基于区块链的多用户环境中公钥可搜索加密方案

郑东，朱天泽，郭瑞
（西安邮电大学网络空间安全学院，陕西 西安 710121）

摘 要：为了满足多用户环境中数据安全共享的需求，提出了一种支持一对多模式的公钥可搜索加密方案。具体地，数据拥有者执行一次加密算法可以指定多位用户对密文进行检索，实现更加灵活的密文数据共享。此外，还设计了具体的文件加密密钥传递算法，确保用户在检索到密文后能够正确解密并获取明文。结合区块链技术，利用智能合约执行检测算法保证了检索结果的正确性。在安全性方面，基于判定性双线性 Diffie-Hellman 假设和修改的判定性双线性 Diffie-Hellman 假设，证明了在随机谕言机模型下所提方案满足密文关键词不可区分性和陷门信息不可区分性的安全要求，并且可以抵御内部关键词猜测攻击。最后，使用 jPBC 密码库对所提方案与现有相关方案进行仿真模拟，测试结果表明所提方案具有较高的计算效率。

关键词：可搜索加密；多用户环境；密文安全共享；区块链；智能合约
中图分类号：TP309
文献标识码：A
DOI: 10.11959/j.issn.1000−436x.2021130

Public key searchable encryption scheme in blockchain-enabled multi-user environment

ZHENG Dong, ZHU Tianze, GUO Rui
College of Cyberspace Security, Xi’an University of Posts and Telecommunications, Xi’an 710121, China

Abstract: In order to meet the needs of data security sharing in multi-user environment, a public key searchable encryption scheme supporting one-to-many mode was proposed. Specifically, the data owner could specify multiple users to retrieve the ciphertext by executing the encryption algorithm once, so as to achieve more flexible sharing of ciphertext data. In addition, the specific file encryption key transfer algorithm was designed to ensure that the user could decrypt and obtain the plaintext correctly after retrieving the ciphertext. Combined with the blockchain technology, the smart contract execution detection algorithm was used to ensure the correctness of the retrieval results. In terms of security, based on the decisional bilinear Diffie-Hellman hypothesis and the modified bilinear Diffie-Hellman hypothesis, it was proved that the proposed scheme satisfies the security requirements of keyword indistinguishability and trapdoor information indistinguishability under the random oracle model, and could resist the internal keyword guessing attack. Finally, the proposed scheme and the existing related schemes were simulated by using jPBC cryptolibrary, and the test results show that the proposed scheme has high computational efficiency.

Keywords: searchable encryption, multi-user environment, ciphertext secure sharing, blockchain, smart contract

收稿日期：2021−02−24；**修回日期**：2021−05−18
基金项目：国家重点研发计划基金资助项目（No.2017YFB0802000）；国家自然科学基金资助项目（No.62072369，No.62072371，No.61802303，No.61772418）；陕西省重点研发计划基金资助项目（No.2020ZDLGY08-04，No.2019KW-053）；陕西省创新能力支持计划基金资助项目（No.2020KJXX-052，No.2017KJXX-47）；陕西省自然科学基金资助项目（No.2019JQ-866，No.2018JZ6001）；陕西省教育厅科研基金资助项目（No.19JK0803）；青海省基础研究计划基金资助项目（No.2020-ZJ-701）
Foundation Items: The National Key Research and Development Program of China (No.2017YFB0802000), The National Natural Science Foundation of China (No.62072369, No.62072371, No.61802303, No.61772418), The Key Research and Development Program of Shaanxi Province (No.2020ZDLGY08-04, No.2019KW-053), The Innovation Capability Support Plan of Shaanxi Province (No.2020KJXX-052, No.2017KJXX-47), The Natural Science Foundation of Shaanxi Province (No.2019JQ-866, No.2018JZ6001), The Scientific Research Project of Shaanxi Provincial Department of Education (No.19JK0803), The Basic Research Program of Qinghai Province (No.2020-ZJ-701)

1　引言

在云计算时代，人们可以借助云存储服务外包存储数据以降低本地信息管理开销。然而，云存储技术在为人们带来便利的数据服务的同时，也存在一定的安全隐患。云服务器中存储的数据包含大量的用户敏感信息，如私密电子邮件、个人健康记录以及财务信息等。这些敏感数据一经上传至云端，便完全脱离了用户的控制，其安全性受到了极大威胁。如何为用户敏感信息提供安全保护，已成为云计算亟须解决的关键问题之一。考虑到作为存储载体的半可信云服务器可以轻松绕过访问控制策略查看用户的数据，因此必须将数据加密后再存储到云服务器上才能真正实现安全存储。对于明文信息，可以使用传统的关键词搜索技术来找到所需的数据。然而，这种方式并不适用于对密态数据的检索。如何实现对密态数据的快速搜索便于用户对所需数据进行准确定位，是云存储技术中的研究热点。

可搜索加密技术可以通过使用特定的关键词来检索加密后的文件，实现密态数据的检索功能，为云平台中存储的敏感数据提供安全性保护。Song 等[1]给出了第一个对称可搜索加密算法。Boneh 等[2]构造了第一个公钥可搜索加密（PEKS, public key encryption with keyword search）方案。如图 1 所示，PEKS 包含以下 3 个步骤：1) 数据拥有者利用自己的明文数据提取明文关键词集，使用数据接收者的公钥生成密文关键词集和密文数据，将密文关键词与密文数据建立索引后，连同密文数据外包存储在服务器上；2) 当数据接收者想对存储在云服务器上的数据进行搜索时，利用自己的私钥生成陷门信息，并将陷门信息发送给云服务器；3) 云服务器接收到陷门信息后进行相关搜索工作，在此过程中不进行数据解密。将搜索到的相关密文数据发送给数据接收者后，数据接收者利用自己的私钥对文档进行解密。然而，在传统公钥可搜索加密体系中服务器是半可信且好奇的。为了节约计算资源，即使服务器通过检测算法检索到了用户所需的文件，也可能会返回错误的或部分检索结果。

目前，解决可搜索加密中第三方的半可信问题，多采用可搜索加密与区块链技术结合的方式。区块链本质上是一种去中心化、分布式存储的数据库，链上存储的数据公开透明，并依托密码技术确保存储在链上的数据不可篡改、可追溯。智能合约是一套以数字形式定义的承诺，合约参与方可以在上面执行所承诺的协议，并且，智能合约允许在没有第三方的情况下进行可信操作，执行的操作可追踪且不可逆转。结合区块链上数据的不可篡改性，利用智能合约将存储在链上的密文关键词与接收到的陷门信息进行匹配，可以保证文件检索结果的正确性且可以帮助用户验证服务器返回的文件是否被篡改。

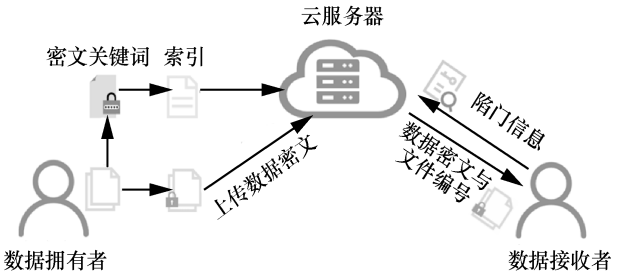

图 1　公钥可搜索加密

1.1　相关工作

为了在云服务器上对密文数据进行检索，Song 等[1]给出了第一个可搜索加密方案。该方案基于对称密码体制，将明文数据中每个单词进行加密后上传，用户想要搜索时，利用关键词生成密文发送给服务器，服务器通过扫描密文数据并与关键词密文进行对比来返回检索结果。然而，这种搜索方式需要遍历全部密文，计算代价大且效率较低。Boneh 等[2]首次将公钥密码体制引入可搜索加密领域，构造了第一个 PEKS。该方案使用索引的结构来访问隐私数据，服务器将数据接收者发送的陷门与密文关键词进行匹配，匹配成功后利用索引返回对应的密文。同时，密文关键词中包含数据接收者的公钥使整个检索过程中用户之间不需要交互，公钥可搜索加密的应用场景也因此更广阔。然而，该方案需在安全信道中进行陷门传递，限制了其应用范围。针对此问题，Baek 等[3]首次提出了可以在公开信道下传输数据的公钥可搜索加密方案（SCF-PEKS）。通过在密文关键词中加入指定服务器的公钥来确保可以在公开信道中传输数据。Fang 等[4]提出了在无随机谕言机模型下证明 PEKS 和 SCF-PEKS 的关键词安全性。此外，Fang 等[5]还设计了一种效率高于 SCF-PEKS 的公钥可搜索加密方案。

然而，上述 PEKS 和 SCF-PEKS 都存在关键词隐私性不足的问题，无法抵抗关键词猜测攻击[6-8]。

为了应对此种攻击，Tang 等[9]设计了可以抵抗离线关键词猜测攻击的方案。Rhee 等[10]首次提出了“陷门不可区分性”的概念，并指出公钥可搜索加密满足陷门不可区分性是对抗离线关键词猜测攻击的充分条件。Qin 等[11]将 Boneh 等提出的密文不可区分性扩展成为了多密文不可区分性。在实际应用中，一份文件通常不仅仅包含一个关键词，而且同一个关键词也可能出现多次。给定一组关键词加密，数据拥有者不希望其他人知道 2 个文件是否包含相同的关键词，或者同一个文件包含多少个相同的关键词。因此，公钥可搜索加密方案需要考虑多密文不可区分性。

