轻量级可搜索医疗数据共享方案

殷新春[1,2]，王梦宇[1]，宁建廷[3,4]

（1. 扬州大学信息工程学院，江苏 扬州 225127；2. 扬州大学广陵学院，江苏 扬州 225128；
3. 福建师范大学计算机与网络空间安全学院，福建 福州 350007；
4. 中国科学院信息安全国家重点实验室，北京 100093）

摘 要：支持策略隐藏和关键字搜索的属性基加密方案在医疗场景中具有良好的应用前景。然而，现有的此类方案大多不支持大属性域或采用“与门”结构，限制了访问控制的可扩展性和灵活性，并且许多方案无法抵抗离线字典猜测攻击。此外，属性基加密涉及大量的双线性配对运算，对于计算资源受限的用户设备来说使用非常不便。提出一种轻量级可搜索医疗数据共享方案。该方案在支持关键字搜索和策略隐藏的基础上采用大属性域和线性秘密共享结构，提高了访问控制的可扩展性和灵活性；采用 Intel SGX 技术对数据进行重加密，实现抗离线字典猜测攻击；将解密计算开销降低到恒定的常数级，适用于计算资源受限的用户设备。最后证明了所提方案具备选择明文不可区分安全性并且可以抵抗离线字典猜测攻击。

关键词：轻量级；策略隐藏；关键字搜索；属性基加密；大属性域；离线字典猜测攻击

中图分类号：TP309

文献标志码：A

DOI: 10.11959/j.issn.1000−436x.2022090

Lightweight searchable medical data sharing scheme

YIN Xinchun[1,2], WANG Mengyu[1], NING Jianting[3,4]

1. College of Information Engineering, Yangzhou University, Yangzhou 225127, China
2. Guangling College of Yangzhou University, Yangzhou 225128, China
3. College of Computer and Cyber Security, Fujian Normal University, Fuzhou 350007, China
4. State Key Laboratory of Information Security, Chinese Academy of Sciences, Beijing 100093, China

Abstract: The attribute-based encryption scheme supporting policy hiding and keyword search had a good application prospect in medical scenarios. However, most of the existing schemes did not support large attribute university or adopt the "AND gate" structure, which limited their scalability and flexibility of access control, and many schemes could not resist offline dictionary guessing attacks. In addition, attribute-based encryption involved a large number of bilinear pairing operations, which was inconvenient for user equipment with limited computing resources. A lightweight searchable medical data sharing scheme was proposed. Based on the support for keyword search and policy hiding, a large attribute university and a linear secret sharing structure were adopted to improve the scalability and flexibility of access control. The Intel SGX was used to re-encrypt data to achieve anti-offline dictionary guessing attack. The computational overhead of decryption was reduced to a constant level, which was suitable for user equipment with limited computing resources. Finally, it is proved that the proposed scheme has the security of selecting plaintext indistinguishable and can resist offline dictionary guessing attacks.

Keywords: lightweight, policy hiding, keyword search, attribute-based encryption, large attribute university, offline dictionary guessing attack

收稿日期：2022−01−19；**修回日期：**2022−03−19

基金项目：国家自然科学基金资助项目（No.62032005, No.61972094）；福建省自然科学基金资助项目（No.2020J02016）

Foundation Items: The National Natural Science Foundation of China (No.62032005, No.61972094), The Natural Science Foundation of Fujian Province (No.2020J02016)

0　引言

近年来，随着云计算和物联网技术的发展，电子医疗得到了人们的高度重视并处于迅猛发展的阶段[1]。通过使用无线传感器技术和通信网络，可穿戴医疗设备可以实时采集患者的健康数据并远程传输到云服务器（CS, cloud server）中，既节省了患者的时间和精力，也便于医护人员了解患者的身体健康状况并采取相应的诊疗措施。然而，由于数据中包含患者的身份、病症和病史等隐私信息，一旦泄露会给患者带来极大的安全隐患[1]。因此，患者的健康数据必须以密文的形式存储在系统中。然而，在密态数据中进行检索是困难的。因此，如何对加密数据进行检索成为一项特别具有挑战性的工作[2]。

一种简单的解决方法是患者或者医生提前下载数据到本地，解密后再一一检索，但这种方式的效率太低。可搜索加密（SE, searchable encryption）[3-4]的提出有效地解决了这一问题。在可搜索加密系统中，数据所有者（DO, data owner）根据数据中的信息生成相应的索引，与加密数据一起存储在云服务器中。当用户访问数据时，构建一个搜索令牌并将其提交给服务器。服务器利用搜索令牌执行搜索，但无法获取除密文和索引外任何与搜索内容相关的信息。因此，可搜索加密使用户可以在未解密数据的情况下搜索加密数据，既保证了数据的安全，也提高了检索的效率。但部分传统的可搜索加密方案[5-7]不支持细粒度访问控制，并不适用于医疗数据共享的应用场景。例如，在日常生活中，每个医生诊治不同的患者，每个患者可以选择不同科室和不同级别的医生就诊。因此，细粒度访问控制在电子医疗系统中是非常必要的。

属性基加密（ABE, attribute-based encryption）[8]可以有效地实现细粒度和非交互式的访问控制机制。ABE 可以分为密钥策略的属性基加密（KP-ABE, key-policy attribute-based encryption）[9]和密文策略的属性基加密（CP-ABE, ciphertext-policy attribute-based encryption）[10]。在 KP-ABE 机制中，访问策略和属性集合分别嵌入密钥和密文中；在 CP-ABE 机制中，访问策略和属性集合分别嵌入密文和密钥中。相比之下，CP-ABE 更加适用于云存储与细粒度数据共享。基于密文策略属性基加密的关键字搜索（CP-ABKS, ciphertext-policy attribute- based encryption keyword searchable）方案在细粒度数据共享的基础上实现了关键字搜索，可以满足上述基本需求。Zheng 等[11]提出了一种可验证的 CP-ABKS 方案，该方案支持搜索验证，但该方案采用效率较低的树形结构且无法抵抗关键字猜测攻击。Sun 等[12]提出了一种支持多用户的可验证 CP-ABKS 方案，该方案支持用户撤销并被证明可以抵抗选择关键字攻击。在属性基加密系统中，访问策略用于指定用户的访问权限。在大多数属性基加密方案中，为了实现解密操作，访问策略通常需要显式地附加到密文中。但是，访问策略通常会包含一些敏感信息，云服务器或恶意用户可能通过访问策略推断这些信息。然而，上述方案均不支持策略隐藏。

