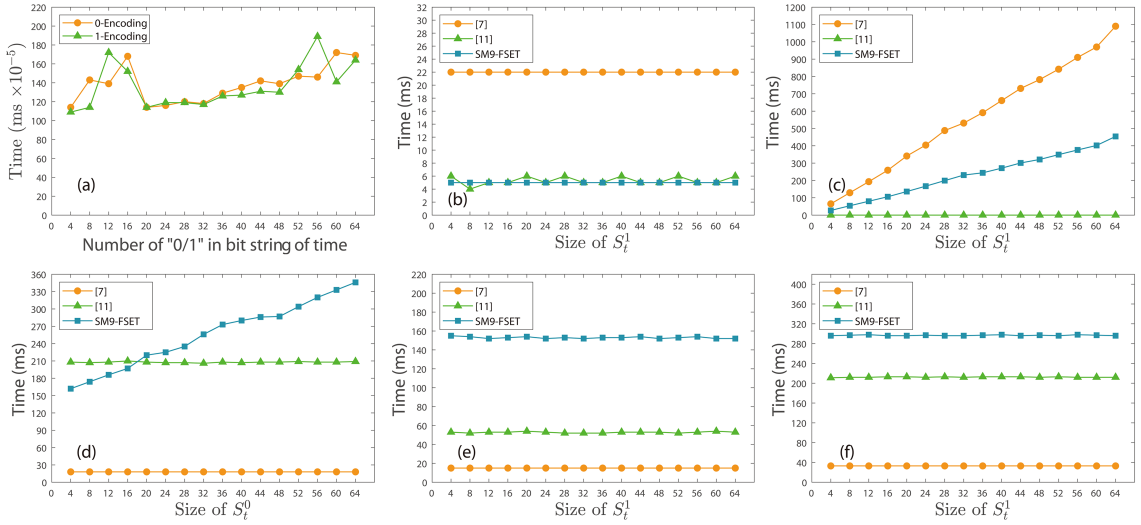

图 1　(网络版彩图) 时间比特串在包含不同 0/1 个数下的 0/1 编码时间消耗和相关方案算法在不同时间编码集合大小下的运行时间. (a) 0/1-Encoding; (b) KeyGen; (c) Encrypt; (d) TdGen; (e) Decrypt; (f) Test.

Figure 1　(Color online) Running time of 0/1-Encoding under different number of 0/1 in bit string of time and related algorithms under different size of S_t^0 or S_t^1. (a) 0/1-Encoding; (b) KeyGen; (c) Encrypt; (d) TdGen; (e) Decrypt; (f) Test.

间戳. 在实验仿真阶段所使用的环境下, 对算法中涉及的 0/1 编码技术的时间消耗进行了测试 (运行十万次取平均时间作为结果). 测试方法是通过获取时间的二进制比特串, 然后从右向左遍历每一位, 利用字符串截取操作获取对应的编码集合. 结果如图 1(a) 所示, 其中 0-Encoding 和 1-Encoding 分别表示在时间二进制串中不同数量的 0/1 进行 0/1 编码的时间消耗. 图中横坐标表示时间戳二进制串中 0 (对 0-Encoding) 或 1 (对 1-Encoding) 的个数, 纵坐标表示 0/1 编码耗时 (单位: ms $\times 10^{-5}$), 从图中可知 0/1 编码的耗时约为 0.1×10^{-5}~0.2×10^{-5} s, 可以忽略不计.