将基于身份的加密（IBE, identity-based encryption）算法与可搜索加密算法相结合，可以大大降低 PEKS 方案中的证书管理开销。Abdalla 等[12]完善了 PEKS 的定义，给出了其与基于身份的匿名加密方案之间的转化关系，并提出了一种新的基于临时关键词搜索的公钥可搜索加密方案。Rhee 等[13]构造了指定测试者的基于身份的公钥可搜索加密（dIBEKS, IBEKS with designated tester）的 2 种通用结构。Emura 等[14]给出了自适应安全的公开信道下公钥可搜索加密的通用构造方法，并构造了基于标签和一次性签名的 IBEKS 通用结构[15]。王少辉等[16]首次提出了基于身份密码系统下的指定测试者可搜索加密方案的定义和安全需求，特别证明了 dIBEKS 密文不可区分性是抵御离线关键词猜测攻击的充分条件并设计了一个高效的 dIBEKS 新方案可以有效抵御离线关键词猜测攻击。Ma 等[17]在物联网（IoT, Internet of things）环境中设计了一种新的基于身份的无证书可搜索加密方案，用于 IoT 数据在云存储服务器中安全外包存储与共享。方案证明了可以针对选择一个公钥替换用户公钥和可以被给予系统主密钥这两类敌手做到抵抗选择关键词攻击。牛淑芬等[18]将 PEKS 与 IBE 相结合，在有效解决了邮件通信系统中对加密邮件检索问题的同时，也降低了公钥可搜索加密方案中复杂的密钥管理开销。此外，通过邮件发送方和接收方之间的相互认证并指定服务器执行检测算法以抵御离线关键词猜测攻击。为了优化传统公钥可搜索加密方案中大量运用双线性映射导致的低效率问题，杨宁滨等[19]构建了无双线性映射运算的公钥认证可搜索加密方案来提升运算的效率。

一对一模式的可搜索加密无法满足同时向多位用户进行密文数据安全共享的需求。为了解决此类问题，Curtmola 等[20]结合广播加密技术首次提出了一对多模式的公钥可搜索加密方案但其密钥撤销代价较大。杜瑞忠等[21]提出了一种基于区块链的公钥可搜索加密方案实现了一对多模式的数据共享，结合智能合约来解决传统方案中服务器不完全可信的问题，而且因为陷门信息无需上传至服务器也防止了半可信的服务器进行内部关键词猜测攻击。然而，其陷门信息在公开信道下传输时无法抵御离线关键词猜测攻击；此外，方案中智能合约在完成检索后将加密数据文件的对称密钥直接发送给用户也存在安全隐患。张玉磊等[22]利用代理重加密构造了一对多模式的 PEKS 方案以适应多用户环境下的需求，其方案中需要数据拥有者持续在线以处理来自数据接收者发起的请求，在实际应用中存在较大限制。

在云计算环境下，为了节省计算资源，服务器可能会返回错误的检索结果或部分检索结果，严重影响用户的使用。智能合约与区块链技术的快速发展可以为用户提供可信云服务[23-24]，在公钥可搜索加密中能够帮助用户验证云服务器提供的密态数据检索结果是否准确。Zhang 等[25]提出了一种基于区块链的个人健康数据安全共享方案，为了保证数据安全性和可用性，对加密后的个人体征数据使用公钥可搜索加密技术实现检索。高梦婕等[26]提出了一种基于区块链的可搜索医疗数据共享方案，方案使用对称可搜索加密并结合区块链与秘密共享技术，在确保参与者能安全获取共享密钥的同时也保证了共享数据的一致性。Li 等[27]提出了一种基于区块链的可搜索加密方案，方案使用对称加密算法实现对密文的检索。在该方案中，Li 等不仅给出了基于区块链的对称可搜索加密方案模型，还提出了 2 种不同方案以处理不同规模的数据。之后，Li 等[28]还对文献[25]所提方案进行了改进以提高其可靠性。Chen 等[29]也使用对称可搜索加密并结合区块链技术提出了一种电子医疗记录共享方案，其使用智能合约作为可信第三方用以确保云平台提供的服务可信。牛淑芬等[30]基于联盟链提出了一种可搜索电子病历共享方案。该方案将代理重加密的思想引入可搜索加密中，并与区块链技术相结合，通过使用服务器存储电子病历密文、私有链存储密文哈希值、联盟链存储关键词索引的方式，实现了电子病历的安全存储与共享。

1.2　本文贡献

为了实现在半可信云存储环境下的多用户密文数据安全共享，本文提出了一种基于区块链的公钥可搜索加密方案，主要贡献如下。

1) 将 PEKS 与 IBE 相结合，实现了一对多模式的公钥可搜索加密方案，并设计了文件加密密钥的传递算法，保证用户在检索到密文后能够正确解密并获取明文。在该模型中，数据共享者只需一次加密上传就可以向多位指定的用户共享数据，能够灵活运用于实际场景中。利用 IBE 的优势也降低了证书管理开销。此外，引入区块链与智能合约，将索引存储在区块链上并利用智能合约进行检索操作，在保证返回正确文件检索结果的同时帮助用户进行数据验证，解决了传统第三方存储中存在的半可信问题，保障了数据的可靠性。

2) 在随机谕言机模型下，基于判定性双线性 Diffie-Hellman 假设（DBDH, decisional bilinear Diffie-Hellman）和修改的判定性双线性 Diffie-Hellman 假设（MBDH, modified bilinear Diffie-Hellman），证明了本文方案的密文关键词和陷门信息满足选择关键词攻击下的不可区分性，并可以抵抗内部关键词猜测攻击。同时，分析了区块链对保护数据完整性的作用。

3) 利用 jPBC 密码库，对本文方案和其他相关方案进行模拟实验。其仿真结果表明，本文方案具有较高的运行效率和较强的实用价值。

2　基础知识

2.1　双线性映射

设 G_1、G_2 均是阶数为素数 p 的群，其中 G_1 是加法群，G_2 是乘法群，P 为群 G_1 的生成元。满足如下 3 个性质的映射称为一个双线性映射 $\hat{e}:G_1\times G_1\to G_2$。

1) 双线性：对于 $\forall P,Q\in G_1$，$\forall a,b\in Z_q^*$，$\hat{e}(aP,bQ)=\hat{e}(P,Q)^{ab}$。

2) 非退化性：$\exists P,Q\in G_1$，使 $\hat{e}(P,Q)\neq 1$。

3) 可计算性：对所有的 $P,Q\in G_1$，存在有效的算法 $\hat{e}(P,Q)$。

2.2　判定性双线性 Diffie-Hellman 问题

设 G_1、G_2 均是阶数为素数 p 的群，其中 G_1 是加法群，G_2 是乘法群，P 为群 G_1 的生成元，双线性映射 $\hat{e}:G_1\times G_1\to G_2$。DBDH 问题可表述为给定 $(P,aP,bP,cP)\in G_1$，$E\in G_2$，判断 $E=\hat{e}(g,g)^{abc}$ 还是 $E=\hat{e}(g,g)^{z}$，其中 $a,b,c,z\leftarrow_R Z_p^*$。

2.3　修改的判定性双线性 Diffie-Hellman 问题

设 G_1、G_2 均是阶数为素数 p 的群，其中 G_1 是加法群，G_2 是乘法群，P 为群 G_1 的生成元，双线性映射 $\hat{e}:G_1\times G_1\to G_2$。MBDH 问题可表述为给定 $(P,aP,bP,cP)\in G_1$，$E\in G_2$，判断 $E=\hat{e}(g,g)^{\frac{ab}{c}}$ 还是 $E=\hat{e}(g,g)^{z}$，其中 $a,b,c,z\leftarrow_R Z_p^*$。

3　方案模型

3.1　系统模型

如图 2 所示，本文方案主要包括密钥生成中心（PKG, private key generator）、用户（user）、联盟链与智能合约、云服务器（CS, cloud sever）4 个实体。

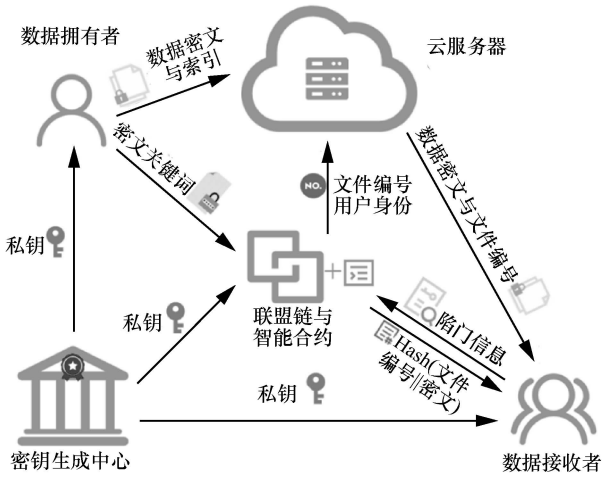

图 2　基于区块链的多用户环境中公钥可搜索加密方案系统框架

1) 密钥生成中心。PKG 为每个用户和管理联盟链的权威中心（以下简称权威中心）生成私钥并公布系统公共参数。

2) 用户。根据系统内用户的不同行为，每个用户既可以是数据拥有者也可以是数据接收者。

① 数据拥有者是系统中向其他用户进行数据共享的用户。数据拥有者将要分享的数据密文编号并建立索引，然后将密文数据与索引一起上传至 CS。同时，生成密文关键词存储在联盟链上供智能合约执行检测算法时调用，密文关键词中包含文件编号与文件哈希值。

② 数据接收者是有检索需求的用户。数据接收者利用自己私钥和想要搜索的关键词生成陷门

信息发送给智能合约，利用返回值进行文件正确性与完整性验证并解密密文数据。

3) 联盟链与智能合约。本文方案采用区块链应用模式中的联盟链进行构造。联盟链在保有区块链基本特性的基础上增加了认证机制，只有经过授权的用户才能参与其中，符合基于身份的加密系统。系统内权威中心对联盟链进行管理并部署智能合约。联盟链存储数据拥有者上传的密文关键词，智能合约收到陷门信息时调用数据拥有者存储在联盟链上的密文关键词并执行检测算法，将数据接收者的身份和检索到的文件编号发送给 CS 并为用户返回哈希值，用来进行文件正确性与完整性验证。