针对访问策略可能会泄露数据所有者敏感信息的问题，Nishide 等[13]提出了一种支持部分策略隐藏的解决方案，使用多值通配符来表示所要隐藏的属性，但该方案仅支持“与门”结构，灵活性差且计算开销较大。为了提高访问策略的灵活性，Lai 等[14]采用线性秘密共享方案（LSSS, linear secret sharing scheme）结构，提出了一种基于合数阶群的部分策略隐藏方案并证明该方案是完全安全的，但其运算效率不高。随后，Cui 等[15]在 Lai 等[14]方案的基础上提出了一种基于素数阶群支持访问策略隐藏的方案，但该方案的计算开销较大且不支持解密测试，解密效率较低。此外，上述方案均不支持关键字搜索。Qiu 等[16]提出了支持策略隐藏和抗关键字猜测攻击的 CP-ABKS 方案，并且该方案可以限制未授权用户的搜索，但该方案未涉及数据加密，仅支持单个数据所有者。随后，Wang 等[17]进一步提出了支持策略隐藏的可搜索和可撤销的数据所有者属性基加密方案，该方案将同一访问策略嵌入密文和关键字中，实现多所有者数据共享。Miao 等[18]提出了支持多所有者合作且具备隐私保护的关键字搜索方案，该方案采用多所有者合作加密模式，支持策略隐藏和用户追踪。但上述 3 种方案均是基于 Nishide 等[13]方案实现的部分策略隐藏，采用“与门”结构且不支持大属性域，访问控制的灵活性和可扩展性较低。Zhang 等[19]采用交互式的在线隐私保护测试，提出一种针对属性值猜测攻击的部分策略隐藏方案，可以抵抗离线字典猜测攻击，并且证明了 Nishide 等[13]方案实际上无法抵抗离线字典猜测攻击，侧面说明 Qiu 等[16]、Wang 等[17]

和 Miao 等[18]方案均无法抵抗离线字典猜测攻击。

考虑到云服务器是不完全可信的，并且存在系统故障和黑客攻击的安全隐患，用户收到数据的完整性和正确性是有待验证的。Miao 等[20]提出针对动态数据所有者的可验证多关键字搜索加密数据方案，该方案支持搜索结果验证，但验证过程需要设置私人审计服务器，并且需要与服务器交互才可验证。不仅增加了通信和计算开销，且该方案不支持策略隐藏。

为了解决上述问题，本文提出了一种轻量级可搜索医疗数据共享方案，采用大属性域、LSSS 结构和 Intel SGX（software guard extension）技术，实现了关键字搜索、策略隐藏、数据验证和抗离线字典猜测攻击的功能。本文方案在降低计算开销的同时，保证了数据的安全性。

本文主要研究工作围绕系统的功能、开销和安全 3 个方面展开，具体内容如下。

1) 提高可扩展性和灵活性。本文方案支持大属性域，系统不需要预先设定属性域，不仅提高了访问控制的可扩展性，也降低了系统在初始化阶段的开销。此外，本文方案采用 LSSS 结构，相较于传统的“与门”结构，具备更加灵活的访问控制能力。

2) 降低开销。本文方案实现了关键字搜索和策略隐藏的功能，将关键字与属性值绑定，在执行数据搜索的同时验证了用户的解密能力，防止未授权用户的访问。相较于传统的策略隐藏方案，本文方案不需要生成测试密文和测试密钥即可验证用户的解密能力，降低了加密和密钥生成阶段的开销，并且用户只需要常数级的计算即可解密密文，适用于计算资源受限的用户设备。此外，本文方案采用非交互式数据验证，避免用户与服务器的交互，降低了数据验证的开销。

3) 保障安全性。由于服务器可能会根据属性名、应用场景等信息预设出一个离线字典，猜测出索引和搜索令牌所包含的属性值和关键字。本文方案采用 Intel SGX 技术，在系统中开辟一个安全容器 Enclave，对数据进行重加密，改变了数据的结构，从而实现抗离线字典猜测攻击，同时避免用户与服务器在搜索过程中的交互，适用于可搜索加密方案。安全性分析证明了本文方案具备选择明文攻击不可区分（IND-CPA, indistinguishable choose plaintext attack）安全性并且可以抵抗离线字典猜测攻击。

1 预备知识

1.1 访问结构

设 $\{P_1,P_2,\cdots,P_n\}$ 是由 n 个参与者组成的实体集，对于集合 $A\subseteq 2^{\{P_1,P_2,\cdots,P_n\}}$，如果 $\forall B,C,B\in A,B\subseteq C$，有 $C\in A$，那么称 A 是单调的。如果集合 A 是 $\{P_1,P_2,\cdots,P_n\}$ 的非空子集，即 $A\subseteq 2^{\{P_1,P_2,\cdots,P_n\}}\setminus\{\varnothing\}$，那么 A 就是一个访问结构，所有包含在 A 中的集合称为授权集合，不包含在 A 中的集合称为非授权集合。

1.2 双线性映射

令 G 和 G_T 为 2 个阶为素数 p 的循环群，g 为群 G 的生成元，$e:G\times G\to G_T$ 为双线性映射，其中双线性映射 e 满足如下条件。

1) 双线性：对于 $g\in G,a,b\in Z_p$，有 $e(g^a,g^b)=e(g,g)^{ab}$。

2) 非退化性：$e(g,g)\neq 1$。

如果群 G 上的计算和双线性映射 $e:G\times G\to G_T$ 可以有效地执行，那么称群 G 是一个双线性群。注意到映射 e 是对称的，因为 $e(g^a,g^b)=e(g,g)^{ab}=e(g^b,g^a)$。

