基于区块链的公钥可搜索加密方案

杜瑞忠[1,2]，谭艾伦[1]，田俊峰[1,2]

（1. 河北大学网络空间安全与计算机学院，河北 保定 071002；2. 河北省高可信信息系统重点实验室，河北 保定 071002）

摘 要：针对公钥加密方案的陷门安全问题，引入随机数构造陷门与索引，用于抵御来自服务器内部的关键字猜测攻击，避免因服务器好奇行为带来的数据泄露。对第三方的可信问题进行研究，将区块链技术与可搜索加密方案相结合，使用智能合约作为可信第三方进行检索工作，既可以防止服务器内部的关键字猜测攻击，又可以保证检索结果的正确性，从而限制服务器在下发数据时的恶意行为。通过安全性分析，验证了所提方案满足 IND-KGA 安全性。经过与其他方案进行实验对比，证明了所提方案在时间开销上具有一定的优势。

关键词：可搜索加密；区块链；智能合约；公钥加密

中图分类号：TP309

文献标识码：A

doi: 10.11959/j.issn.1000−436x.2020070

Public key searchable encryption scheme based on blockchain

DU Ruizhong[1,2], TAN Ailun[1], TIAN Junfeng[1,2]

1. Cyberspace Security and Computer College, Hebei University, Baoding 071002, China

2. Key Laboratory on High Trusted Information System in Hebei Province, Baoding 071002, China

Abstract: Aiming at the trapdoor security problem of the public key encryption scheme, a random number constructing trapdoor and index was introduced to defend against keyword guessing attacks from the server and avoid data leakage caused by server curious behavior. Research on trusted issues of third parties, the blockchain mechanism with a searchable encryption scheme was combined, and smart contracts as trusted third parties for retrieval was used, which could prevent keyword guessing attacks inside the server and ensure retrieval. The correctness of the results, thereby limiting the malicious behavior of the server when sending data. The solution was analyzed for security and the verification scheme satisfies IND-KGA security. Experiments in real data sets, compared with other programs, prove that the program has certain advantages in time overhead.

Key words: searchable encryption, blockchain, smart contract, public key encryption

1 引言

云存储是当下一种主流的在线存储方式，在免去用户本地存储的硬件与管理开销的同时，使数据脱离了用户的物理控制，因此数据的安全受到了巨大威胁。为了解决云存储的数据安全问题，一般采用数据加密的方式，但是加密后的数据在云服务器中会面临检索困难的问题。

安全搜索通常指对加密数据的有效搜索，为了解决加密数据存储在云端时，服务器在不完全可信的前提条件下如何利用服务器来完成安全的关键字搜索问题，学者们提出了将可搜索加密作为安全搜索的核心技术。

可搜索加密是一种支持用户在密文中进行关键字检索的新技术，主要解决云存储环境下如何利用不可信服务器实现基于关键字的安全搜索，使

收稿日期：2019−10−21；**修回日期：**2020−01−30

基金项目：国家自然科学基金资助项目（No.61572170, No.61170254）；河北省自然科学基金重点资助项目（No.F2019201290）；河北省自然科学基金资助项目（No.F2018201153）

Foundation Items: The National Natural Science Foundation of China (No.61572170, No.61170254), The Key Projects of Natural Science Foundation of Hebei Province (No.F2019201290), The Natural Science Foundation of Hebei Province (No.F2018201153)

用户能够将加密的数据存储到云中，并通过密文域来执行关键字搜索，有选择地从云中检索感兴趣的文档。

2　相关工作

2000 年，Dawn 等[1]为了增强数据在服务器上的安全性，提出一对一模式的可搜索加密方案，从此引发人们对可搜索加密的研究。由于一对一的模式不能满足人们的需求，Boneh 等[2]于 2004 年提出多对一的模式可搜索加密模型，给出了基于公钥的可搜索加密（PEKS, public key encryption with keyword search）的概念，并定义了公钥加密下可搜索加密的安全性。但在某些特性的环境下，多对一模式并不实用。2011 年，Curtmola 等[3]基于 Naor 广播加密技术构造出一对多的可搜索加密模型，但是该模型中用户密钥更换时需要极大的开销。在大型网络环境中，数据的传输是复杂的，Wang 等[4]基于 Shamir 的秘密共享技术和文献[2]中基于身份的加密技术构造了多对多模式的加密方案，实现了多用户在服务器中的交互式检索。Yuan 等[5]为了有效地解决多接收者时变密文的检索问题，提出了一种一对多公钥密文时间释放可搜索加密（PKTRSE_(OM)）的密码模型，在 PKTRSE_(OM)模型中，发送方将加密消息发送给云服务器，只有预定的授权用户组成员才能搜索到包含指定关键字的目标密文，但是直到将来发布时才能解密。Zhong 等[6]提出一种多对一的同态加密方案，克服了传统同态加密一对一的局限性。

在可搜索加密方案的安全性方面，Boneh 等[2]证明了公钥可搜索加密是基于语义安全的，但却不能抵御关键字猜测攻击（KGA, keyword guess attack）。2009 年，Tang 等[7]提出了一种基于公钥加密的注册关键字搜索方案，该方案可以抵御 KGA，但是必须预先注册关键字，这使方案性能并不高。2013 年，Fang 等[8]提出一种可以抵御关键字攻击的公钥加密方案，该方案定义了一个公钥可搜索加密模型和 2 个重要的安全概念。这 2 个安全概念中，一个针对内部攻击，另一个针对外部攻击，但是大量的双线性配对计算导致 Fang 等[8]方案的效率较低。