4) 云服务器。外包存储数据拥有者上传的数据密文与索引。

3.2 形式化定义

本文方案主要包括以下步骤。

步骤 1 系统初始化算法 SetUp(λ)。该算法以安全参数 λ 为输入。PKG 生成公开参数 params，保留系统私钥 msk。权威中心将智能合约部署在联盟链上。

步骤 2 密钥生成算法 KeyGen (params, $\mathrm{ID_U}$, $\mathrm{ID_A}$, msk)。该算法以系统公共参数 params、用户身份 $\mathrm{ID_U}$、权威中心标识 $\mathrm{ID_A}$、系统私钥 msk 为输入。PKG 为用户计算生成 $\mathrm{sk_U}$，用户自己计算生成 D_{U}，将 $\mathrm{sk_U}$、D_{U} 作为自己的私钥；PKG 为权威中心生成 $\mathrm{sk_A}$，权威中心将 $\mathrm{sk_A}$ 作为私钥。PKG 计算生成用户身份公钥 $T_{\mathrm{ID_U}}$。

步骤 3 密文关键词生成算法 KeywordGen (params, w, M)。该算法以系统公共参数 params、一个明文关键词 w、明文数据 M 为输入。用户生成一个密文关键词 C_{kw}、一个索引 I 和一个数据密文 C_m。

步骤 4 陷门信息生成算法 TrapdoorGen (params, w', T_i, D_{U}, $\mathrm{ID_A}$)。该算法以系统公共参数 params、用户想要搜索的关键词 w'、T_i、用户私钥 D_{U} 以及权威中心身份标识 $\mathrm{ID_A}$ 为输入，用户生成陷门信息 T_w。

步骤 5 检测算法 Test(params, C_{kw}, T_w, $\mathrm{sk_A}$)。该算法以系统公共参数 params、密文关键词 C_{kw}、陷门信息 T_w 和权威中心私钥 $\mathrm{sk_A}$ 为输入，返回检索结果。

步骤 6 数据密文解密算法 Dec (N, n', C'_m)。该算法以密文关键词 C_{kw} 中的 N 以及服务器返回的密文关键词 n' 和数据密文 C'_m 为输入，用户验证数据完整性。若验证通过则解密 C'_m，此时 $C_m = C'_m$。

3.3 安全模型

定义 1 如果对任何多项式时间敌手 $\mathcal{A}$，存在一个可忽略的函数 $\varepsilon(K)$，K 为安全参数，使 $\mathrm{Adv}_{\Pi\mathcal{A}}^{\mathrm{CKA}}(K) \leqslant \varepsilon(K)$，那么就称这个加密算法 Π 是语义安全的，或称为在选择关键词攻击下具有不可区分性，简称为 IND-CKA（indistinguishability-chosen keyword attack）安全。

定义 2 如果对任何多项式时间敌手 $\mathcal{A}$，存在一个可忽略的函数 $\varepsilon(K)$，使 $\mathrm{Adv}_{\Pi\mathcal{A}}^{\mathrm{CPA}}(K) \leqslant \varepsilon(K)$，那么就称这个加密算法 Π 是语义安全的，或称为在选择明文攻击下具有不可区分性，简称为 IND-CPA（indistinguishability-chosen plaintext attack）安全。

定义 3 设 G_1、G_2 是阶为大素数 p 的群，g 为 G_1 的生成元，映射 $\hat{e}: G_1 \times G_1 \to G_2$，$a,b,c,z \leftarrow_{\mathrm{R}} Z_p^*$，则随机五元组 $R = (g, g^a, g^b, g^c, \hat{e}(g,g)^z)$ 与 DBDH 五元组 $D = (g, g^a, g^b, g^c, \hat{e}(g,g)^{abc})$ 是计算上不可区分的，称为 DBDH 假设。

具体地，对任一敌手 $\mathcal{B}$，$\mathcal{B}$ 区分 R 和 D 的优势 $\mathrm{Adv}_{\mathcal{B}}^{\mathrm{DBDH}} = |\Pr[\mathcal{B}(R)=1] - \Pr[\mathcal{B}(D)=1]|$ 是可忽略的，即 $\mathrm{Adv}_{\mathcal{B}}^{\mathrm{DBDH}} \leqslant \varepsilon(K)$。

定义 4 设 G_1、G_2 是阶为大素数 p 的群，g 为 G_1 的生成元，映射 $\hat{e}: G_1 \times G_1 \to G_2$，$a,b,c,z \leftarrow_{\mathrm{R}} Z_p^*$，则随机五元组 $R = (g, g^a, g^b, g^c, \hat{e}(g,g)^z)$ 与 MBDH 五元组 $M = (g, g^a, g^b, g^c, \hat{e}(g,g)^{\frac{ab}{c}})$ 是计算上不可区分的，称为 MBDH 假设。

具体地，对任一敌手 $\mathcal{B}$，$\mathcal{B}$ 区分 R 和 M 的优势 $\mathrm{Adv}_{\mathcal{B}}^{\mathrm{MBDH}} = |\Pr[\mathcal{B}(R)=1] - \Pr[\mathcal{B}(M)=1]|$ 是可忽略的，即 $\mathrm{Adv}_{\mathcal{B}}^{\mathrm{MBDH}} \leqslant \varepsilon(K)$。

3.3.1 密文关键词的不可区分性

下面，通过敌手 $\mathcal{A}$ 与挑战者之间的安全游戏来定义本文方案的密文关键词不可区分性。密文关键词 C_{kw} 的 IND-CKA 游戏定义如下。

初始化 挑战者输入安全参数 K，产生公开的系统参数 params 和保密的主密钥。

阶段 1（训练） 敌手 $\mathcal{A}$ 发出对 ID 的密钥产生询问并声明意欲挑战的身份 ID^*，否则记为 $\mathrm{ID_U}$。挑战者运行密钥生成算法，根据是否为要挑战的 ID

返回不同的身份公钥T与私钥给敌手$\mathcal{A}$，这一过程可重复多项式有界次。

挑战　敌手$\mathcal{A}$输出 2 个长度相等的关键词w_0、w_1和一个意欲挑战的公开钥ID^*。唯一的限制是ID^*在阶段 1 中的任何密钥询问中出现时挑战者报错并退出。挑战者随机选择一个比特值$\beta \leftarrow_{\mathrm{R}} \{0,1\}$，用$\mathrm{ID}^*$加密$w_\beta$得到$C_{\mathrm{kw}}^*$，并将$C_{\mathrm{kw}}^*$发送给敌手$\mathcal{A}$。

阶段 2（训练）　敌手$\mathcal{A}$发出对另外 ID 的密钥产生询问，唯一的限制是$\mathrm{ID} \neq \mathrm{ID}^*$，挑战者以阶段 1 中的方式进行回应，这一过程可重复多项式有界次。

猜测　敌手$\mathcal{A}$输出猜测$\beta' \in \{0,1\}$，如果$\beta' = \beta$，则敌手$\mathcal{A}$攻击成功。

敌手$\mathcal{A}$的优势定义为安全参数K的函数

$$\mathrm{Adv}_{C_{\mathrm{kw}},\mathcal{A}}^{\mathrm{CKA}}(K) = |\Pr[\mathrm{Succ}] - \frac{1}{2}|$$

3.3.2　陷门信息的不可区分性

下面，通过敌手$\mathcal{A}$与挑战者之间的安全游戏来定义本文方案的陷门信息不可区分性。陷门信息T_{w}的 IND-CKA 游戏定义如下。

初始化　挑战者输入安全参数K，产生公开的系统参数 params 和保密的主密钥。

阶段 1（训练）　敌手$\mathcal{A}$发出对 ID 的密钥产生询问并声明意欲挑战的身份ID^*，否则记为ID_i。挑战者运行密钥生成算法，根据是否为要挑战的 ID 用不同的方式返回$H_1(\mathrm{ID})$与私钥给敌手$\mathcal{A}$，这一过程可重复多项式有界次。

挑战　敌手$\mathcal{A}$输出 2 个长度相等的明文w_0、w_1和一个意欲挑战的公开钥ID^*。唯一的限制是ID^*在阶段 1 中的任何密钥询问中出现时挑战者报错并退出。挑战者随机选择一个比特值$\beta \leftarrow_{\mathrm{R}} \{0,1\}$，用$\mathrm{ID}^*$加密$w_\beta$得到$T_w^*$，并将$T_w^*$发送给敌手$\mathcal{A}$。

阶段 2（训练）　敌手$\mathcal{A}$发出对另外 ID 的密钥产生询问，唯一的限制是$\mathrm{ID} \neq \mathrm{ID}^*$，挑战者以阶段 1 中的方式进行回应，这一过程可重复多项式有界次。

猜测　敌手$\mathcal{A}$输出猜测$\beta' \in \{0,1\}$，如果$\beta' = \beta$，则敌手$\mathcal{A}$攻击成功。

敌手$\mathcal{A}$的优势定义为安全参数K的函数

$$\mathrm{Adv}_{T_w,\mathcal{A}}^{\mathrm{CKA}}(K) = |\Pr[\mathrm{Succ}] - \frac{1}{2}|$$

3.3.3　密钥保护信息的不可区分性

下面，通过敌手$\mathcal{A}$与挑战者之间的安全游戏来定义本文方案的密钥保护信息不可区分性。密钥保护信息K的 IND-CPA 游戏定义如下。

初始化　挑战者输入安全参数K，产生公开的系统参数 params 和保密的主密钥。

阶段 1（训练）　敌手$\mathcal{A}$发出对 ID 的密钥产生询问并声明意欲挑战的身份ID^*，否则记为ID_i。挑战者运行密钥生成算法，根据是否为要挑战的 ID 用不同的方式返回$H_1(\mathrm{ID})$与私钥给敌手$\mathcal{A}$，这一过程可重复多项式有界次。