1.3 LSSS

设 p 是一个大素数，U 是一个属性集。对于 U 上的每个单调访问结构，都可以找到一个矩阵 $\boldsymbol{M}_{l\times n}$，该矩阵有 l 行 n 列，称为共享生成矩阵。$(\boldsymbol{M}_{l\times n},\rho)$ 是访问结构，ρ 是一个将矩阵 $\boldsymbol{M}_{l\times n}$ 的每一行映射到一个属性 $\mathrm{Attr}\in U$ 的映射函数。其中，$i\in[1,l]$。线性秘密共享方案由以下 2 种算法组成。

1) 秘密分享。当 $\boldsymbol{M}_{l\times n}$ 共享一个秘密值 $s\in Z_p$ 时，首先设置 n 维列向量 $\vec{\boldsymbol{v}}=(s,y_2,y_3,\cdots,y_n)^{\mathrm{T}}$，其中 $y_2,y_3,\cdots,y_n\in Z_p$。该算法计算 s 的共享向量 $\vec{\boldsymbol{\lambda}}=\boldsymbol{M}_{l\times n}\vec{\boldsymbol{v}}$，即 $\lambda_i=\boldsymbol{M}_i\vec{v}$，其中 $i\in[1,l]$，$\boldsymbol{M}_i$ 表示 $\boldsymbol{M}_{l\times n}$ 的第 i 行，λ_i 表示属性名索引 $\rho(i)$ 所持有的共享值。

2) 秘密值重构。设 A 是任意授权集合，$I=\{i\mid\rho(i)\in A\}\subseteq\{1,2,\cdots,l\}$，存在常数集 $\{\omega_i\in Z_p\}_{1\leqslant i\leqslant I}$，使 $\sum_{i\in I}\omega_i\boldsymbol{M}_i=(1,0,\cdots,0)$，从而有 $\sum_{i\in I}\omega_i\lambda_i=s$。如果 A' 是未经授权的集合，并且 $I'=\{i\mid\rho(i)\in A'\}$，存在系数元组 $\{\omega'_i\}_{i\in I'}$ 且 ω'_1 为非零元素，使方程 $\sum_{i\in I'}\omega_i\lambda_i=0$。如果集合 $\{1,2,\cdots,l\}$ 的一个子集 I 满足 $(\boldsymbol{M}_{l\times n},\rho)$，并且任意 $I'\subset I$ 都不满

足 $(\boldsymbol{M}_{l\times n},\rho)$，那么称 I 为 $(\boldsymbol{M}_{l\times n},\rho)$ 的最小授权集。

1.4　基于 LSSS 的部分隐藏结构

将一个属性分为 2 个部分，与矩阵 $\boldsymbol{M}_{l\times n}$ 第 i 行相关联，即 $\mathrm{Attr}_i \to \{c_i, v_i\}$。其中，$c_i$ 表示属性名，v_i 表示属性值。存在 2 个映射函数 ρ 和 π。其中，ρ 可以将矩阵第 i 行映射到一个属性名，即 $\rho(i)\to c_i$；π 可以将矩阵第 i 行映射到一个属性值，即 $\pi(i)\to v_i$。$(\boldsymbol{M}_{l\times n},\rho)$ 是以明文形式存在于访问结构中的，π 是被隐藏在密文中的。所以，$((\boldsymbol{M}_{l\times n},\rho),\pi)$ 作为本文方案的部分隐藏访问结构。

1.5　困难性问题

1) q-parallel BDHE（q-parallel bilinear Diffie-Hellman exponent）假设

给定 2 个阶为素数 p 的乘法循环群 G, G_T，g 是 G 的生成元，双线性映射 $e: G\times G\to G_T$。随机选取 $a, s, \{o_j\}_{j=1}^{q}\in Z_p$，给出 T，$\vec{\boldsymbol{p}}=(G, e, p, g, g^s, \{g^{a^i}\}_{1\leqslant i\leqslant 2q, i\neq q+1}, \{g^{so_j}\}_{i\leqslant j\leqslant q}, \{g^{\frac{b^i}{o_j}}\}_{1\leqslant i\leqslant 2q, i\neq q+1, 1\leqslant j\leqslant q}, \{g^{\frac{b^i so_k}{o_j}}\}_{1\leqslant i,k\leqslant q, k\neq j})$。

对于任意多项式时间攻击者来说，区分 $T=e(g,g)^{sa^{q+1}}$ 和 $T=R$ 是困难的。其中 $R\in G_T$。

2) CDH（computational Diffie-Hellman problem）假设

给定一个 p 阶的乘法循环群 G，g 是 G 的生成元。随机选择 2 个元素 $a, b\in Z_p$。CDH 假设可表述为：通过 g^a 和 g^b 来计算 g^{ab} 是困难的。

1.6　Intel SGX

Intel SGX 是在原有 Intel 架构上扩展的一组新的指令集和内存访问机制[21]，允许应用程序创建一个叫做 Enclave 的隔离执行环境。Enclave 作为一个可信和安全的实体，用来存储数据和执行代码。Enclave 具有 3 个安全特性：隔离、密封和认证[22]。隔离限制了对硬件保护的内存区域的访问，只有特定的 Enclave 可以访问它。同一处理器上的任何其他进程，甚至是操作系统、管理程序等其他外部实体，都不能访问该内存。密封提供了一个将 Enclave 秘密加密到磁盘上的持久存储的方法，以便即使 Enclave 被拆毁也能检索秘密。加密是使用特定 Enclave 私有的密封密钥执行的，除了完全相同的 Enclave 之外，没有任何进程可以对其解密或修改。认证可以使验证者验证检测代码是否在 Enclave 内安全运行且未被修改。SGX 提供 2 种身份认证方式，分别是本地认证和远程认证[23]。本地认证用于同一平台上的 2 个 Enclave 之间的认证，同一平台上的 2 个 Enclave 可以使用它们之间共享的根密封密钥派生一个共享密钥。远程认证使 Enclave 能够生成任何远程实体都可以验证的报告。

2　系统定义

2.1　系统模型

本文提出了一种的轻量级可搜索医疗数据共享方案。系统模型如图 1 所示，该模型由 5 个实体组成：授权中心（AC, authorization center）、云服务器、Enclave、数据所有者和数据用户（DU, data user）。每个实体在该系统中具有不同的分工，具体如下。