近年来，在内部攻击方面，学者们进行了很多研究。2013 年，Xu 等[9]提出了带有 2 个陷门（模糊陷门和精确陷门）的方案，并称该方案可以抵抗内部 KGA。在该方案中，敌手只能获得模糊陷门，因此无法提取关于陷门对应关键字的确切信息，受到了安全性和效率方面的限制。2015 年，Chen 等[10]引入了一种新的框架以防止内部 KGA，该框架使用 2 个服务器，并要求 2 个服务器不能相互“勾结”。但是，任何人都可以生成关键字的合法陷门，这将影响数据的隐私安全。Shao 等[11]提出并解决了服务器进行离线 KGA 的问题，重新定义了 dPEKS（designated tester public key encryption with keyword search）对 KGA 的安全性并提出了 IND-KGA-SERVER 安全性，根据公钥基础设施证书颁发机构和确定性数字签名的存在情况，演示当 KGA 攻击者是服务器时如何构造安全的 dPEKS。2016 年，Chen 等[12]提出一种使用 2 个云服务器的方案来抵御内部 KGA，并且方案具有较高的效率，但是由于假设条件中要求 2 个云服务器不能串联，这在实践中很难实现。2017 年，Huang 等[13]基于关键字搜索提出了一个公钥认证加密方案，该方案的密文生成过程中需要数据所有者的密钥，虽然方案可以抵抗内部 KGA，但无法实现所选关键字密文的不可区分性。Kang 等[14]提出一种利用双线性对和 TF/IDF（term frequency/inverse document frequency）算法构成的完全安全的公钥加密方案，该方案在静态假设条件下达到了安全，与传统的可搜索加密方案相比，该方案在可搜索效率、密文完整性和安全性方面有较好的性能。2018 年，Wu 等[15]提出了一种高效安全的、具有隐私保护的可搜索公钥加密方案，该方案使用了 Diffie-Hellman 共享密钥，并被证明能够抵抗 KGA。

在最新的研究成果中，公钥可搜索加密被运用于各种环境，Wu 等[16]在物联网环境中，提出了一种无证书的可搜索公钥认证加密方案，在能抵御 KGA 的同时，也具有较高的效率。Ma 等[17]设计了一种新的基于多关键字的无证书公钥加密方案，用于 IoT（Internet of things）部署。Lu 等[18]针对电子病历系统，提出了 2 种安全的无信道 PEKS 方案，后来经过证明，2 种方案都存在由 KGA 引起的安全漏洞。针对这一问题，Lu 等[18]又提出了一种新的 PEKS 方案，该方案不仅能抵抗现有的 3 种类型的关键字猜测攻击，还改善了指定服务器的缺点。

随着区块链的发展，可搜索加密与区块链技术相结合，解决了传统方案中可信第三方的问题，极大地提高了可搜索加密的可实现性。Li 等[19]提出了

一种基于区块链的对称可搜索加密方案，该方案不仅提出了基于区块链的可搜索加密模型，还针对不同大小的数据提出了 2 种方案。2019 年，Li 等[20]对文献[19]的方案进行了改进，提高了可实现性。Chen 等[21]基于区块链机制提出了一个用于电子医疗记录分享的可搜索加密方案，该方案同样通过对称加密的方法，使用智能合约作为方案中的权威可信方，保证方案中服务器的可信度。

针对现有方案中第三方的可信问题，本文引入区块链，构建区块链环境下的公钥可搜索加密方案，旨在解决私有云环境中一对多的数据分享问题。本文的主要贡献如下。

1) 在密文检索方案中引入区块链机制，利用区块链解决传统方案中第三方的可信问题；将检索工作放到区块链中进行计算，保证检索结果的正确性；利用区块链的不可篡改性，对文件进行编号，防止在云服务器错误时发送数据，或恶意发送错误的数据。

2) 针对私有云环境，构造一对多的公钥可搜索加密方案。该方案中使用 DBDH（decisional bilinear Diffie-Hellman）困难问题的构建方式，使同一个关键字多次加密结果不同，可以有效抵御 KGA，保证索引及陷门不会泄露关键字信息。

3) 对所提方案进行安全性证明，验证了方案可抵御 KGA，同时分析区块链的安全性在方案中的作用。本文基于 PBC（pairing based cryptography）库环境，在数据集上进行实验，得出方案的索引与陷门构造以及查询时间，证明了所提方案具有较高的效率。

3 预备知识

3.1 双线性映射

假设群 G 与群 G_T 是阶为素数 p 的循环群，g 是群 G 的生成元，存在双线性映射 $\hat{e}:G\times G\to G_T$ 并满足以下性质。

1) 双线性。对任意的 $x,y\in G$，$a,b\in G_T$，存在 $\hat{e}(x^a,y^b)=\hat{e}(x^b,y^a)=\hat{e}(x,y)^{ab}$。