挑战　敌手$\mathcal{A}$输出两个长度相等的明文M_0、M_1和一个意欲挑战的公开钥ID^*。唯一的限制是ID^*在阶段 1 中的任何密钥询问中出现时挑战者报错并退出。挑战者随机选择一个比特值$\beta \leftarrow_{\mathrm{R}} \{0,1\}$，用$\mathrm{ID}^*$加密$M_\beta$得到$K^*$，并将$K^*$发送给敌手$\mathcal{A}$。

阶段 2（训练）　敌手$\mathcal{A}$发出对另外 ID 的密钥产生询问，唯一的限制是$\mathrm{ID} \neq \mathrm{ID}^*$，挑战者以阶段 1 中的方式进行回应，这一过程可重复多项式有界

猜测　敌手$\mathcal{A}$输出猜测$\beta' \in \{0,1\}$，如果$\beta' = \beta$，则敌手$\mathcal{A}$攻击成功。

敌手$\mathcal{A}$的优势定义为安全参数K的函数

$$\mathrm{Adv}_{K,\mathcal{A}}^{\mathrm{CPA}}(K) = |\Pr[\mathrm{Succ}] - \frac{1}{2}|$$

4　具体方案

4.1　方案描述

基于区块链的多用户环境中公钥可搜索加密方案的运行过程如图 3 所示，本文方案具体执行细节如下。

1) 系统初始化算法 SetUp(λ)。该算法以安全参数λ为输入。

① PKG 生成阶数为大素数$p(p > 2^\lambda)$的乘法群G_1、G_2，g为G_1的生成元。生成一个双线性映射$\hat{e}: G_1 \times G_1 \to G_2$。选择 5 个抗碰撞哈希函数$H_1: \{0,1\}^* \to G_1$、$H_2: G_1 \to Z_p^*$、$H_3: \{0,1\}^* \to G_2$、$H_4: G_2 \to \{0,1\}^k$和$H_5: \{0,1\}^* \to \{0,1\}^k$。随机选取$y \leftarrow_{\mathrm{R}} Z_p^*$作为系统私钥，计算系统公钥为$Y = \hat{e}(g,g)^y$和$Y' = g^y$。

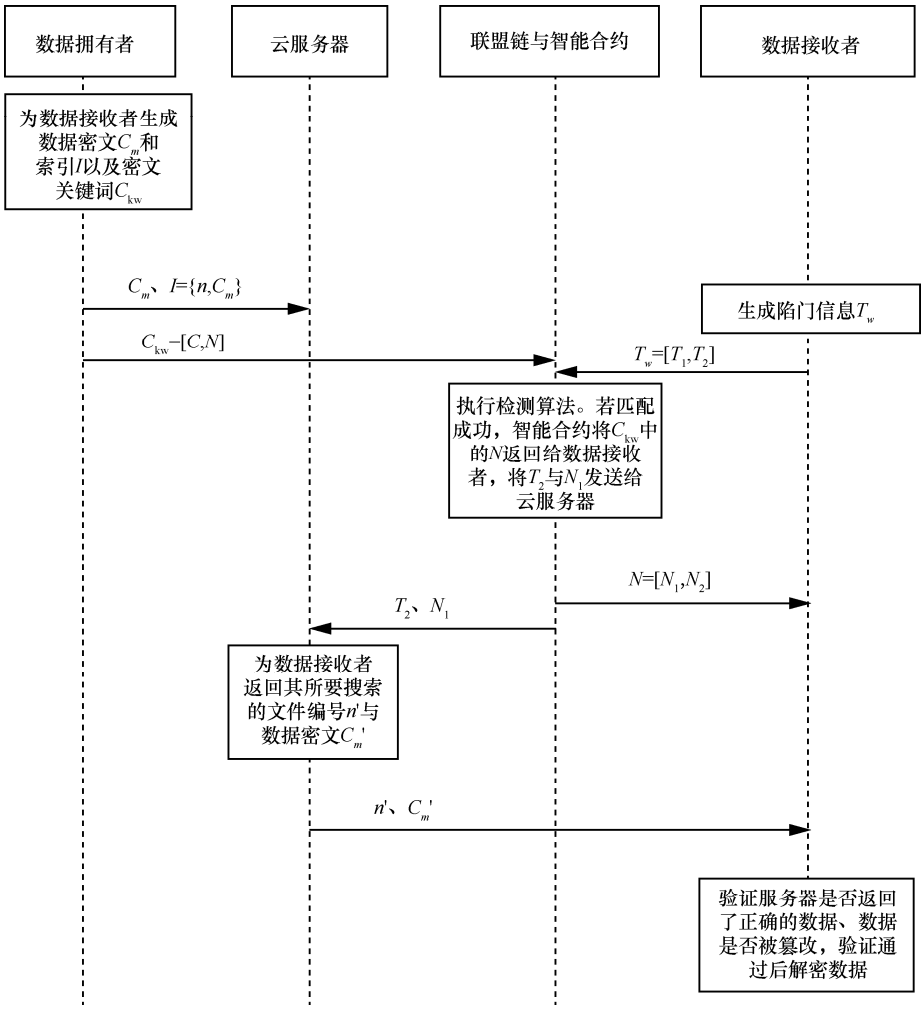

图 3　基于区块链的多用户环境中公钥可搜索加密方案运行过程

② PKG 公布系统公共参数 $\text{params}=(p,G_1,G_2,g,\hat{e},H_1,H_2,H_3,H_4,H_5,Y,Y')$，权威中心将智能合约部署在联盟链上。

2) 密钥生成算法 KeyGen (params， ID_U , ID_A , y)。该算法以系统公共参数 params 、用户身份 ID_U 、权威中心标识 ID_A 、系统私钥 y 为输入。设群组内共有 m 个用户(m 可动态增加)，则有 $1\leqslant U\leqslant m$ ，$m\in N_+$ 。

① 用户 ID_U 随机选取 $u\leftarrow{}_\text{R}Z_p^*$，利用自己身份计算 $[\text{ID}_\text{U},H_1(\text{ID}_\text{U})^u]$ 发送给 PKG，PKG 计算 $\text{sk}_\text{U}=H_1(\text{ID}_\text{U})^y$ 发送给用户，然后计算用户身份公钥 $T_{\text{ID}_\text{U}}=g^{x_\text{U}}$ ，其中 $x_\text{U}=H_2[(H_1(\text{ID}_\text{U})^u)^y]$ 。用户收到 sk_U 后计算 $D_\text{U}=(Y')^{(x'_\text{U})^{-1}}=g^{\frac{y}{x'_\text{U}}}$ ，将 sk_U 、 D_U 作为自己的私钥并保存，其中 $x'_\text{U}=H_2(\text{sk}_\text{U}^\text{u})=H_2(H_1(\text{ID}_\text{U})^{yu})$ 。

② 权威中心将自己的标识 ID_A 发送给 PKG，PKG 计算 $\text{sk}_\text{A}=H_1(\text{ID}_\text{A})^y$ 发送给权威中心，权威中心收到后将 sk_A 作为自己的私钥并保存。

3) 密文关键词生成算法 KeywordGen (params , w , M)。该算法以系统公共参数 params 、一个明文关键词 w 、明文数据 M 为输入。

① 数据拥有者首先指定可以检索密文的 $i(i\leqslant m,i\in N_+)$ 个用户的集合 $\theta=\{1\leqslant U\leqslant i\,|\,\text{ID}_\text{U}$ ，然后随机选取 $r_1\leftarrow{}_\text{R}Z_p^*$ ，并使用 θ 对应的 i 个用户的身份公钥 T_{ID_U} 计算集合 $T_i=\{1\leqslant U\leqslant i\,|\,T_{\text{ID}_\text{U}}^{r_1}\}$ 。最后，生成密文 $C=[C_1,C_2,C_3]=[H_3(w)Y^{r_1},\theta,T_i]$ 。

② 数据拥有者采用对称加密算法加密明文数据 M 。随机选择 $\eta\leftarrow{}_\text{R}G_2$ 并计算对称加密密钥 $k=H_4(\eta)$ ，将 k 作为对称加密密钥加密 M 得到 M' 。然后，共享者计算密钥保护信息 $K=\eta Y^{r_1}$ 并将 K 拼接在 M' 头部形成一个大文件作为数据密文

C_m ($C_m = K \| M'$) ，如图 4 所示。

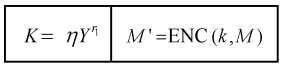

图 4　数据密文 C_m 结构

③ 数据拥有者从正整数集上选择一个数 n 作为要上传的密文关键词的编号，使用编号 n 与密文 C_m 生成 $N=[N_1,N_2]=[n,H_5(n\|C_m)]$。

④ 数据拥有者将密文关键词 $C_{\mathrm{kw}}=[C,N]$ 上传至联盟链，将 n 与数据密文 C_m 建立索引关系 $I=\{n,C_m\}$ (能够通过 n 查询到 C_m)后将 I 和 C_m 上传至 CS。

4) 陷门信息生成算法 TrapdoorGen (params , w' , T_i , D_{U} , $\mathrm{ID_A}$)。该算法以系统公共参数 params 、用户想要搜索的关键词 w' 、T_i 、用户私钥 $\mathrm{ID_U}$ 以及权威中心身份标识 $\mathrm{ID_A}$ 为输入。

数据接收者首先查看密文关键词 C_{kw} 中设置的集合 θ 中是否有自己的身份。若有，则使用想要搜索的关键词 w' 和 T_i 中对应的 $T_{\mathrm{ID_U}}^{r_1}$ 计算陷门信息 $T_w=[T_1,T_2]=[H_3(w')\hat{e}(T_{\mathrm{ID_U}}^{r_1},D_{\mathrm{U}})\ \hat{e}(\mathrm{sk_U},H_1(\mathrm{ID_A})),\mathrm{ID_U}]$。