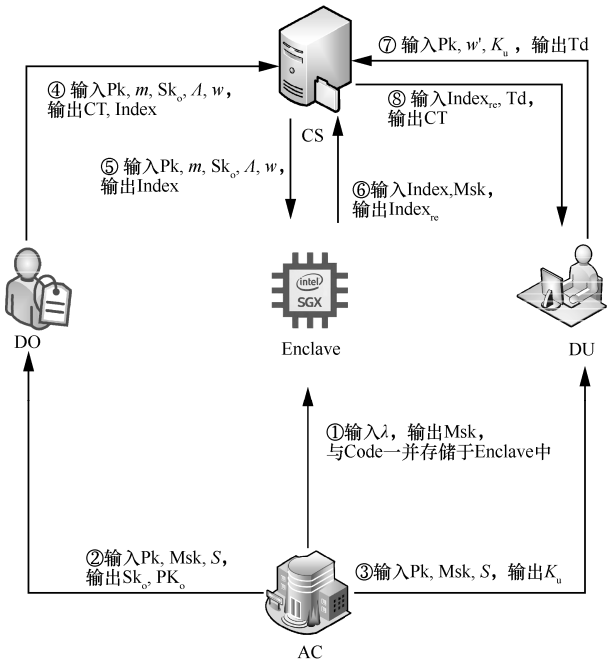


图 1　系统模型

授权中心。AC 是完全可信的，负责初始化系统公私钥、发布密钥。

云服务器。CS 是“诚实且好奇”的半可信服务器，负责存储 DO 的数据，处理 DU 的搜索请求。CS 会按协议诚实地执行任务，但也会尽可能地收集用户的敏感数据。

Enclave。Enclave 是完全可信的，是通过 Intel SGX 技术在计算机内存中创建的一个隔离环境，即使是在不可信的主机上，它也可以安全地执行程序和保存数据。在本文方案中 Enclave 被用来存储系

统主密钥 Msk 和重加密代码 Code。

数据所有者。DO 是诚实的，负责加密数据和关键字，并将加密后的数据和索引上传到 CS。

数据用户。DU 是不可信的，被授权的 DU 负责生成搜索令牌，解密共享数据，验证解密的正确性。未授权的 DU 可能会与其他用户共谋，获取自身无法访问的数据。

2.2 算法定义

本文方案由以下 7 种算法组成。

1) 系统初始化算法。$\text{Setup}(\lambda)\to(\text{Pk},\text{Msk})$：该算法由 AC 执行，输入安全参数 λ，输出系统公钥 Pk，保存系统主密钥 Msk。

2) 密钥生成算法。$\text{KeyGen}(\text{Pk},\text{Msk},S)\to(\text{Sk}_\text{o},\text{Pk}_\text{o},K_\text{u})$：该算法由 AC 执行，输入系统公钥 Pk、主密钥 Msk、DU 的属性集合 S，输出 DO 的私钥 Sk_o、公钥 Pk_o 和 DU 的密钥 K_u，通过安全信道传送给 DU 和 DO。

3) 加密算法。$\text{Encrypt}(\text{Pk},m,\text{Sk}_\text{o},\Lambda,w)\to(\text{CT},\text{Index})$：该算法由 DO 执行，输入系统公钥 Pk、消息 m、DO 的私钥 Sk_o、访问策略 Λ、关键字 w，输出密文 CT 和关键字索引 Index。

4) 重加密算法。$\text{ReEncrypt}(\text{Index},\text{Msk})\to\text{Index}_\text{re}$：该算法由 Enclave 执行，输入关键字索引 Index 和部分系统主密钥 Msk，输出重加密索引 Index_re。

5) 搜索令牌生成算法。$\text{TokenGen}(\text{Pk},w',K_\text{u})\to(\text{Td},\text{Sk}_\text{u})$：该算法由 DU 执行，输入系统公钥 Pk、关键字 w'、DU 的密钥 K_u，输出搜索令牌 Td 和 DU 的私钥 Sk_u。

6) 搜索算法。$\text{Search}(\text{Index}_\text{re},\text{Td})\to\text{CT or }\perp$：该算法由 CS 执行，输入重加密索引 Index_re、搜索令牌 Td。如果搜索成功，输出密文 CT；否则，算法终止。

7) 解密算法。$\text{Decrypt}(\text{CT},\text{Sk}_\text{u})\to m\text{ or }\perp$：该算法由 DU 执行，输入 DU 的私钥 Sk_u 和密文 CT，如果解密结果正确，输出消息 m；否则，算法终止。

2.3 安全模型

本文提出的方案具备选择明文攻击不可区分安全性和抵抗离线字典猜测攻击安全性。

2.3.1 选择明文攻击不可区分安全性

通过攻击者 A 和挑战者 C 之间的安全游戏来定义选择明文攻击不可区分安全性模型，其详细描述如下。

初始化阶段。攻击者 A 提交一个访问策略 $\Lambda=((\boldsymbol{M}_{l\times n},\rho),\pi)$ 和一个关键字 w 给挑战者 C。

系统建立阶段。挑战者 C 先运行 Setup 算法，输入安全参数 λ，然后将公共参数 Pk 发送给攻击者 A 并保存主密钥 Msk。

查询阶段 1。攻击者 A 发起关于一系列属性集 $\text{AS}=\{S_1,S_2,\cdots,S_\tau\}$ 和关键字集 $W=\{w_1,w_2,\cdots,w_\tau\}$ 的密钥和搜索令牌请求。挑战者 C 运行 KeyGen 和 TokenGen 算法将生成的密钥和搜索令牌发送给攻击者 A。其中，AS 的所有属性集均不满足访问策略 Λ 且 $w\notin W$。

挑战阶段。攻击者 A 向挑战者 C 提交 2 个等长的明文消息 m_0,m_1。挑战者 C 随机投掷一枚硬币 $b\in\{0,1\}$，利用访问策略 Λ、关键字 w 和消息 m_b，运行 Encrypt 和 ReEncrypt 算法生成密文 CT、索引 Index 和重加密索引 Index_re，将 CT 和 Index_re 发送给攻击者 A。