2) 非退化性。存在 $g\in G$，使 $\hat{e}(g,g)\neq 1$。

3) 可计算性。对所有的 $x,y\in G$，存在有效的算法来计算 $\hat{e}(x,y)$。

3.2 判定性双线性 Diffie-Hellman 假设

设群 G_1、G_2 及双线性映射 $\hat{e}:G_1\times G_1\to G_2$，$g$ 是群 G_1 的生成元，随机生成$(a,b,c,z)\leftarrow_R Z_p$，生成 2 个五元组 $T_0=(g, A=g^a, B=g^b, C=g^c, Z=\hat{e}(g,g)^z)$与 $T_1=(g, A=g^a, B=g^b, C=g^c, Z=\hat{e}(g,g)^{abc})$。其中，$a$、$b$、$c$、$z$ 表示所生成的随机数，${}_R Z_p$ 表示随机的实数空间将 2 个五元组分别记为

$$P_{\mathrm{BDH}}=\{(g,g^a,g^b,g^c,e(g,g)^{abc})\}$$

$$R_{\mathrm{BDH}}=\{(g,g^a,g^b,g^c,e(g,g)^z)\}$$

DBDH 假设指没有多项式时间的敌手，能以不可忽略的优势 ε 来区分五元组 P_{BDH} 与 R_{BDH}。

3.3 智能合约

智能合约是一种旨在以数字方式执行合同谈判的计算机程序。智能合约与传统合同不一定相同，可以是任何类型的计算机程序，利用加密算法和各种安全协议，实现不同类型的智能合约。智能合约有助于交易的可靠执行，而不涉及某些第三方，并且所有交易都是可追踪和不可逆转的。换句话说，智能合约提供的安全性优于传统的合同，并降低了与合同相关的交易成本。

以太坊是一个运行智能合约的分散式平台。在以太坊中，智能合约用于在区块链上执行一些通用计算。由于区块链的特性，所有操作在以太坊中都是透明和可靠的，这意味着理论上可以使用以太坊智能合约来执行任何计算任务。

3.4 符号及其含义

本文的符号及其含义如下。

C_m：对明文 m 对称加密后的密文。

H：哈希函数 h 对密文 C 与编号密文 N 的运算结果。

I：数据文件索引。

k：对称加密密钥。

N：私钥加密后的文件编号。

T_w：关键字陷门。

T_{w_i}：索引中的关键字陷门。

w：文件中的关键字。

w_i：用户查询的关键字，i=1,2,…,I。

$offer：检索单价。

$user：用户账户。

$deposit：押金账户。

4 系统模型

4.1 系统简介

本文具体系统主要由 4 个部分组成：数据拥有者（DO, data owner）、云服务器（CS, cloud server）、智能合约（smart contract）、用户（U, user）。系统模型及其流程如图 1 所示。

1) 数据拥有者。主要工作是计算索引和密文数据，然后将索引上传给智能合约，将密文数据上传给云服务器。

2) 云服务器。主要工作是存储由数据拥有者上传的密文数据，接收由用户上传的数据下发请求，并与智能合约进行交互，得到下发验证结果。

3) 智能合约。主要工作是接收数据拥有者上传的索引与用户上传的陷门，并且进行查询，得到查询结果，通过交易将结果下发给用户。然后通过与服务器的交易，告知服务器是否下发密文。

4) 用户。主要工作是计算陷门，并上传给智能合约，同时向服务器发送密文请求，最后得到数据并解密。

4.2　算法定义

1) $\text{setup}(1^{\lambda}) \to \text{par}$。系统初始化由安全参数 λ 生成公开参数 par。

2) $\text{Enc}(n,m,w,k) \to (C_T, I, N)$。该算法由数据拥有者计算，由明文数据 m 与对称加密密钥 k，以及文件中的关键字 w 生成密文 C_m 与数据文件索引 I。其中明文 m 与对称加密密钥 k 加密后得到密文 C_m，加密关键字 w 时会将密钥 k 一起加密，生成索引 I。将明文进行编号，每个编号 n 使用对称加密后得到密文状态下的编号数据 N。最后使用哈希函数对 N 和 C_m 进行计算得到结果 H，将 N、H 打包后得到 C_T。

3) $T(w_i) \to T_{w_i}$。该算法由用户进行计算，用户选取文件中的关键字 w，加密计算后得到关键字陷门 T_w，并将 T_w 上传给智能合约。

4) $\text{search}(I, T_{w_i}) \to (k, N, H)$。该算法由智能合约进行计算，智能合约在接收到 T_w 与 I 后进行计算，若匹配成功，则得到明文的对称加密密钥 k、编号密文 N 及哈希结果 H。

5) $\text{verify}(C_T, N) \to 0$ 或1。该算法由用户进行计算。用户收到服务器下发的密文 C_T 以后，需要验证服务器是否错误下发密文以及是否恶意破坏或者篡改密文数据。首先验证 C_T 中文件编号 N 与 H 是否同智能合约下发的一致，然后将 C_m 与 N 进行哈希运算，得到结果 H_1，若 H=H_1，则验证成功，输出 1，否则输出 0。