5) 检测算法 Test(params , C_{kw} , T_w , $\mathrm{sk_A}$)。该算法以系统公共参数 params 、密文关键词 C_{kw} 、陷门信息 T_w 和权威中心私钥 $\mathrm{sk_A}$ 为输入，返回检索结果 0 或 1。

① 数据接收者有检索需求时，将生成的陷门信息 T_w 发送给智能合约。智能合约从 T_w 中解析出 T_1、T_2，从 C_{kw} 中解析出 C_1，利用权威中心提供的私钥 $\mathrm{sk_A}$ 验证等式 $\hat{e}(\mathrm{sk_A},H_1(T_2))C_1=T_1$ 是否成立。若不成立输出 0，若成立输出 1。

② 当输出 1 时，智能合约将 C_{kw} 中的 N 返回给数据接收者，同时将其身份 $T_2=\mathrm{ID_U}$ 与密文关键词编号 $N_1=n$ 发送给 CS。

③ CS 收到 N_1 后，通过 $N_1=n$ 查询对应的 C_m，随后将 n' 与 C_m' 返回给 T_2 所示用户。其中，n' 为 CS 返回的密文关键词编号，C_m' 为 CS 返回的数据密文。

6) 数据密文解密算法 Dec (N,n',C_m')。该算法以密文关键词 C_{kw} 中的 N 以及服务器返回的 n' 和 C_m' 为输入。

① 用户收到 N、n'、C_m' 后，首先验证等式 $N_1=n'$ 是否成立，若成立表明 CS 返回了用户所要搜索的文件。然后，验证等式 N_2=$H_5(n'\|C_m')$ 是否成立，若成立表明密文数据没有被篡改。

② 若上述验证通过，用户读取 C_m' 的头文件得到 K，然后使用集合 T_i 中的 $T_{\mathrm{ID_U}}^{r_1}$ 和自己私钥 D_{U} 计算 $K/\hat{e}(T_{\mathrm{ID_U}}^{r_1},D_{\mathrm{U}})$ 可恢复出 η，进一步可计算 $k=H_4(\eta)$ 从而恢复出明文数据 M，如图 5 所示。

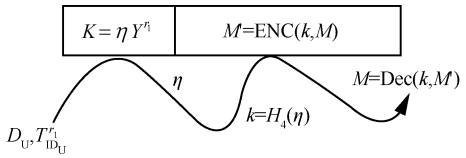

图 5　数据恢复过程

4.2　正确性验证

检测算法正确性

$$\hat{e}(\mathrm{sk_A},H_1(T_2))C_1=\hat{e}(H_1(\mathrm{ID_A})^y,H_1(\mathrm{ID_U}))H_3(w)Y^{r_1}=$$
$$\hat{e}(H_1(\mathrm{ID_A}),H_1(\mathrm{ID_U}))^y H_3(w)\hat{e}(g,g)^{r_1 y}$$
$$T_1=H_3(w')\hat{e}(T_{\mathrm{ID_U}}^{r_1},D_{\mathrm{U}})\hat{e}(\mathrm{sk_U},H_1(\mathrm{ID_A}))=$$
$$H_3(w')\hat{e}(g^{H_2(H_1(\mathrm{ID_U})^{uy})r_1},g^{\frac{y}{H_2(H_1(\mathrm{ID_U})^{yu})}})$$
$$\hat{e}(H_1(\mathrm{ID_U})^y,H_1(\mathrm{ID_A}))=$$
$$H_3(w')\hat{e}(g,g)^{r_1 y}\hat{e}(H_1(\mathrm{ID_U}),H_1(\mathrm{ID_A}))^y$$

显然，当陷门 T_w 中包含的关键词 w' 与密文关键词 C_{kw} 中包含的关键词 w 相同时等式成立。

数据密文解密算法正确性

$$\frac{K}{\hat{e}(T_{\mathrm{ID_U}}^{r_1},D_{\mathrm{U}})}=$$
$$\frac{\eta\hat{e}(g,g)^{r_1 y}}{\hat{e}\left(g^{H_2(H_1(\mathrm{ID_U})^{uy})r_1},g^{\frac{y}{H_2(H_1(\mathrm{ID_U})^{yu})}}\right)}=\frac{\eta\hat{e}(g,g)^{r_1 y}}{\hat{e}(g,g)^{r_1 y}}=\eta$$

因此，用户可以计算出 k 正确解密密文数据。

5　安全性分析

5.1　密文关键词的不可区分性

定理 1　在随机谕言机模型下，若 MBDH 假设成立，则本方案的密文关键词 C_{kw} 是 IND-CKA 安全的，即满足选择关键词攻击下的不可区分性。

证明　要证明 C_{kw} 的 IND-CKA 安全性只需证明 C_{kw} 中的 C 满足 IND-CKA 安全性即可。

挑战者先建立 MBDH 问题，设 $\mathcal{B}$ 是一个攻击 MBDH 问题的多项式时间敌手，$\mathcal{A}$ 是一个攻击 C 的多项式时间敌手，敌手 $\mathcal{B}$ 以敌手 $\mathcal{A}$ 为子程序，具体挑战过程如下。

初始化 挑战者建立 MBDH 问题，$\mathcal{B}$ 获得一个五元组实例 $T=(g,g^a,g^b,g^c,E)$，其中 $E=\hat{e}(g,g)^{\frac{ab}{c}}$ 或 $E=\hat{e}(g,g)^z$，$a,b,c,z\leftarrow_{\mathrm{R}}Z_p^*$。当 $E=\hat{e}(g,g)^{\frac{ab}{c}}$ 时记 $\mu=0$，当 $E=\hat{e}(g,g)^z$ 时记 $\mu=1$。$\mathcal{B}$ 用实例 T 中的 g^a 生成系统公钥 $Y=\hat{e}(g,g^a)=\hat{e}(g,g)^a$ 和 $Y'=g^a$，其余参数与挑战者方案相同。公开系统公共参数 $\mathrm{params}=(p,G_1,G_2,g,e,H_1,H_2,H_3,H_4,H_5,Y,Y')$。

阶段 1 在此阶段，敌手 $\mathcal{B}$ 承担敌手 $\mathcal{A}$ 的挑战者。敌手 $\mathcal{B}$ 首先获得 $\mathcal{A}$ 意欲挑战的身份 ID^*，否则记为 ID_{U}，然后敌手 $\mathcal{B}$ 模拟一个随机谕言机对敌手 $\mathcal{A}$ 发出的询问进行应答。

敌手 $\mathcal{A}$ 向敌手 $\mathcal{B}$ 发起对 ID^* 的询问。首先，敌手 $\mathcal{A}$ 随机选取 $d\leftarrow_{\mathrm{R}}Z_p^*$ 计算 $[\mathrm{ID}^*,H_1(\mathrm{ID}^*)^d]$ 发送给敌手 $\mathcal{B}$，敌手 $\mathcal{B}$ 随机选择 $r\leftarrow_{\mathrm{R}}Z_p^*$，令 $T_{\mathrm{ID}}{}^*=g^{cr}$。

敌手 $\mathcal{A}$ 向敌手 $\mathcal{B}$ 发起对 ID_{U} 的询问。首先，敌手 $\mathcal{A}$ 随机选取 $u\leftarrow_{\mathrm{R}}Z_p^*$ 计算 $[\mathrm{ID}_{\mathrm{U}},H_1(\mathrm{ID}_{\mathrm{U}})^u]$ 发送给敌手 $\mathcal{B}$，敌手 $\mathcal{B}$ 随机选择 $v\leftarrow_{\mathrm{R}}Z_p^*$ 计算 $\mathrm{sk}_{\mathrm{U}}=H_1(\mathrm{ID}_{\mathrm{U}})^v$ 作为对 $\mathcal{A}$ 的应答，然后公开 $T_{\mathrm{ID}_{\mathrm{U}}}=g^{x_{\mathrm{U}}}$，其中 $x_{\mathrm{U}}=H_2(H_1(\mathrm{ID}_{\mathrm{U}})^{uv})$。$\mathcal{A}$ 令 $D_{\mathrm{U}}=(Y')^{(x'_{\mathrm{U}})^{-1}}=g^{\frac{y}{x'_{\mathrm{U}}}}$，其中 $x'_{\mathrm{U}}=H_2(\mathrm{sk}_{\mathrm{U}}^u)=H_2(H_1(\mathrm{ID}_{\mathrm{U}})^{vu})$。在敌手 $\mathcal{A}$ 看来，所有参数均为随机的且 $\mathcal{B}$ 为其产生的密钥相比于真实方案中的密钥是有效的，这是因为 $\hat{e}(T_{\mathrm{ID}_{\mathrm{U}}},D_{\mathrm{U}})=Y$。

挑战 敌手 $\mathcal{A}$ 输出 2 个长度相等的明文关键词 w_0、w_1 发送给敌手 $\mathcal{B}$。$\mathcal{B}$ 随机选择 $\beta\leftarrow_{\mathrm{R}}\{0,1\}$，计算 w_β 的密文 $C^*=[H_3(w_\beta)E,T^*,\ T_i^*=g^{br}]$ 返回给敌手 $\mathcal{A}$。

阶段 2 与阶段 1 一样。

猜测 敌手 $\mathcal{A}$ 输出猜测 $\beta'\in\{0,1\}$。若 $\beta'=\beta$，$\mathcal{B}$ 输出 $\mu'=0$，表示实例 T 是 MBDH 五元组 M；如果 $\beta'\neq\beta$，$\mathcal{B}$ 输出 $\mu'=1$，表示实例 T 是随机五元组 R。