查询阶段 2。与查询阶段 1 相同。

猜测阶段。攻击者 A 输出猜测 $b\in\{0,1\}$。如果 $b'=b$，攻击者 A 将赢得游戏。

攻击者 A 赢得这个游戏的优势定义为 $\text{Adv}_A=|\Pr[b=b']-\frac{1}{2}|$。若攻击者 A 不能够在概率多项式时间（PPT, probabilistic polynomial time）内以不可忽略的优势 Adv_A 打破上述安全性游戏，说明方案具备选择明文攻击不可区分安全性。

2.3.2 抗离线字典猜测攻击安全性

离线字典猜测攻击通常是针对关键字或属性值发起的。在本文方案中，关键字和属性值是嵌入关键字索引和搜索令牌中的。所以，攻击者会针对这两部分发起离线字典猜测攻击。首先，攻击者在多项式时间内找到一个判别式。然后，攻击者将离线字典中的关键字和属性值放入判别式中，以检查判别式是否成立。如果成立，说明攻击者可以打破抗离线字典猜测攻击安全；否则，说明本文方案具备抵抗离线字典猜测攻击安全性。

3 方案设计

本文方案所涉及的符号定义如表 1 所示。

1) $\text{Setup}(\lambda)$。该算法由 AC 执行。首先，该算法输入安全参数 λ，输出一组双线性对密码参数 (G,G_T,e,p,g,g_1)，其中 G 和 G_T 是阶为素数 p 的乘法循环群，e 是一个双线性映射，g,g_1 是群 G 的 2 个生成元。然后，该算法随机选择 $\alpha,a,d,e,k,z\in Z_p$ 和

表 1　　符号定义

参数	含义
p	大素数
G, G_T	p 阶循环群
g	群 G 的生成元
e	双线性映射
Z_p	p 阶整数群
$\Lambda=((\boldsymbol{M}_{l\times n},\rho),\pi)$	$\boldsymbol{M}_{l\times n}$ 为访问矩阵，ρ 为属性名映射，π 为属性值映射
$S=\{c_i,v_i\}$	c_i 为属性名，v_i 为属性值
$H:\{0,1\}^*\to G$	将字符映射为群 G 元素的哈希函数
$H_1:\{0,1\}^*\to Z_p$	将字符映射为群 Z_p 元素的哈希函数

2 个哈希函数 $H:\{0,1\}^*\to G, H_1:\{0,1\}^*\to Z_p$，计算系统公钥和系统主密钥分别为

$$\mathrm{Pk}=(g,g_1,e(g,g)^{\alpha},g^a,g^d,g^e,H,H_1)$$
$$\mathrm{Msk}=(\alpha,a,d,e,k,z)$$

AC 公布 Pk，保存 Msk 并将其中的 z 预存在 Enclave 中。

2) KeyGen(Pk, Msk, S)。该算法由 AC 执行，分别为 DO 和 DU 生成相关的密钥。

a) $\mathrm{KeyGen}_{\mathrm{DO}}$。该算法随机选择 $\gamma\in Z_p$，计算 DO 的公私钥 $\mathrm{Sk}_{\mathrm{o}}=\gamma$，$\mathrm{Pk}_{\mathrm{o}}=g_1^{\gamma}$。

b) $\mathrm{KeyGen}_{\mathrm{DU}}$。该算法随机选择 $t,t'\in Z_p$。$S=\{c_i,v_i\}_{i=1}^h$ 是属性集，其中 $\{c_i\}_{i=1}^h$ 表示属性名集，$\{v_i\}_{i=1}^h$ 表示属性值。计算 DU 的密钥为

$$K_u=(K_1,K_2,K_3,\{K_i\}_{i=1}^h)$$

其中，

$$K_1=g^{\alpha}g^{akdt}$$
$$K_2=g^{kdt}$$
$$K_3=g^{\frac{akdt}{e}}$$
$$\{K_i=H(v_i)^{ktz}\}_{i=1}^h$$

3) Encrypt(Pk, m, Sk_{o}, Λ, w)。该算法由 DO 执行，分别对数据和关键字加密，生成密文 CT 和关键字索引 Index。该算法选择一个访问策略 $\Lambda=((\boldsymbol{M}_{l\times n},\rho),\pi)$ 和向量 $\vec{\boldsymbol{v}}=(s,y_2,\cdots,y_n)^{\mathrm{T}}\in Z_p^n$。其中，$\rho$ 和 π 是 2 个映射函数，分别将矩阵第 i 行映射到一个属性名和属性值，s 是用于分享的随机秘密值，计算 $\lambda_i=\boldsymbol{M}_i\vec{\boldsymbol{v}}$。

a) $\mathrm{Encrypt}_{\mathrm{Data}}$。该算法选择一个明文消息 m，利用私钥 Sk_{o} 计算验证参数 $x=g_1^{\frac{1}{H_1(m)+\gamma}}$。计算密文为

$$\mathrm{CT}=(y,C_0,C_1)$$

其中，

$$y=e(g_1,g_1)^{\frac{1}{H_1(m)+\gamma}}$$
$$C_0=(x\,\|\,m)e(g,g)^{\alpha s}$$
$$C_1=g^s$$

b) $\mathrm{Encrypt}_{\mathrm{Index}}$。该算法选择关键字 w，随机选择 $r_i\in Z_p$，计算索引为

$$\mathrm{Index}=(I_1,\{I_{i,1},I_{i,2}\}_{i=1}^l)$$

其中，

$$I_1=g^{es},I_{i,1}=g^{dr_i},I_{i,2}=g^{a\lambda_i}H(\pi_i)^{\frac{-r_i}{H_1(w)}}$$

4) ReEncrypt(Index, Msk)。该算法由 Enclave 执行。CS 将 $I_{1,i}$ 发送给 Enclave。Enclave 调用预存的系统主密钥 z，计算 $I_{i,1}{*}=I_{i,1}^{\frac{1}{z}}=g^{\frac{dr_i}{z}}$。