4.3　安全模型

4.3.1　关键字隐私安全游戏

如果不存在敌手 A 能够在概率多项式时间内从密文关键字或陷门值推断出关键字明文信息，则关键字的隐私安全可以得到保证。定义关键字隐私安全游戏如下。

1) 初始化。给定安全参数 λ，挑战者 C 执行初始化算法 $\text{Init}(1^{\lambda})$，生成 par。

2) 阶段 1。敌手 A 多次运行陷门生成算法。

3) 挑战。敌手 A 从关键字空间随机选取 2 个关键字，发送给挑战者，挑战者执行陷门生成算法，然后随机选取一个陷门发送给敌手 A。

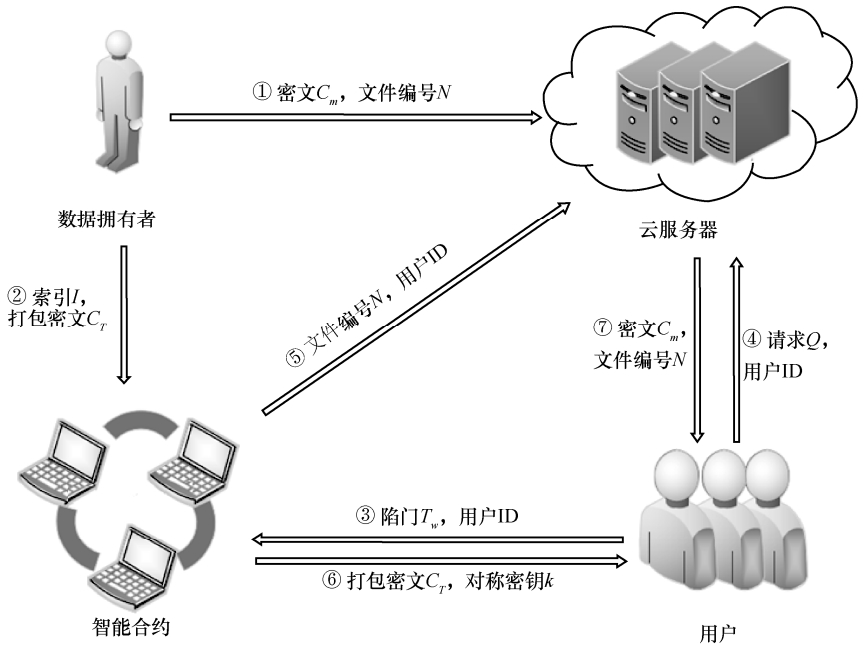

图 1　系统模型及其流程

4) 猜测。敌手 A 查询了 τ 个不同的关键字后，进行分析猜测，如果敌手能够猜对陷门，则敌手 A 在安全游戏中获胜。

4.3.2 判定性双线性 Diffie-Hellman 假设困难问题规约证明

如果存在敌手 A 能够在多项式时间内以优势 δ 破解方案，则敌手 A 能在多项式时间内以优势 δ 解决 DBDH 困难问题。定义判定性双线性 Diffie-Hellman 假设困难问题规约证明如下。

1) 初始化。给定群组 G_1、G_2 及映射 $\hat{e}: G_1 \times G_1 \to G_2$。挑战者 C 随机生成（$a$, b, c, z）$\leftarrow_R Zp$，生成 2 个五元组 T_0、T_1。

2) 阶段 1。敌手 A 多次运行加密算法。

3) 挑战。挑战者 C 随机选取明文 m，要求 m 在阶段 1 未被查询，并生成密文 C_m，将密文传给敌手 A。

4) 猜测。敌手 A 对密文 C_m 进行分析解密，如果敌手 A 能够解密密文 C_m 且得到正确的明文 m，则敌手 A 在游戏中获胜。

5) 证明。敌手 A 能够对密文进行解密，则敌手 A 也能解决判定性双线性 Diffie-Hellman 假设困难问题。

5 具体系统描述

1) 初始化阶段

$\text{Setup}(1^{\lambda}) \to \text{par}$。系统初始化，由安全参数 λ 生成公开参数 par，其中包括循环群 G、G_T；g 是群 G 的生成元，g_1 是群 G 的元素；哈希运算 h，随机参数 a；计算得到 $g_2=g_a$；双线性映射 $\hat{e}$：$G\times G\to G_T$。

$$\text{par} = \{g, g_1, g_2, a, G, \hat{e}, h\}$$

智能合约初始化，数据拥有者设置检索单价 \$offer。用户使用 ID 注册账户\$user 并存款，区块链系统设置押金账户\$deposit。

2) 密文加密与上传

$\text{Enc}(m, w, k) \to (C_m, I)$。明文 m 由对称加密密钥 k 经过加密得到密文 C_m，然后将对称加密密钥 k 加入索引 I 的计算过程中。首先选择一个随机数 $r\leftarrow_R Z_P$，然后将文件的关键字 w 进行哈希运算得到 $H(w)$，计算后得到

$$I = \left(\hat{e}\left(g_1, g_2\right)^r k, g^r, H\left(w\right)^r\right)$$

数据拥有者将加密后的密文 C_m 与索引 I 进行编号，并将编号使用私钥进行加密，得到密文状态的文件编号 N，将文件编号 N 与密文 C_m 存储在一起后进行哈希运算得到结果 H，将文件编号 N、结果 H 打包为密文 C_T。将文件编号 N 和密文 C_m 上传给服务器进行存储操作，将打包的密文 C_T 与索引 I 上传给智能合约进行查询操作。