当 μ=1 时，$E=\hat{e}(g,g)^z$，有 $H_3(w_\beta)E=H_3(w_\beta)\hat{e}(g,g)^z$。由于 z 是随机的，因此在 $\mathcal{A}$ 看来 C^* 也是 G_2 中随机的元素。因为 C^* 是随机的，敌手 $\mathcal{A}$ 没有获得有关 β 的任何信息，所以 $\Pr[\beta'\neq\beta\,|\,\mu=1]=\frac{1}{2}$。而当 $\beta'\neq\beta$ 时，$\mathcal{B}$ 猜测 $\mu'=1$，所以 $\Pr[\mu'=\mu\,|\,\mu=1]=\frac{1}{2}$。

当 μ=0 时，$E=\hat{e}(g,g)^{\frac{ab}{c}}$。如果设 $r'=\frac{b}{c}$，$H_3(w_\beta)E=H_3(w_\beta)\hat{e}(g,g)^{\frac{ab}{c}}=H_3(w_\beta)\hat{e}(g,g)^{ar'}=H_3(w_\beta)Y^{r'}$，$T_i^*=g^{br}=g^{\frac{b}{c}cr}=g^{r'cr}$。所以该密文是消息 $H_3(w_\beta)$ 在公钥 T^* 下加密的结果，$\Pr[\beta'=\beta\,|\,\mu=0]=\Pr[\mathrm{Succ}]$。而当 $\beta'=\beta$ 时，$\mathcal{B}$ 猜测 $\mu'=0$，$\Pr[\mu'=\mu\,|\,\mu=0]=\Pr[\mathrm{Succ}]$。

所以，有

$$\Pr[\mu'=\mu\,|\,\mu=0]-\Pr[\mu'=\mu\,|\,\mu=1]=\Pr[\mathrm{Succ}]-\frac{1}{2}$$

$$\Pr[\mathcal{B}(T)=1\,|\,M]-\Pr[\mathcal{B}(T)=1\,|\,R]=\Pr[\mathrm{Succ}]-\frac{1}{2}$$

$$\Pr[\mathcal{B}(M)=1]-\Pr[\mathcal{B}(R)=1]=\Pr[\mathrm{Succ}]-\frac{1}{2}$$

等式左边与定义 3 中敌手 $\mathcal{B}$ 解决 MBDH 问题的优势定义一致，等式右边为敌手 $\mathcal{A}$ 区分 C 的优势。因此 $\mathrm{Adv}_{\mathcal{B}}^{\mathrm{MBDH}}=\mathrm{Adv}_{C,\mathcal{A}}^{\mathrm{CKA}}$。

由定义 3 可知，$\mathrm{Adv}_{\mathcal{B}}^{\mathrm{MBDH}}\leqslant\varepsilon(K)$ 是可忽略的，因此 $\mathrm{Adv}_{C,\mathcal{A}}^{\mathrm{CKA}}$ 也是可以忽略的，即 C 是 IND-CKA 安全的。因此 C_{kw} 是 IND-CKA 安全的。

证毕。

5.2 陷门信息的不可区分性

定理 2 在随机谕言机模型下，若 DBDH 假设成立，则本文方案的陷门信息 T_w 是 IND-CKA 安全的，即满足选择关键词攻击下的不可区分性。

证明 挑战者先建立 DBDH 问题，设 $\mathcal{B}$ 是一个攻击 DBDH 问题的多项式时间敌手，$\mathcal{A}$ 是一个攻击陷门信息 T_w 的多项式时间敌手，敌手 $\mathcal{B}$ 以敌手 $\mathcal{A}$ 为子程序，具体挑战过程如下。

初始化 挑战者建立 DBDH 问题，$\mathcal{B}$ 获得一个五元组实例 $T=(g,g^a,g^b,g^c,E)$，其中 $E=\hat{e}(g,g)^{abc}$ 或 $E=\hat{e}(g,g)^z$，$a,b,c,z\leftarrow_{\mathrm{R}}Z_p^*$。$\mathcal{B}$ 以实例 T 中的 g^a 作为系统公钥 Y'，其余公共参数用与挑战者方案相同的方式产生，公开系统公共参数 $\mathrm{params}=(p,G_1,G_2,g,\hat{e},H_1,H_2,H_3,H_4,\ H_5,Y')$。

阶段 1 在此阶段，敌手 $\mathcal{B}$ 承担敌手 $\mathcal{A}$ 的挑战者。敌手 $\mathcal{B}$ 模拟一个随机谕言机对敌手 $\mathcal{A}$ 发出的询问进行应答。

H_1 询问　敌手 $\mathcal{B}$ 建立一个初始元素为空的列表 $H_1^{\text{list}}=\{(\cdot,\cdot,\cdot,\cdot)\}$，列表元素为四元组 $(\text{ID},H_1(\text{ID}),r,\text{coin})$。这里 coin 作为 $\mathcal{B}$ 的猜测，$\text{coin}=0$ 表示 $\mathcal{A}$ 将对这次询问的 ID 发起攻击，记要攻击的 ID 为 ID_{U}^* 和 ID_{A}^*；否则 $\text{coin}=1$ 并记 ID 为 ID_i。

若 $\text{coin}=0$，$\text{ID}=\text{ID}_{\text{A}}^*$ 或 $\text{ID}=\text{ID}_{\text{U}}^*$，$\mathcal{B}$ 以实例 T 中的 g^b 和 g^c 作为对 $\mathcal{A}$ 发起的 H_1 询问的应答，即 $H_1(\text{ID}_{\text{A}}^*)=g^b$、$H_1(\text{ID}_{\text{U}}^*)=g^c$。然后，向 H_1^{list} 中添加元组 $(\text{ID}_{\text{A}}^*,g^b,\backslash,0)$、$(\text{ID}_{\text{U}}^*,g^c,\backslash,0)$；否则，$\mathcal{B}$ 随机选取 $r_i\leftarrow_{\text{R}}Z_p^*$，返回 $H_1(\text{ID}_i)=g^{r_i}$ 作为对 $\mathcal{A}$ 发起的 H_1 询问的应答。然后，向 H_1^{list} 中添加元组 $(\text{ID}_i,g^{r_i},r_i,1)$。

密钥提取询问　$\mathcal{A}$ 向 $\mathcal{B}$ 询问 ID 对应的私钥 sk。$\mathcal{B}$ 收到 $\mathcal{A}$ 发送的 ID 后，在 H_1^{list} 查找对应的元组并读取元组中的记录。若 $\text{coin}=0$，$\mathcal{B}$ 报错并退出；否则，$\mathcal{B}$ 从元组中取出 ID_i 对应的 r_i 计算 $\text{sk}_{\text{ID}_i}=(Y')^{r_i}=g^{ar_i}$ 作为对 ID_i 密钥提取询问的应答，这是因为 $\text{sk}_{\text{ID}_i}=(Y')^{r_i}=g^{ar_i}=(g^{r_i})^a=H_1(\text{ID}_i)^a$。

挑战　敌手 $\mathcal{A}$ 输出 2 个长度相等的明文关键词 w_0'、w_1' 和要挑战的身份 ID_{U}^*、ID_{A}^*，随机选择 $x\leftarrow_{\text{R}}Z_p^*$ 计算 $H_3(w_0')Y^x$ 和 $H_3(w_1')Y^x$ 发送给敌手 $\mathcal{B}$，唯一的限制是若 ID_{U}^* 和 ID_{A}^* 在阶段 1 中的密钥提取问中出现时 $\mathcal{B}$ 报错并退出。敌手 $\mathcal{A}$ 计算 $H_3(w_\beta')Y^x$ 发送给敌手 $\mathcal{B}$ 是因为其与真实方案中的 $H_3(w')\hat{e}(T_{\text{U}}^{r_1},D_{\text{U}})=H_3(w')\hat{e}(g,g)^{r_1y}=H_3(w')Y^{r_1}$ 等效。$\mathcal{B}$ 随机选取 $\beta\leftarrow_{\text{R}}\{0,1\}$ 后计算陷门 $T_w^*=H_3(w_\beta')Y^xE$ 返回给敌手 $\mathcal{A}$。

阶段 2　与阶段 1 一样。

猜测　敌手 $\mathcal{A}$ 输出猜测 $\beta'\in\{0,1\}$。若 $\beta'=\beta$，则 $\mathcal{A}$ 挑战成功并输出 1（用 Succ 表示该事件），否则输出 0。同时，$\mathcal{B}$ 也把 $\mathcal{A}$ 的输出作为自己的输出。

在以上规约过程中，如果 $\mathcal{B}$ 不中断，则 $\mathcal{B}$ 的模拟是完备的。这是因为，① $\mathcal{B}$ 对 $\mathcal{A}$ 的 H_1 询问的应答：$\text{coin}=0$ 时，$H_1(\text{ID}_{\text{A}}^*)=g^b$、$H_1(\text{ID}_{\text{U}}^*)=g^c$；$\text{coin}=1$ 时，$H_1(\text{ID}_i)=g^{r_i}$。由于 b、c 和 r_i 都是随机的，因此 $H_1(\text{ID}_{\text{A}}^*)$、$H_1(\text{ID}_{\text{U}}^*)$ 和 $H_1(\text{ID}_i)$ 是随机均匀的，$\mathcal{B}$ 对 $\mathcal{A}$ 的应答达到了随机谕言机的效果。② $\mathcal{B}$ 对 $\mathcal{A}$ 的密钥提取询问的应答 $\text{sk}_{\text{ID}_i}=(Y')^{r_i}=g^{ar_i}=(g^{r_i})^a=H_1(\text{ID}_i)^a$，因而是有效的。

假设 $\mathcal{B}$ 不中断。当 T 为随机五元组 R 即 $E=\hat{e}(g,g)^z$ 时，因为 E 在 G_2 中是随机均匀分布的，所以 T_w^* 的第一部分在 G_2 中也是均匀分布的，$\mathcal{A}$ 不能以超过 1/2 的概率输出 1。而 $\mathcal{B}$ 输出 1 当且仅当 $\mathcal{A}$ 成功，所以 $\Pr[\mathcal{B}(T)=1\mid R]=\frac{1}{2}$。