然后，将其返回 CS。CS 更新索引为

$$\mathrm{Index}_{\mathrm{re}}=(I_1,\{I_{i,1}{*},I_{i,2}\}_{i=1}^l)$$

5) TokenGen(Pk, w', K_u)。该算法由 DU 执行。该算法选择关键字 w'，随机选择 $\theta\in Z_p$，令 $\mathrm{Sk}_u=\theta$，计算搜索令牌 $\mathrm{Td}=(\mathrm{Td}_1,\mathrm{Td}_2\{\mathrm{Td}_i\}_{i=1}^h)$。其中，

$$\mathrm{Td}_1=K_2^{\theta}=g^{kdt\theta}$$
$$\mathrm{Td}_2=K_3^{\theta}=g^{\frac{akdt\theta}{e}}$$
$$\mathrm{Td}_i=K_i^{\frac{\theta}{H_1(w')}}=H(v_i)^{\frac{ktz\theta}{H_1(w')}}$$

6) Search($\mathrm{Index}_{\mathrm{re}}$, Td)。该算法由 CS 执行。DU 将搜索令牌 Td 发送给 CS。如果 DU 的属性名集合满足访问策略，CS 取 $I=\{i\,|\,\rho(i)\in\{c_j\}_{j=1}^h\}\subset\{1,2,\cdots,l\}$，计算 $\omega_i\in Z_p$，其中 $\sum_{i\in I}\omega_i M_i=(1,0,\cdots,0)$。然后，执行如下搜索计算

$$E_1=e(I_1,\mathrm{Td}_2)=e(g,g)^{akdt\theta s}$$
$$E_2=\prod_{i\in I}(e(I_{i,2},\mathrm{Td}_1)e(I_{i,1}{*},\mathrm{Td}_i))^{\omega_i}=$$
$$\prod_{i\in I}\left(e(g^{a\lambda_i}H(\pi_i)^{\frac{-r_i}{H_1(w)}},g^{kdt\theta})e\left(g^{\frac{dr_i}{z}},H(v_i)^{\frac{ktz\theta}{H_1(w')}}\right)\right)^{\omega_i}=$$

$$\prod_{i\in I}\left(e(g,g)^{akdt\theta\lambda_i}e(H(\pi_i),g)^{\frac{-r_i kdt\theta}{H_1(w)}}e\left(g,H(v_i)^{\frac{r_i kdt\theta}{H_1(w')}}\right)^{\omega_i}\right)=$$

$$e(g,g)^{akdt\theta\sum_{i\in I}\omega_i\lambda_i}=e(g,g)^{akdt\theta s}$$

如果 $E_1=E_2$，说明 CS 搜索密文 CT 成功，将 CT和E_1发送给 DU；否则，算法终止。

7) Decrypt(CT,$\mathrm{Sk_u}$)。该算法由 DU 执行。

首先，DU 计算

$$x\,\|\,m=\frac{C_0E_1^{\frac{1}{\theta}}}{e(C_1,K_1)}$$

然后，验证如下等式

$$y=e(g_1,x)$$

$$e(g^{H_1(m)}\mathrm{Pk_o},x)=e(g_1,g_1)$$

如果上述等式成立，说明结果正确，输出消息 m；否则，算法终止。

4 安全性分析

4.1 选择明文不可区分安全性证明

定理 1 如果 q-parallel BDHE 假设成立，则本文方案是选择明文不可区分安全的。

证明 假设存在一个 PPT 攻击者 A，具有不可忽略的优势 Adv_A 攻破本文方案。利用 q-parallel BDHE 假设条件 $\{\vec{\boldsymbol{p}},T\}$，构建一个与攻击者 A 在安全游戏中交互的挑战者 C。在存在攻击者 A 的情况下，证明挑战者 C 在打破 q-parallel BDHE 假设方面具有不可忽略的优势。

初始化阶段。挑战者 C 接收 q-parallel BDHE 假设的条件 $\{\vec{\boldsymbol{p}},T\}$、攻击者 A 提交的一个访问策略 $\Lambda=((\boldsymbol{M}_{l\times n},\rho),\pi)$ 和一个关键字 w，其中，$l,n<q$。

系统设置阶段。挑战者 C 随机选择 $\alpha,a,d,e,k,z,d_1\in Z_p$ 并选择 2 个哈希映射函数 $H:\{0,1\}^*\to G_1,H_1:\{0,1\}^*\to Z_p$。挑战者 C 计算 $e(g^a,g^{a^q})e(g,g)^{\alpha'}=e(g,g)^{\alpha},g_1=g^{d_1}$ 并隐式设置 $\alpha=\alpha'+a^{q+1}$。挑战者 C 将 $\{g,g_1,e(g,g)^{\alpha},g^a,g^d,g^e\}$ 设置为系统公钥。对于每一个属性 $\{c,v\}$，挑战者 C 随机选择 $\tilde{v}\in Z_p$，按照如下规则设计 H。

如果 $I\neq\varnothing$，挑战者 C 返回 $H(v)=g^{\tilde{v}}\prod_{i\in I}(g^{\frac{a}{o_i}})^{M_{i,1}}(g^{\frac{a^2}{o_i}})^{M_{i,2}}\cdots(g^{\frac{a^n}{o_i}})^{M_{i,n}}$。如果 $I=\varnothing$，挑战者 C 返回 $H(V)=g^{\tilde{v}}$，其中 $I=\{i\,|\,\{\rho_i,\pi_i\}=\{c,v\}\}_{i=1}^{l}$。

查询阶段 1。查询阶段分为密钥查询和搜索令牌查询。攻击者 A 提交一系列属性集 $\mathrm{AS}=\{S_1=\{c_i,v_{1,i}\}_{i=1}^{h},S_2=\{c_i,v_{2,i}\}_{i=1}^{h},\cdots,S_\tau=\{c_i,v_{\tau,i}\}_{i=1}^{h}\}$ 和关键字集 $W=\{w_1,w_2,\cdots,w_\tau\}$，分别向挑战者 C 发起密钥和搜索令牌请求。其中，AS 中所有属性集均不满足访问策略 Λ 且 $w\notin W$。