3) 陷门加密与上传

$T(w_i) \to T_{w_i}$。用户计算得到索引文件中的关键字 w 的陷门 T_{w_i}。首先将文件中的关键字进行哈希运算得到 $H(w)$，再选择一个随机数 $t\leftarrow_R Z_P$，计算得到

$$T_{w_i} = \{g_1{}^a H\left(w_i\right)^t g^t\}$$

用户上传陷门到智能合约，并由自身账户余额向区块链系统进行存款操作\$user→\$deposit。

4) 查询阶段

$\text{search}(I, T_{w_i}) \to k$。智能合约通过交易来接收用户的索引 I，检查用户 ID 是否合法。然后系统检查押金账户\$deposit 中用户预存的押金是否满足一次搜索，当押金满足时将陷门和数据文件索引进行计算，计算过程如下

$$\hat{e}(g_1, g_2)^r k \frac{\hat{e}(g^t, H(w)^r)}{\hat{e}(g_1{}^a H(w_i)^t, g^r)} =$$

$$\hat{e}(g_1, g_2)^r k \frac{\hat{e}(g^t, H(w)^r)}{\hat{e}(H(w_i)^t, g^r)\hat{e}(g_1{}^a, g^r)} =$$

$$\hat{e}(g_1, g_2)^r k \frac{e(g, H(w))^{tr}}{\hat{e}(H(w_i), g)^{tr}\hat{e}(g_1, g^a)^r} =$$

$$\hat{e}(g_1, g_2)^r k \frac{e(g, H(w))^{tr}}{\hat{e}(H(w_i), g)^{tr}\hat{e}(g_1, g_2)^r} =$$

$$k \frac{e(g, H(w))^{tr}}{\hat{e}(H(w_i), g)^{tr}}$$

如果 w=w_i，则最后的结果为对称加密密钥 k，智能合约将会记下文件编号 N，然后开始下一次的查询，直到所有的文件都检索完毕。

5) 验证阶段

$\text{verify}(C_T, N) \to 0$ 或者 1。智能合约在已经检索出的文件集中进行下一个关键字的检索操作，同时从押金账户中扣去对应的检索单价\$offer，直到押金账户\$deposit 的金额不足以进行一次检索，区块链系统就会返回用户押金不足信息：\$deposit←\$deposit − \$offer。

如果押金账户中的金额能够满足用户上传的所有关键字的检索操作，智能合约将所有满足用户

关键字请求的文件编号 N 以及用户 ID 发送给云服务器，云服务器接收后，根据文件编号 N 将密文 C_m 下发给用户。

同时在智能合约与用户的数据交互过程中，智能合约检索成功后得到密文 C_T 与对称加密密钥 k，然后发送给用户，用户验证 $N_{\mathrm{BS}}=N_{\mathrm{CS}}$。其中，$N_{\mathrm{BS}}$ 表示区块链系统发送的文件编号，N_{CS} 表示云服务器发送的文件编号。

若从服务器与智能合约接收的文件编号相同，则证明服务器没有错误下发数据，然后验证

$$H_1=h(N,C_m)$$

$$H=H_1$$

将密文 C_m 与文件编号 N 进行哈希运算，若得到的结果 H_1 与 C_T 中的 H 相等，则证明服务器没有对密文数据进行篡改，最后用密钥 k 对密文 C_m 进行解密，得到明文 m。

6　安全性分析

将检索过程放在区块链系统中运行，可以保证以下几个方面的安全。

1) 公正性。由于区块链与每个用户进行交互时都在基于交易的基础上，每次交易都是透明的，那么可以保证每次查询的结果是正确的，且不会存在恶意篡改结果的情况。同时由于每次交易需要一定的费用，可以有效地防止恶意用户破坏方案正常工作的情况。

2) 可信性。区块链给出的检索结果一定是诚实可信的，同时也能以这个结果为基准，有效地防止恶意服务器对本文方案造成的威胁。用户可以有效地验证服务器操作的正确性，从而获得正确的检索文件。

3) 安全性。本文方案能够保证关键字的安全性，由于关键字陷门是随机加密的，因此满足 IND-KGA 安全。此外，由于关键的数据文件索引 I 的构造是按照判定性双线性 Diffie-Hellman 假设困难问题中的五元组的构造方式来进行的，密文的安全性可以规约为判定性双线性 Diffie-Hellman 假设困难问题。

定理 1　基于一般的双线性群，本文方案在随机预言模型下是满足 IND-KGA（indistinguish keyword guess attack）安全的。

初始化　挑战者 C 生成随机数 $a, b\leftarrow_{\mathrm{R}}Z_{\mathrm{P}}$，公开参数 $\mathrm{par}=\{g,g_1,g_2,a,G,\hat{e},h\}$。

1) 阶段 1。敌手选取关键字集合$(w_1,w_2,\cdots,w_n)$，发送给挑战者 C，挑战者输出关键字集合生成的陷门集合$(T_{w_1},T_{w_2},\cdots,T_{w_n})$，并发送给敌手 A。

2) 挑战 1。敌手 A 随机选取 2 个关键字 w_0、w_1，并要求 w_0、w_1 没有在第一阶段被查询过。然后将 2 个关键字发给挑战者。挑战者选择随机数 p，运行陷门生成算法，计算 $T_{w_0}=(g^p,w_0{}^p)$、$T_{w_1}=(g^p,w_1{}^p)$，然后选取随机数 $\mu\leftarrow(0,1)^{\lambda}$，将 T_{w_μ} 发送给敌手 A。