表 2 显示, SM9-FSET 的密钥生成、解密和等值测试算法的时间消耗保持在固定水平, 密钥生成效率与文献 [11] 相同且优于文献 [7], 等值测试算法的效率与文献 [7,11] 相近. SM9-FSET 的陷门生成、加密和解密算法的计算效率低于文献 [7,11], 主要因为文献 [7] 所提方案未能有效防止好奇的测试者对指定时间之外的密文执行等值测试, 存在隐私泄露风险; 文献 [11] 中的方案未实现密文等值测试的前向安全性, 且一个陷门只能匹配一个密文, 效率较低. 表 3 显示, SM9-FSET 的私钥大小与文献 [7,11]

相同, 陷门大小与文献 [7] 相近. 尽管 SM9-FSET 的陷门和密文大小略大于文献 [11], 但 SM9-FSET 实现了密文等值测试的前向安全性, 在安全性和隐私保护方面具有更强的优势. 同时, SM9-FSET 与文献 [11] 均具备 IND-ID-CCA 和 OW-ID-CCA 的安全性保障, 而文献 [7] 中所提方案由于未考虑敌手通过任意临时陷门进行等值测试而攻破方案, 因此并未实现 IND-ID-CCA 的安全性. 这一关键缺陷使得文献 [7] 的方案在实际应用中面临严重的安全威胁.

5.2　实验仿真

本小节从实验方面对本文提出的 SM9-FSET 方案与文献 [7] (支持 Tem-Tem 等值测试) 和文献 [11] 中方案的各算法运行时间进行了对比分析. 实验环境为 64 位 Windows 11 操作系统, 硬件平台为 8 核 AMD Ryzen i7-5800H @ 3.20 GHz 处理器, 配备 16 GB 内存. 每个算法运行 1000 次, 取平均运行时间作为结果. 实验基于 JPBC (Java pairing-based cryptography) 密码学库进行实现, 在 80-bit 安全级别下, 使用 Type A 实现文献 [7] 中的方案, 使用 Type F 实现 SM9-FSET 与文献 [11] 中的方案. 实验结果如图 1(b)∼(f) 所示. 图中横坐标表示时间编码集合大小 $|S_t^i|$ $(i = [0, 1])$, 时间为 64 位毫秒级时间戳, 因此 $|S_t^i|$ 的最大值为 64. 纵坐标表示算法千次运行平均耗时 (单位 ms).

图 1(b) 表明, SM9-FSET 与文献 [7,11] 中方案的密钥生成的时间均不受时间编码集合大小影响. 文献 [7] 中的私钥生成过程需要执行哈希到群操作, 这一操作相对耗时; 而 SM9-FSET 与文献 [11] 的私钥生成过程无需执行哈希到群操作, 耗时约为 5 ms, 显著优于文献 [7] 的 22 ms. 由于在陷门中嵌入时间属性, SM9-FSET 和文献 [7] 的陷门生成算法时间消耗如图 1(c) 所示均与时间 1 编码集合的大小呈线性关系, 但 SM9-FSET 展现出更优的效率. 即使在时间 1 编码集合的大小达到最大值 64 时, SM9-FSET 的陷门生成算法时间消耗约为 454 ms, 远优于文献 [7] 的 1090 ms. 此外, 文献 [11] 未实现密文等值测试的前向安全性且其陷门是私钥的一部分, 因此陷门生成算法的时间消耗为 0.

为实现密文等值测试的前向安全性, SM9-FSET 在加密过程中需嵌入时间编码集合中的每个元素, 导致其加密耗时如图 1(d) 所示与时间 0 编码集合的大小呈线性关系. 然而, 即使 0 编码集合大小达到最大值 64 时, SM9-FSET 的加密耗时约为 346 ms, 仍处于实际应用可接受的范围. 相比之下, 文献 [7,11] 因未实现密文等值测试的前向安全性, 其加密耗时均与时间编码集合的大小无关. 等值测试算法通常将密文分为解密和等值测试两部分. 因此, 解密算法的时间消耗通常不涉及密文中的等值测试部分, 这使得 SM9-FSET 和文献 [7] 的解密时间消耗均与时间编码集合大小无关. 结果如图 1(e) 所示, SM9-FSET 的解密时间消耗约为 154 ms, 高于文献 [7] 的 15 ms, 但仍处于可接受范围内. SM9-FSET 的等值测试算法虽然涉及时间编码处理过程, 但实验表明, 该过程的时间开销几乎可以忽略不计 (如图 1(a) 所示), 因此等值测试的整体时间消耗基本保持在固定水平, 与时间编码集合的大小无关. 如图 1(f) 所示, SM9-FSET 的等值测试耗时约为 299 ms, 高于文献 [7] 的 33 ms 和文献 [11] 的 212 ms, 但在提升安全性的前提下, 这一时间开销仍处于可接受范围.