当 T 为 DBDH 五元组 D 即 $E=\hat{e}(g,g)^{abc}$ 时，由 $Y'=g^a$、$H_1(\text{ID}_{\text{A}}^*)=g^b$，$H_1(\text{ID}_{\text{U}}^*)=g^c$ 得 $E=\hat{e}(g,g)^{abc}$ 与 $\hat{e}(H_1(\text{ID}_{\text{U}}^*)^a,H_1(\text{ID}_{\text{A}}^*))=\hat{e}(g,g)^{abc}$ 相等。这意味着 $\mathcal{B}$ 区分实例 T 时获得的公钥与密文的分布与 $\mathcal{A}$ 区分 T_w^* 时获得的公钥与密文的分布相同，所以 $\mathcal{B}$ 输出 1 当且仅当 $\mathcal{A}$ 成功，此时有 $\Pr[\mathcal{B}(T)=1\mid D]=\Pr[\text{Succ}]$。

所以，有

$$\Pr[\mathcal{B}(T)=1]=\Pr[D]\Pr[\mathcal{B}(T)=1\mid D]+\Pr[R]\cdot$$

$$\Pr[\mathcal{B}(T)=1\mid R]=\frac{1}{2}\Pr[\text{Succ}]+\frac{1}{2}\times\frac{1}{2}$$

$$\Pr[\mathcal{B}(T)=0]=\Pr[D]\Pr[\mathcal{B}(T)=0\mid D]+\Pr[R]\cdot$$

$$\Pr[\mathcal{B}(T)=0\mid R]=\frac{1}{2}[1-\Pr[\text{Succ}]]+\frac{1}{2}\times\frac{1}{2}$$

所以，有

$$|\Pr[\mathcal{B}(T)=1]-\Pr[\mathcal{B}(T)=0]|=|\Pr[\text{Succ}]-\frac{1}{2}|$$

$$|\Pr[\mathcal{B}(T)=1]-\Pr[\mathcal{B}(\bar{T})=1]|=|\Pr[\text{Succ}]-\frac{1}{2}|$$

$$|\Pr[\mathcal{B}(R)=1]-\Pr[\mathcal{B}(D)=1]|=|\Pr[\text{Succ}]-\frac{1}{2}|$$

等式左边与定义 2 中敌手 $\mathcal{B}$ 解决 DBDH 问题的优势定义一致，等式右边为敌手 $\mathcal{A}$ 区分 T_w 的优势。因此 $\text{Adv}_{\mathcal{B}}^{\text{DBDH}}=\text{Adv}_{\text{T}_w,\mathcal{A}}^{\text{CKA}}$。

由定义 2 可知，$\text{Adv}_{\mathcal{B}}^{\text{DBDH}}\leqslant\varepsilon(K)$ 是可忽略的，因此 $\text{Adv}_{\text{T}_w,\mathcal{A}}^{\text{CKA}}$ 也是可以忽略的，即 T_w 是 IND-CKA 安全的。

此外，陷门信息中 T_w 中包含的 $H_3(w')\ \hat{e}(T_{\text{U}}^{r_1},D_{\text{U}})=H_3(w')Y^{r_1}$ 对于权威中心来说也是不可区分的，这一点在 5.1 节中已进行了详细的证明，因此陷门信息 T_w 可以抵御内部关键词猜测攻击。同时，由于陷门信息中包含了权威中心公钥 $H_1(\text{ID}_{\text{A}})$，只有利用智能合约才可以执行检测算法，因此本文方案也满足指定可搜索性可以在公开信道下进行数据传输。

5.3 密钥保护信息的不可区分性

定理 3 在随机谕言机模型下，若 MBDH 假设成立，则本文方案的密钥保护信息 K 是 IND-CPA 安全的，即满足选择明文攻击下的不可区分性。

证明 K 的 IND-CPA 安全性证明与 C 的相同，具体证明过程省略。

证毕。

6 效率分析

实验设备的处理器使用 Intel(R) Core(TM) i5-10210U CPU@1.60 GHz 2.11 GHz，内存为 16 GB。为了让效率分析的结果更加准确，在 Windows10 系统环境下使用 IDEA 编程软件利用 jPBC 密码库对参与比较的文献进行了仿真实现，其中采用了库中的 Type A 类曲线构造对称质数阶双线性群，群的阶数为 512 bit，域的阶数为 160 bit。

本文方案主要与文献[16-18, 21]方案进行对比。实验分别记录各方案生成 100 次密文关键词、生成 100 次陷门信息与执行 100 次检测算法的累计耗时，然后计算其单次运行平均耗时使数据更加可靠，最后进行总结分析。

6.1 密文关键词生成算法的时间

各方案密文关键词生成时间如图 6 所示，与文献[16-18]方案相比，本文方案在同样满足密文关键词不可区分性的情况下还可以适用于一对多通信模式的多用户环境中且生成密文关键词时的效率更高。随着分享用户数量的增加，本文方案的优势将会更加明显。与文献[21]方案相比，在二者都适用于多用户环境中的情况下，本文方案不仅在效率上具有一定优势且文件加密密钥的传输过程更加安全。图 7 展示了各方案生成一个密文关键词的平均耗时。

6.2 陷门信息生成与执行检测算法的时间

记生成一次陷门信息与执行一次检测算法时间的总和为一次交互时间。各方案陷门生成与执行检测算法的时间如图 8 所示。分析结果表明，与文献[16-18]方案的效率相比，本文方案具有较大优势，可以更快地为用户反馈检索结果，改善用户体验。在与文献[21]方案比较时，虽然运行效率相近，但文献[21]方案的陷门信息无法抵抗关键词猜测攻击，在安全性方面不及本文方案。图 9 展示了各方案单次交互平均耗时即运行一次陷门信息生成算法与执行一次检测算法时间总和的平均值。

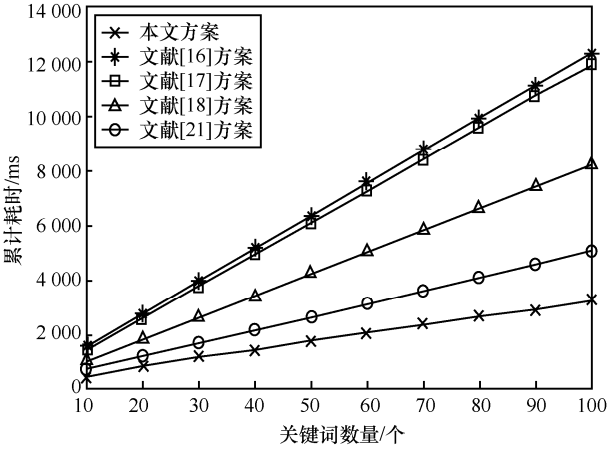

图 6 各方案密文关键词生成时间

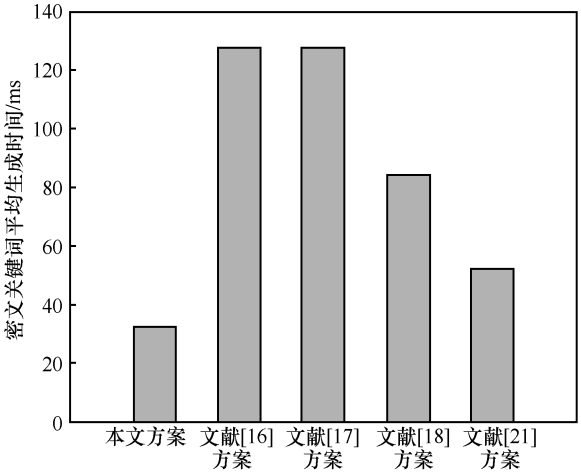

图 7 各方案密文关键词平均生成时间

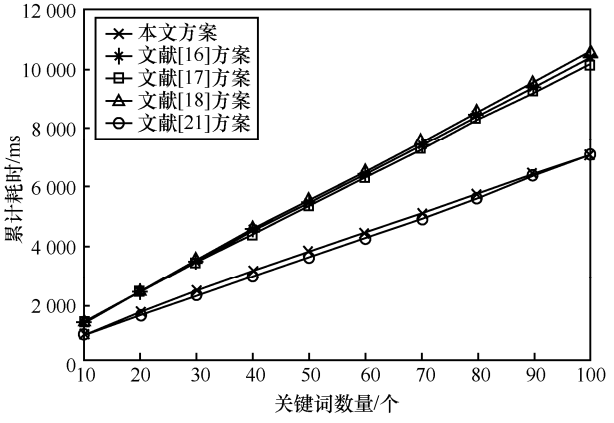

图 8 各方案陷门生成与执行检测算法的时间

6.3 使用不同长度关键词生成密文与陷门的时间

使用不同字符长度的关键词运行本文方案的密文关键词生成算法（取 $i=3$）与陷门信息生成算法 100 次计算平均值如图 10 所示。结果表明，本文方案在使用不同长度关键词时的运行效率相近，具有很高的灵活性。