3) 猜测。敌手 A 对阶段 1 与阶段 2 中查询的关键字陷门进行分析，输出 μ'，如果 $\mu'=\mu$，则敌手 A 赢得游戏。

证明　本文方案是支持关键字隐私安全的，由于关键字陷门在加密时引入了随机数，导致同一个关键字生成的陷门不同，可有效抵御统计分析攻击。敌手 A 在安全游戏中获胜的概率最多是 $\dfrac{1}{|\Psi|-n}+\varepsilon$。其中，$n$ 表示关键字集的个数，ε 表示在安全参数 λ 下可以忽略的概率，Ψ 表示关键字的空间。

证毕。

定理 2　基于一般的双线性群，本文方案的安全性可以规约到判定性双线性 Diffie-Hellman 假设困难问题。如果敌手 A 能够在多项式时间内以优势 δ 破解方案，则敌手 A 能在多项式时间内解决 DBDH 困难问题。

初始化　建立系统，生成安全参数 λ，然后运行算法 setup(1^{λ})，得到安全参数 par。

$$\mathrm{par}=\{g,g_1,g_2,a,G,\hat{e},h\}$$

1) 阶段 1。敌手 A 多次运行索引加密算法。

2) 挑战 1。挑战者 C 选取 2 个密钥 k_1 和 k_2，要求它们在阶段 1 不能被敌手 A 查询。运行加密算法，同时生成随机数 t，计算得到 $I_1=(\hat{e}(g_1,g_2)^r k_1,g^r,H(w)^r)$，$I_2=(\hat{e}(g_1,g_2)^t k_2,g^t,H(w)^t)$。然后挑战者 C 随机发送一个索引 I^*给敌手 A。

3) 猜测。敌手 A 收到索引 I^*以后，对密文进行分析解算。然后敌手输出猜测的结果 I'，如果 $I'=I^*$，则敌手 A 赢得游戏。如果敌手 A 能够对密文 I'进行正确解密，那么敌手 A 就能区分密文 I'中的 $\hat{e}(g_1,g_2)^r$。

4) 阶段 2。敌手 A 尝试破解判定性双线性 Diffie-Hellman 假设中的 2 个五元组。敌手 A 多次运行算法计算这 2 个五元组。

5) 挑战 2。挑战者 C 随机选择$(a,b,c,z)\leftarrow_{\mathrm{R}}Z_{\mathrm{P}}$，

生成 2 个五元组，T_0 是 BHD 五元组，T_1 是随机五元组，具体如下。

$$T_0=(g,A=g^a,B=g^b,C=g^c,Z=e(g,g)^z)$$

$$T_1=(g,A=g^a,B=g^b,C=g^c,Z=e(g,g)^{abc})$$

挑战者 C 随机生成 $\mu\leftarrow R\{0,1\}$，若 μ=0，则输出 T_0，若 μ=1，则输出 T_1。挑战者将得到的五元组发送给敌手 A。

6) 猜测。敌手 A 接收到挑战者 C 发出的五元组 T^*进行分析，然后输出 μ'，如果 $\mu'=\mu$，则敌手 A 在游戏中获胜。

由上述分析可得，敌手 A 能够对密文中的 $\hat{e}(g_1,g_2)^t$ 进行区分，然而密文中的 $\hat{e}(g_1,g_2)^r=\hat{e}(g,g)^{abt}$，也就是说，敌手 A 能够对 $\hat{e}(g,g)^{abr}$ 进行区分，则敌手 A 能够对挑战者 C 发送的五元组 T^*中的 $\hat{e}(g,g)^*$进行区分。因此，敌手 A 能够正确输出对 μ 的猜测值。

证明 敌手 A 能够对索引 I 进行解密，则敌手 A 能够解决判定性双线性 Diffie-Hellman 假设困难问题。综上所述，方案的密文安全性可以规约到判定性双线性 Diffie-Hellman 假设困难问题。

证毕。

7 实验分析

实验环境为 64 bit Windows 操作系统、Intel® Core(TM) i5-4570 CPU 3.20 GHz、内存 16 GB，本文实验主要利用本地的虚拟机 VMware 加载开源项目 OpenStack 来进行性能测试，使用 C++语言，加密函数由 PBC 函数库提供。

本节实验将本文方案与文献[12-13, 15]这 3 种方案进行对比，分别对比了陷门生成时间、索引生成时间和关键字检索时间。实验中的关键字数量以 50 为步长，从 50 依次递增到 500，对每一个关键字数量进行 50 次反复实验，求出时间开销的平均值，保证实验结果的有效性。同时进行字符串字符数量与时间开销关系的实验，得到字符串的数据复杂度与时间开销无关的结论，本文实验选取 8 个字母的单词作为关键字。

本文使用的数据集由复旦大学国际数据库中心自然语言处理小组提供，其中测试语料共有 9 833 篇文档，训练语料共有 9 804 篇文档。

本节还将本文方案在区块链引入之前和引入之后进行对比实验。利用 testrpc 软件进行本地以太坊网络环境的搭建，然后将本文方案编写为智能合约，并设置挖矿时间为 0，以排除其他时间对结果的影响。