值得注意的是, 图 1(d)∼(f) 显示, SM9-FSET 加密、解密和等值测试算法的执行效率劣于文献 [7,11], 其原因在于文献 [7,11] 未能防御好奇的测试者对任意密文进行等值测试的攻击, 而本文所提方案有效防止了好奇的测试者对指定时间之外的密文进行等值测试, 具备更高的安全性, 更适用于高安全需求的场景.

6　结论

基于我国商用密码 SM9 无法实现密文等值测试算法的前向安全问题, 本文提出了一种前向安全的 SM9 密文等值测试方案 SM9-FSET. 方案实现了加密者可以在密文中嵌入时间属性, 当且仅当陷门中嵌入的时间大于密文中嵌入的时间时, 密文才可进行等值测试. 提高了密文等值测试算法的安全性和灵活性, 有效推广了 SM9 的发展. 并在随机谕言机模型中证明了 SM9-FSET 满足 IND-ID-CCA

和 OW-ID-CCA 的安全性. 最后, 将 SM9-FSET 与相关方案 [7,11] 在理论和实验两方面进行比较, 结果表明 SM9-FSET 方案在计算效率、通信效率和安全性方面与相关方案 [7,11] 是可比的且更具实用性.

参考文献

1 Shamir A. Identity-based cryptosystems and signature schemes. In: Proceedings of CRYPTO, Santa Barbara, 1984. 47–53

2 Boneh D, Franklin M K. Identity-based encryption from the weil pairing. In: Proceedings of the 21st Annual International Cryptology Conference, Santa Barbara, 2001. 213–229

3 Sun Y X, Chatterjee P, Chen Y, et al. Efficient identity-based encryption with revocation for data privacy in Internet of Things. IEEE Int Things J, 2022, 9: 2734–2743

4 Chiku S, Hara K, Shikata J. Chosen ciphertext security for (hierarchical) identity-based matchmaking encryption. IACR Cryptol ePrint Arch, 2024. https://eprint.iacr.org/2024/1292.pdf

5 Song D X, Wagner D A, Perrig A. Practical techniques for searches on encrypted data. In: Proceeding IEEE Symposium on Security and Privacy, 2000. 44–55

6 Yang G M, Tan C H, Huang Q, et al. Probabilistic public key encryption with equality test. In: Proceedings of the Cryptographers' Track at the RSA Conference, San Francisco, 2010. 119–131

7 Ma S, Ye Z Q, Huang Q, et al. Controllable forward secure identity-based encryption with equality test in privacy-preserving text similarity analysis. Inf Sci, 2024, 660: 120099

8 Lin H-Y, Tzeng W-G. An efficient solution to the millionaires' problem based on homomorphic encryption. In: Proceedings of the 3rd International Conference on Applied Cryptography and Network Security, New York, 2005. 456–466

9 GM/T0044-2016. Identity-based cryptographic algorithm SM9. 2016 [GM/T0044-2016. SM9 标识密码算法. 2016] https://www.math.pku.edu.cn/teachers/qiuzy/qfiles/SM9.pdf

10 Pu L, Lin C, Wu W, et al. A public-key encryption with keyword search scheme from SM9. J Cyber Secur, 2023, 8: 108–118 [蒲浪, 林超, 伍玮, 等. 基于 SM9 的公钥可搜索加密方案. 信息安全学报, 2023, 8: 108–118]

11 Xiong H, Lin Y, Yao T. SM9 identity-based encryption scheme with equality test and cryptographic reverse firewalls. J Comput Res Dev, 2024, 61: 1070–1084 [熊虎, 林烨, 姚婷. 支持等式测试及密码逆向防火墙的 SM9 标识加密方案. 计算机研究与发展, 2024, 61: 1070–1084]