密钥查询。攻击者 A 选取一个属性集 $S_x\in\mathrm{AS}$ 向挑战者 C 发起密钥请求。挑战者 C 选择一个向量 $\vec{\boldsymbol{\omega}}=(\omega_1,\omega_2,\cdots,\omega_n)^{\mathrm{T}}$。其中 $\omega_1=-1$，$\boldsymbol{M}_i\vec{\boldsymbol{\omega}}=0$，$I'=\{i\,|\,\{\rho_i,\pi_i\}\in S_x\}_{i=1}^{l}$。挑战者 C 随机选择 $\tilde{t}\in Z_p$ 并隐式设置 $t=\tilde{t}+\omega_1a^q+\omega_2a^{q-1}+\cdots+\omega_na^{q-n+1}$。然后，计算密钥

$$K_1=g^{\alpha}g^{adkt}=$$

$$g^{\alpha'}g^{a^{q+1}}(g^a)^{\tilde{t}dk}\left(\prod_{i=1}^{n}(g^{a^{q-i+2}})^{\omega_i}\right)^{dk}=$$

$$g^{\alpha'}g^{a^{q+1}}(g^a)^{\tilde{t}dk}g^{-a^{q+1}}\left(\prod_{i=2}^{n}(g^{a^{q-i+2}})^{\omega_i}\right)^{dk}=$$

$$g^{\alpha'}(g^a)^{\tilde{t}dk}\left(\prod_{i=2}^{n}(g^{a^{q-i+2}})^{\omega_i}\right)^{dk}$$

$$K_2=g^{dkt}=g^{\tilde{t}dk}\left(\prod_{i=1}^{n}(g^{a^{q-i+1}})^{\omega_i}\right)^{dk}$$

$$K_3=g^{\frac{adkt}{e}}=(g^a)^{\frac{\tilde{t}dk}{e}}\left(\prod_{i=1}^{n}(g^{a^{q-i+1}})^{\omega_i}\right)^{\frac{dk}{e}}$$

假设 $\{c_\eta,v_{x,\eta}\}\in S_x,I_{x,\eta}=\{i\,|\,v_i=v_{x,\eta}\}_{i=1}^{l}\subseteq I'$，计算

$$K_\eta=H(v_{x,\eta})^{ktz}=$$

$$(g^t)^{kz\tilde{v}_i}\prod_{i=1}^{I_{x,\eta}}\prod_{j=1}^{n}\left(g^{\frac{a^i}{o_i}}\right)^{M_{i,j}\tilde{t}kz}$$

$$\left(\prod_{i=1}^{I_{x,\eta}}\prod_{j=1}^{n}\prod_{\xi=1}^{n}\left(g^{\frac{a^{q-\xi+1+j}}{o_i}}\right)^{\omega_\xi M_{i,j}\tilde{t}}\right)^{kz}$$

当 $\xi=j$ 时，

$$\prod_{i=1}^{I_{x,\eta}}\prod_{j=1,\xi=j}^{n}\left(g^{\frac{a^{q+1}}{o_i}}\right)^{\omega_\xi M_{i,j}}=\prod_{i=1}^{I_{x,\eta}}\left(g^{\frac{a^{q+1}}{o_i}}\right)^{\sum_{j=1}^{n}\omega_jM_{i,j}}=1$$

所以，挑战者 C 计算 K_η 为

$$K_\eta=(g^t)^{kz\tilde{v}_{x,\eta}}\prod_{i=1}^{I_{x,\eta}}\prod_{j=1}^{n}\left(g^{\frac{a^i}{o_i}}\right)^{M_{i,j}kz}\cdot$$
$$\left(\prod_{i=1}^{I_{x,\eta}}\prod_{j=1}^{n}\prod_{\xi=1,\xi\neq j}^{n}\left(g^{\frac{a^{q-\xi+1+j}}{o_i}}\right)^{\omega_\xi M_{i,j}}\right)^{kz}$$

最后，挑战者 C 将密钥 $K_1,K_2,K_3,\{K_\eta\}_{\eta=1}^{h}$ 发送给攻击者 A。

搜索令牌查询。攻击者 A 选取一个关键字 $w_x\in W$ 向挑战者发起搜索令牌请求。挑战者 C 随机选择 $\theta\in Z_p$，计算搜索令牌

$$\mathrm{Td}_1=K_2^\theta=g^{kdt\theta},\ \mathrm{Td}_2=K_3^\theta=g^{\frac{akdt\theta}{e}},\ \mathrm{Td}_n=K_\eta^{\frac{\theta}{H_1(w_x)}}$$

最后，挑战者 C 将搜索令牌 $\mathrm{Td}_1,\mathrm{Td}_2,\{\mathrm{Td}_\eta\}_{\eta=1}^{h}$ 发送给攻击者 A。

挑战阶段。攻击者 A 提交 2 个等长消息 m_0 和 m_1 给挑战者。挑战者 C 随机选择 $b\in\{0,1\}$，利用访问策略 Λ、关键字 w 对 m_b 进行加密处理。挑战者 C 随机选择 $\gamma\in Z_p$，计算密文

$$x=g_1^{\frac{1}{H_1(m_b)+\gamma}}=g^{\frac{d_1}{H_1(m_b)+\gamma}}$$
$$y=e(g_1,g_1)^{\frac{1}{H_1(m_b)+\gamma}}=e(g,g)^{\frac{d_1^2}{H_1(m_b)+\gamma}}$$
$$C_0=(x\|m_b)Te(g^s,g^{\alpha'})$$
$$C_1=g^s$$

对于 $\xi\in[1,l]$，挑战者 C 随机选择 $\tilde{r}_\xi\in Z_p$，隐式设置 $r_\xi=-\tilde{r}_\xi+so_\xi$，计算索引 $I_1=g^{es}$，$I_{\xi,1}=g^{dr_\xi}=(g^{\tilde{r}_\xi}g^{so_\xi})^d$。

挑战者 C 随机选择 $y_1,y_2,\cdots,y_n\in Z_p$，其中 $y_1=0$。挑战者 C 隐式设置

$$\vec{\boldsymbol{v}}=(s,sa+y_2,\cdots,sa^{n-1}+y_n)^{\mathrm{T}}$$
$$\lambda_\xi=M_\xi\vec{\boldsymbol{v}}=M_{\xi,1}s+M_{\xi,2}(sa+y_2)+\cdots+$$
$$M_{\xi,n}(sa^{n-1}+y_n)=\sum_{j=1}^{n}M_{\xi,j}(sa^{j-1}+y_j)$$