7.1 陷门生成时间

将本文方案与 DS-PEKS[12]、PAEKS[13]和 SPE-PP[15]这 3 种方案进行对比，由图 2 可知，在陷门计算过程中，随着关键字数量的增加，陷门的生成时间也随之增加。对比后发现，本文方案在陷门生成时间上比其他 3 个方案都有一定的优势，并且随着关键字数量的增加，优势越来越大。本文方案中的陷门生成时间不会随着关键字包含的字母数量的增多而增多，这对于查询复杂的字符串有一定的优势。同时在第 6 节已经得到验证，本文方案构造的陷门满足 IND-KGA 安全，关键字的安全性能够得到保障。

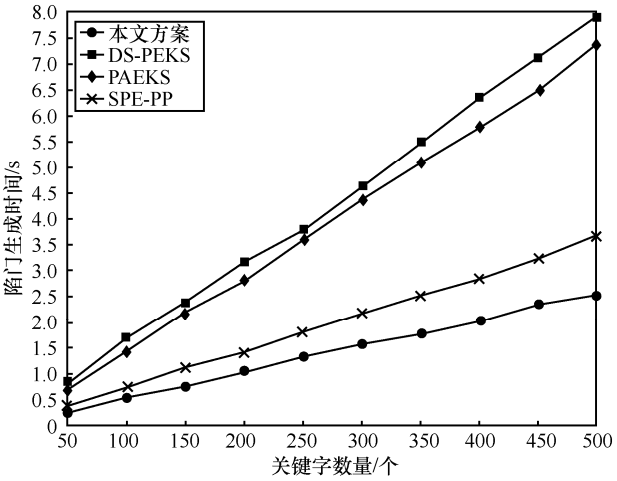

图 2 陷门生成时间

7.2 索引生成时间

由图 3 可知，本文方案与 DS-PEKS 和 PAEKS 方案相比有很大优势，与 SPE-PP 方案相比略有优势。造成这个结果的主要原因是本文方案中索引的计算只需要进行一次双线性计算和一次哈希计算，相比其他 3 种方案，本文方案是具有更简单的构造方案。与 DS-PEKS 和 PAEKS 方案相比，随着关键字的增多，本文方案的优势会越来越大。

7.3 关键字检索时间

本文方案在进行关键字检索时一共进行了 3 次双线性对计算，相比于其他 3 种方案，计算开销较小。由图 4 可知，在关键字数量为 500 个时，本文方案比 PAEKS 和 SPE-PP 方案的效率大约高 25%。