12 Boneh D, Boyen X. Efficient selective-ID secure identity-based encryption without random oracles. In: Proceedings of International Conference on the Theory and Applications of Cryptographic Techniques, Interlaken. 2004. 223–238

13 Gentry C. Practical identity-based encryption without random oracles. In: Proceedings of the 25th Annual International Conference on the Theory and Applications of Cryptographic Techniques, St. Petersburg, 2006. 445–464

14 Tiplea F L, Iftene S, Teseleanu G, et al. Security of identity-based encryption schemes from quadratic residues. IACR Cryptol ePrint Arch, 2016. https://eprint.iacr.org/2022/012.pdf

15 Mitrokotsa K, Mukherjee S, Tomy J. Oblivious identity-based encryption (IBE secure against an adversarial KGC). IACR Cryptol ePrint Arch, 2023. https://eprint.iacr.org/2023/825.pdf

16 Lu Y, Zhang R, Lin D D. Stronger security model for public-key encryption with equality test. In: Proceedings of the 5th International Conference on Pairing-Based Cryptography, Cologne. 2012. 65–82

17 Tang Q. Towards public key encryption scheme supporting equality test with fine-grained authorization. In: Proceedings of the 16th Australasian Conference on Information Security and Privacy, Melbourne. 2011. 389–406

18 Ma S, Huang Q, Zhang M G, et al. Efficient public key encryption with equality test supporting flexible authorization. IEEE Trans Inform Forensic Secur, 2015, 10: 458–470

19 Ma S. Identity-based encryption with outsourced equality test in cloud computing. Inf Sci, 2016, 328: 389–402

20 Nguyen G L D, Duong D H, Le H Q, et al. Lattice-based public key encryption with multi-ciphertexts equality test in cloud computing. IACR Cryptol ePrint Arch, 2022. https://eprint.iacr.org/2022/199.pdf

21 Ma S, Zhong Y J, Huang Q. Efficient public key encryption with outsourced equality test for cloud-based IoT environments. IEEE Trans Inform Forensic Secur, 2022, 17: 3758–3772

22 Günther C G. An identity-based key-exchange protocol. In: Proceedings of Workshop on the Theory and Application of Cryptographic Techniques, Houthalen. 1989. 29–37

23 Diffie W, van Oorschot P C, Wiener M J. Authentication and authenticated key exchanges. Des Codes Crypt, 1992, 2: 107–125

24 Canetti R, Halevi S, Katz J. A forward-secure public-key encryption scheme. In: Proceedings of International Conference on the Theory and Applications of Cryptographic Techniques, Warsaw. 2003. 255–271