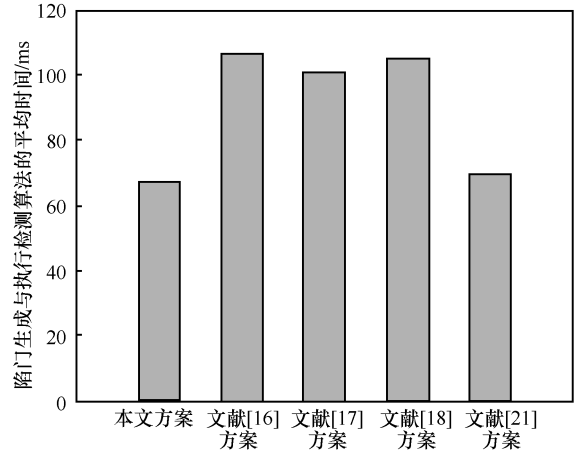

图 9　各方案陷门生成与执行检测算法的平均时间

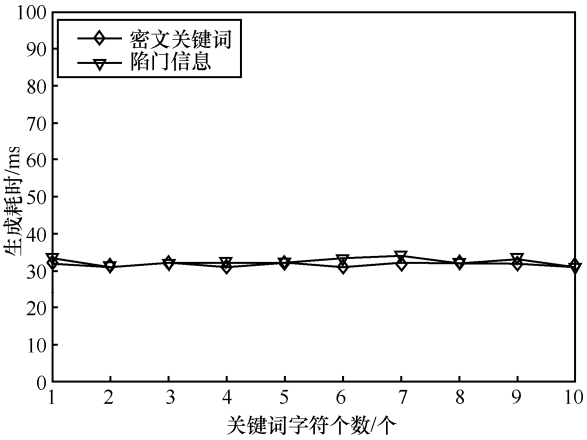

图 10　不同关键词长度下的运行效率

7　结束语

本文将 PEKS 与 IBE 相结合，在多用户环境中实现了一对多模式的公钥可搜索加密方案并设计了文件加密密钥的详细传递算法。数据共享者只需一次加密上传就可以向多位指定的用户进行数据共享，能够灵活运用于实际场景中。利用 IBE 的优势也降低了证书管理开销。此外，引入区块链与智能合约，将索引存储在区块链上并利用智能合约进行检索操作，在保证返回正确检索结果的同时解决了传统第三方存储的半可信问题，保障了数据的可用性。在随机谕言机模型下，基于判定性双线性 Diffie-Hellman 假设和修改的判定性双线性 Diffie-Hellman 假设证明了本文方案的密文关键词和陷门信息满足选择关键词攻击下的不可区分性且可以抵抗内部关键词猜测攻击。同时，分析了引入区块链对保证数据可用性的作用。利用 jPBC 密码库对本文方案和参与效率分析的其他方案进行了仿真实现。结果表明，本文方案在具备安全性与实用性的基础上也具有较高的运行效率。

参考文献：

[1] SONG D X, WAGNER D, PERRIG A. Practical techniques for searches on encrypted data[C]//Proceeding 2000 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2000: 44-55.

[2] BONEH D, DI CRESCENZO G, OSTROVSKY R,et al. Public key encryption with keyword search[C]//Advances in Cryptology - EUROCRYPT 2004. Berlin: Springer, 2004: 506-522.

[3] BAEK J, SAFAVI-NAINI R, SUSILO W. Public key encryption with keyword search revisited[C]//Computational Science and Its Applications – ICCSA 2008. Berlin: Springer, 2008: 1249-1259.

[4] FANG L M, SUSILO W, GE C P, et al. A secure channel free public key encryption with keyword search scheme without random oracle[C]//Cryptology and Network Security. Berlin: Springer, 2009: 248-258.

[5] FANG L M, WANG J D, GE C P, et al. Decryptable public key encryption with keyword search schemes[J]. International Journal of Digital Content Technology and Its Applications, 2010, 4(9): 141-150.

[6] BYUN J W, RHEE H S, PARK H A, et al. Off-line keyword guessing attacks on recent keyword search schemes over encrypted data[C]//Secure Data Management. Berlin: Springer, 2006: 75-83.

[7] YAU W C, HENG S H, GOI B M. Off-line keyword guessing attacks on recent public key encryption with keyword search schemes[C]//Autonomic and Trusted Computing. Berlin: Springer, 2008: 100-105.

[8] RHEE H S, SUSILO W, KIM H J. Secure searchable public key encryption scheme against keyword guessing attacks[J]. IEICE Electronics Express, 2009, 6(5): 237-243.

[9] TANG Q, CHEN L Q. Public-key encryption with registered keyword search[C]//Public Key Infrastructures, Services and Applications. Berlin: Springer, 2010: 163-178.

[10] RHEE H S, PARK J H, SUSILO W, et al. Trapdoor security in a searchable public-key encryption scheme with a designated tester[J]. Journal of Systems and Software, 2010, 83(5): 763-771.

[11] QIN B D, CHEN Y, HUANG Q, et al. Public-key authenticated encryption with keyword search revisited: security model and constructions[J]. Information Sciences, 2020, 516: 515-528.

[12] ABDALLA M, BELLARE M, CATALANO D, et al. Searchable encryption revisited: consistency properties, relation to anonymous IBE, and extensions[J]. Journal of Cryptology, 2008, 21(3): 350-391.

[13] RHEE H S, PARK J H, LEE D H. Generic construction of designated tester public-key encryption with keyword search[J]. Information Sciences, 2012, 205: 93-109.

[14] EMURA K, MIYAJI A, RAHMAN M S, et al. Generic constructions of secure-channel free searchable encryption with adaptive security[J]. Security and Communication Networks, 2015, 8(8): 1547-1560.

[15] SUZUKI T, EMURA K, OHIGASHI T. A generic construction of integrated secure-channel free PEKS and PKE and its application to EMRs in cloud storage[J]. Journal of Medical Systems, 2019, 43(5): 128.

[16] 王少辉, 韩志杰, 肖甫, 等. 指定测试者的基于身份可搜索加密方案[J]. 通信学报, 2014, 35(7): 22-32.
WANG S H, HAN Z J, XIAO F, et al. Identity-based searchable en-

cryption scheme with a designated tester[J]. Journal on Communications, 2014, 35(7): 22-32.

[17] MA M M, HE D B, KUMAR N, et al. Certificateless searchable public key encryption scheme for industrial Internet of Things[J]. IEEE Transactions on Industrial Informatics, 2018, 14(2): 759-767.

[18] 牛淑芬, 谢亚亚, 杨平平, 等. 加密邮件系统中基于身份的可搜索加密方案[J]. 电子与信息学报, 2020, 42(7): 1803-1810.
NIU S F, XIE Y Y, YANG P P, et al. Identity-based searchable encryption scheme for encrypted email system[J]. Journal of Electronics & Information Technology, 2020, 42(7): 1803-1810.

[19] 杨宁滨, 周权, 许舒美. 无配对公钥认证可搜索加密方案[J]. 计算机研究与发展, 2020, 57(10): 2125-2135.
YANG N B, ZHOU Q, XU S M. Public-key authenticated encryption with keyword search without pairings[J]. Journal of Computer Research and Development, 2020, 57(10): 2125-2135.

[20] CURTMOLA R, GARAY J, KAMARA S, et al. Searchable symmetric encryption: improved definitions and efficient constructions[J]. Journal of Computer Security, 2011, 19(5): 895-934.

[21] 杜瑞忠, 谭艾伦, 田俊峰. 基于区块链的公钥可搜索加密方案[J]. 通信学报, 2020, 41(4): 114-122.
DU R Z, TAN A L, TIAN J F. Public key searchable encryption scheme based on blockchain[J]. Journal on Communications, 2020, 41(4): 114-122.

[22] 张玉磊, 文龙, 王浩浩, 等. 多用户环境下无证书认证可搜索加密方案[J]. 电子与信息学报, 2020, 42(5): 1094-1101.
ZHANG Y L, WEN L, WANG H H, et al. Certificateless authentication searchable encryption scheme for multi-user[J]. Journal of Electronics & Information Technology, 2020, 42(5): 1094-1101.

[23] 王文明, 施重阳, 王英豪, 等. 基于区块链技术的交易及其安全性研究[J]. 信息网络安全, 2019(5): 1-9.
WANG W M, SHI C Y, WANG Y H, et al. Research on transaction and security based on blockchain technology[J]. Netinfo Security, 2019(5): 1-9.

[24] 马春光, 安婧, 毕伟, 等. 区块链中的智能合约[J]. 信息网络安全, 2018(11): 8-17.
MA C G, AN J, BI W, et al. Smart contract in blockchain[J]. Netinfo Security, 2018(11): 8-17.

[25] ZHANG A Q, LIN X D. Towards secure and privacy-preserving data sharing in e-health systems via consortium blockchain[J]. Journal of Medical Systems, 2018, 42(8): 140.

[26] 高梦婕, 王化群. 基于区块链的可搜索医疗数据共享方案[J]. 南京邮电大学学报(自然科学版), 2019, 39(6): 94-103.
GAO M J, WANG H Q. Blockchain-based searchable medical data sharing scheme[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2019, 39(6): 94-103.

[27] LI H G, ZHANG F G, HE J J. A searchable symmetric encryption scheme using blockchain[J]. arXiv Preprint, arXiv:1711.01030, 2017.

[28] LI H G, TIAN H B, ZHANG F , et al. Blockchain-based searchable symmetric encryption scheme[J]. Computers & Electrical Engineering, 2019, 73: 32-45.

[29] CHEN L X, LEE W K, CHANG C C, et al. Blockchain based searchable encryption for electronic health record sharing[J]. Future Generation Computer Systems, 2019, 95: 420-429.

[30] 牛淑芬, 刘文科, 陈俐霞, 等. 基于联盟链的可搜索加密电子病历数据共享方案[J]. 通信学报, 2020, 41(8): 204-214.
NIU S F, LIU W K, CHEN L X, et al. Electronic medical record data sharing scheme based on searchable encryption via consortium blockchain[J]. Journal on Communications, 2020, 41(8): 204-214.

[作者简介]

郑东（1964− ），男，山西翼城人，博士，西安邮电大学教授、博士生导师，主要研究方向为密码学理论和网络安全。

朱天泽（1997− ），男，河南郑州人，西安邮电大学硕士生，主要研究方向为云计算安全。

郭瑞（1984− ），男，河南洛阳人，博士，西安邮电大学副教授、硕士生导师，主要研究方向为云计算安全、区块链中的隐私保护技术。