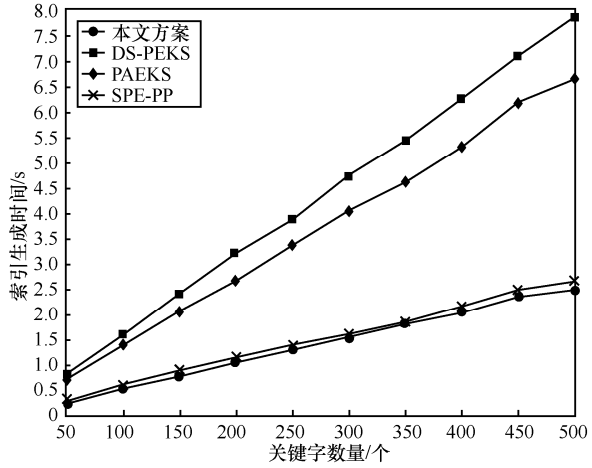

图 3　索引生成时间

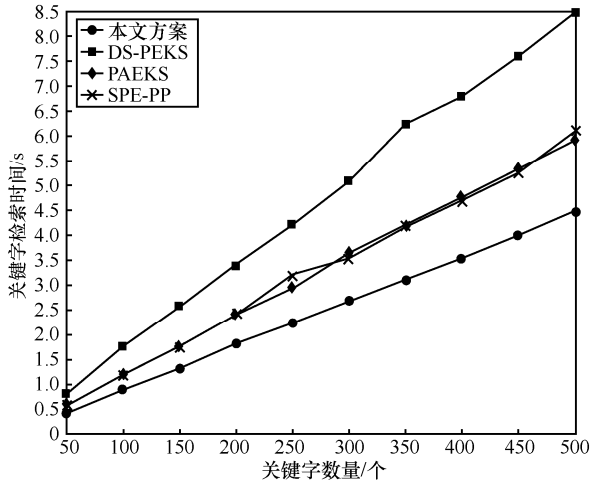

图 4　关键字检索时间

7.4　区块链引入前后对比实验

区块链引入之后会导致检索时间的增加，但是增强了安全性。由图 5 可知，随着关键字数量的增多，引入区块链方案的检索时间增长量逐渐减少。

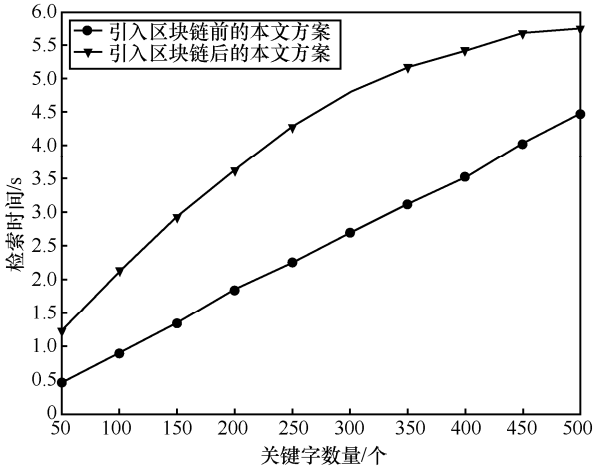

图 5　关键字检索时间

8　结束语

本文提出一种基于区块链的公钥可搜索加密方案，这是一种一对多的可搜索加密方案，主要应用于搭载在公共平台上的私有云环境。方案主要解决 2 个方面的问题：1) 保证陷门的安全性，利用 DBDH 困难问题构造原则使生成索引添加的随机数与生成陷门时添加的随机数不必相同，减少了用户与数据拥有者的通信资源开销，也预防了由信道安全引起的数据泄露问题；2) 利用区块链技术解决了传统方案中第三方的可信问题，同时利用区块链系统的公平公正特点，限制了服务器产生的恶意行为。安全性分析和挑战者游戏证明了本文方案的安全性。针对方案中索引生成时间、陷门生成时间、关键字检索时间进行了实验，关键字数量由 50 增大到 500，并对每个关键字数量各进行 50 次实验，对得到的时间开销取平均值，与文献[12-13,15]中的方案进行对比，证明了本文方案具有较高的效率。

接下来的工作将针对可搜索加密，利用区块链技术在公有环境中进行多对多模型的研究，虽然本文能够利用区块链的高可信度，扮演权威可信机构的角色，限制服务器的恶意行为，但是在安全性与效率方面还需要进行更深入的研究。

参考文献：

[1] DAWN S D, SONG D, WAGNER A P, et al. Practical techniques for searches on encrypted data[C]// Proceedings of the 2000 IEEE Security and Privacy Symposium. Piscataway: IEEE Press, 2000: 44-45.

[2] BONEH D, CRESCENZO G D, OSTROVSKY R, et al. Public key encryption with keyword search[C]// International Conference on the Theory and Applications of Cryptographic Techniques. Berlin: Springer, 2004: 506-522.

[3] 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.

[4] WANG P, WANG H, PIEPRZYK J. Threshold privacy preserving keyword searches[C]// Conference on Sofsem: Theory & Practice of Computer Science. Berlin: Springer, 2008: 646-658.

[5] YUAN K, LIU Z L, JIA C F, et al. Public key timed-release searchable encryption in one-to-many scenarios[J]. Tien Tzu Hsueh Pao/Acta Electronica Sinica, 2015, 43(4): 760-768.

[6] ZHONG H, CUI J, SHI R H, et al. Many-to-one homomorphic e-ncryption scheme[J]. Security & Communication Networks, 2016, 9(10):1007-1015.

[7] TANG Q, CHEN L Q. Public-key encryption with registered keyword search[C]// 6th European Workshop Public Key Infrastructures. Berlin: Springer, 2009: 163-178.

[8] FANG L M, SUSILO W, GE C, et al. Public key encryption with keyword search secure against keyword guessing attacks without random oracle[J]. Information Sciences, 2013, 238: 221-241.

[9] XU P, JIN H, WU Q, et al. Public-key encryption with fuzzy keyword search: a provably secure scheme under keyword guessing attack[J]. IEEE Transactions on Computers, 2013, 62(11): 2266-2277.

[10] CHEN R, MU Y, YANG G, et al. A new general framework for secure public key encryption with keyword search[C]// Australasian Conference on Information Security and Privacy. Berlin: Springer, 2015: 59-76.

[11] SHAO Z Y, YANG B. On security against the server in designated tester public key encryption with keyword search[J]. Information Processing Letters, 2015, 115(12):957-961.

[12] CHEN R, MU Y, YANG G, et al. Dual-server public-key encryption with keyword search for secure cloud storage[J]. IEEE Transactions on Information Forensics and Security, 2016, 11(4):789-798.

[13] HUANG Q, LI H. An efficient public-key searchable encryption scheme secure against inside keyword guessing attacks[J]. Information Sciences, 2017(403-404):1-14.

[14] KANG Y, LIU Z. A fully secure verifiable and out sourced decryption ranked searchable encryption scheme supporting synonym query[C]// IEEE Second International Conference on Data Science in Cyberspace. Piscataway: IEEE Press, 2017:223-231.

[15] WU L, CHEN B, ZEADALLY S, et al. An efficient and secure searchable public key encryption scheme with privacy protection for cloud storage[J]. Soft Computing, 2018, 22(23): 7685-7696.

[16] WU L B, ZHANG Y B, MA M M, et al. Certificateless searchable public key authenticated encryption with designated tester for cloud-assisted medical Internet of things[J]. Annales des Télécommunications, 2019, 74(7-8): 423-434.

[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] LU Y, LI J G. Efficient searchable public key encryption against keyword guessing attacks for cloud-based EMR systems[J]. Cluster Computing, 2019, 22(1): 285-299.

[19] LI H G, ZHANG F G, HE J J, et al. A searchable symmetric encryption scheme using blockchain[J]. CoRR: abs/1711.01030, 2017.

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

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

[作者简介]

杜瑞忠（1975– ），男，河北献县人，博士，河北大学教授、硕士生导师，主要研究方向为可信计算与信息安全。

谭艾伦（1995– ），男，四川广安人，河北大学硕士生，主要研究方向为可信计算与信息安全。

田俊峰（1964– ），河北蠡县人，博士，河北大学教授、博士生导师，主要研究方向为分布计算、可信计算与信息安全。