25 Zhang X J, Xu C X, Wang H X, et al. FS-PEKS: lattice-based forward secure public-key encryption with keyword search for cloud-assisted industrial Internet of Things. IEEE Trans Dependable Secure Comput, 2021, 18: 1019–1032
26 Agrawal S, Boneh D, Boyen X. Lattice basis delegation in fixed dimension and shorter-ciphertext hierarchical IBE. In: Proceedings of the 30th Annual Cryptology Conference, Santa Barbara. 2010. 98–115
27 Yang X M, Chen X J, Huang J Y, et al. FS-IBEKS: forward secure identity-based encryption with keyword search from lattice. Comput Standards Interfaces, 2023, 86: 103732
28 Yang Y T, Cai J L, Zhang X W, et al. Privacy preserving scheme in block chain with provably secure based on SM9 algorithm. J Softw, 2019, 30: 1692–1704 [杨亚涛, 蔡居良, 张筱薇, 等. 基于 SM9 算法可证明安全的区块链隐私保护方案. 软件学报, 2019, 30: 1692–1704]
29 Gan Z W, Liao F Y. Rapid calculation of R-ate bilinear pairing in China state cryptography standard SM9. Comput Eng, 2019, 45: 171–174 [甘植旺, 廖方圆. 国密 SM9 中 R-ate 双线性对快速计算. 计算机工程, 2019, 45: 171–174]
30 Lai J C, Huang X Y, He D B, et al. Security analysis of SM9 digital signature and key encapsulation. Sci Sin Inform, 2021, 51: 1900–1913 [赖建昌, 黄欣沂, 何德彪, 等. 国密 SM9 数字签名和密钥封装算法的安全性分析. 中国科学: 信息科学, 2021, 51: 1900–1913]
31 Lai J C, Huang X Y, He D B, et al. An efficient identity-based signcryption scheme based on SM9. J Cryptol Res, 2021, 8: 314–329 [赖建昌, 黄欣沂, 何德彪, 等. 基于商密 SM9 的高效标识签密. 密码学报, 2021, 8: 314–329]
32 Lai J C, Huang X Y, He D B. An efficient identity-based broadcast encryption scheme based on SM9. Chinese J Comput, 2021, 44: 897–907 [赖建昌, 黄欣沂, 何德彪. 一种基于商密 SM9 的高效标识广播加密方案. 计算机学报, 2021, 44: 897–907]
33 Zhang C, Peng C G, Ding H F, et al. Efficient identity-based encryption with revocation for data privacy in Internet of Things. Comput Eng, 2022, 48: 159–167 [张超, 彭长根, 丁红发, 等. 基于国密 SM9 的可搜索加密方案. 计算机工程, 2022, 48: 159–167]
34 Lai J C, Huang X Y, He D B, et al. Efficient hierarchical identity-based encryption based on SM9. Sci Sin Inform, 2023, 53: 918–930 [赖建昌, 黄欣沂, 何德彪, 等. 基于商用密码 SM9 的高效分层标识加密. 中国科学: 信息科学, 2023, 53: 918–930]
35 Li C, Liang J K, Ding Y J, et al. Hierarchical identity-based broadcast inner product functional encryption based on SM9. Sci Sin Inform, 2024, 54: 1400–1418 [李聪, 梁俊凯, 丁煜甲, 等. 基于 SM9 的分层标识广播内积函数加密. 中国科学: 信息科学, 2024, 54: 1400–1418]

Forward secure equality test based on SM9

Jianlong ZHANG[1], Jianchang LAI[1*], Xinyi HUANG[2], Liquan CHEN[1] & Jinguang HAN[1]

1. *School of Cyber Science and Engineering, Southeast University, Nanjing 211189, China*
2. *College of Cyber Security, Jinan University, Guangzhou 510632, China*
* Corresponding author. E-mail: jclai@seu.edu.cn

Abstract Equality test on ciphertext can effectively compare whether two ciphertexts are encrypted under the same plaintext without decryption, which plays a significant role in database partitioning, encrypted data statistics, deduplication storage, and so on. The SM9 identity-based cryptosystem has become a commercial and national cryptographic standard in China and is widely applied in many fields such as finance and government. However, the existing SM9-based equality test schemes cannot achieve forward security, and the ciphertexts generated at any time can be tested through the previous trapdoor. It fails to meet the forward security requirements for special applications. It therefore becomes a bottleneck for the deployment of SM9. Forward security in this paper refers to only the ciphertexts generated before a specified time embedded in the trapdoor can be tested, while those generated at or after that time cannot be successfully tested. This paper proposes a forward secure equality test scheme based on SM9 (SM9-FSET). Only the ciphertexts that match the time constraints can perform equality tests effectively, thereby enhancing data privacy. We prove that if the q-BDHI assumption holds, our scheme is proven to be IND-ID-CCA and OW-ID-CCA secure in the random oracle model. Finally, we compare our SM9-FSET scheme to the related schemes. The results demonstrate that SM9-FSET is both comparable and more practical to the related schemes in terms of computational cost, communication overhead, and security.

Keywords SM9, equality test, forward security, chosen-ciphertext attack (CCA)