计算

$$I_{\xi,2}=g^{a\lambda_\xi}H(v_\xi)^{\frac{-r_\xi}{H_1(w)}}=$$
$$g^{a\sum_{j=1}^{n}M_{\xi,j}y_j}g^{\sum_{j=1}^{n}M_{\xi,j}sa^j}H(v_\xi)^{\frac{\tilde{r}_\xi}{H(w)}}(g^{-so_\xi})^{\frac{\tilde{v}_\xi}{H_1(w)}}\cdot$$
$$\prod_{i=1}^{I_\xi}\prod_{j=1}^{n}\left(g^{-\frac{sa^jM_{i,j}o_\xi}{o_i}}\right)^{\frac{1}{H_1(w)}},(I_\xi=\{i\,|\,\{\pi_i=v_\xi\}_{\xi=1}^{l})$$

当 $i=\xi$ 时，

$$\prod_{i=1}^{I_\xi}\prod_{j=1}^{n}\left(g^{-\frac{sa^jM_{i,j}o_\xi}{o_i}}\right)^{\frac{1}{H_1(w)}}=\prod_{j=1}^{n}(g^{-sa^iM_{\xi,j}})^{\frac{1}{H_1(w)}}。$$

所以，

$$I_{\xi,2}=g^{a\lambda_\xi}H(v_\xi)^{\frac{-r_\xi}{H_1(w)}}=$$
$$g^{a\sum_{j=1}^{n}M_{\xi,j}y_j}g^{\sum_{j=1}^{n}M_{\xi,j}sa^j}(H(v_\xi)^{\tilde{r}_\xi}(g^{-so_\xi})^{\tilde{v}_\xi}\cdot$$
$$\prod_{i=1,i\neq\xi}^{I_\xi}\prod_{j=1}^{n}(g^{-\frac{sa^jM_{i,j}o_\xi}{o_i}})\prod_{j=1}^{n}(g^{-sa^iM_{\xi,j}}))^{\frac{1}{H_1(w)}}=$$
$$g^{a\sum_{j=1}^{n}M_{\xi,j}y_j}(H(v_\xi)^{\tilde{r}_\xi}(g^{-so_\xi})^{\tilde{v}_\xi}\cdot$$
$$\left(\prod_{i=1,i\neq\xi}^{I_\xi}\prod_{j=1}^{n}\left(g^{-\frac{sa^jM_{i,j}o_\xi}{o_i}}\right)\right)^{\frac{1}{H_1(w)}}$$
$$(I_\xi=\{i\,|\,\pi_i=v_\xi\}_{\xi=1}^{l})$$

挑战者对索引中的 $I_{\xi,1}$ 进行重加密，计算 $I_{\xi,1}*=I_{\xi,1}^{\frac{1}{z}}=g^{\frac{dr_\xi}{z}}(\xi\in[1,l])$。

最后，挑战者 C 将密文 C_0,C_1,y 和索引 $I_1,\{I_{\xi,1}*,I_{\xi,2}\}_{\xi=1}^{l}$ 发送给攻击者 A。

查询阶段 2。与查询阶段 1 相同。

猜测阶段。攻击者 A 提交猜测值 b' 给挑战者 C。

如果 $b'=b$，挑战者 C 输出 0，表明 $T=e(g,g)^{sa^{q+1}}$；否则，C 输出 1，表明 $T=R$。

当 $T=e(g,g)^{sa^{q+1}}$ 时，攻击者 A 获得有效密文，则攻击者 A 的优势为 Adv_A。所以，

$$\Pr[b'=b\,|\,T=e(g,g)^{sa^{q+1}}]=\frac{1}{2}+\mathrm{Adv}_A$$
$$\Pr[C(\vec{\boldsymbol{p}},T=e(g,g)^{sa^{q+1}})=0]=\frac{1}{2}+\mathrm{Adv}_A$$

如果 $T=R$，攻击者 A 不知道关于 m_0 或 m_1 的任何信息。所以，

$$\Pr[b'=b\,|\,T=R]=\frac{1}{2},\Pr[C(\vec{\boldsymbol{p}},T=R)=0]=\frac{1}{2}。$$

从而，

$$\Pr[C(\vec{\boldsymbol{p}},T=e(g,g)^{sa^{q+1}})=0]-\Pr[C(\vec{\boldsymbol{p}},T=R)=0]=\mathrm{Adv}_A$$

所以，如果存在一个攻击者 A 可以在多项式时间打破本文方案的安全游戏，那么存在一个挑战者

C 能够打破 q-parallel BDHE 假设。证毕。

4.2 抗离线字典猜测攻击安全性证明

本节将分别从关键字索引和搜索令牌进行抗离线字典猜测攻击安全性分析。其中，属性 $\{c_i, v_i\}$ 和关键字 w' 均是离线字典中的元素，云服务器看作攻击者。

定理 2 如果 CDH 假设成立，则本文方案可以抵抗针对关键字索引的离线字典猜测攻击。

证明 加密阶段，攻击者收到索引 $\text{Index} = (I_1 = g^{es}, \{I_{i,1} = g^{dr_i}, I_{i,2} = g^{a\lambda_i} H(\pi_i)^{\frac{-r_i}{H_1(w)}}\}_{i=1}^{l})$。在 CDH 假设下，攻击者已知 g, g^d, g^a 计算 g^{ad} 是困难的。所以，攻击者无法通过判别式 $e(I_1, g^{ad}) = \prod_{i=1}^{l}(e(I_{i,2}, g^d) e(I_{i,1}, H(v_i)^{\frac{1}{H_1(w')}}))^{\omega_i}$ 是否成立猜测出索引中包含的属性值和关键字。证毕。

定理 3 本文方案可以抵抗针对搜索令牌的离线字典攻击。

证明 在搜索阶段，攻击者收到搜索令牌 $\text{Td} = \left(\text{Td}_1 = g^{kdt\theta}, \text{Td}_2 = g^{\frac{akdt\theta}{e}}, \left\{ \text{Td}_i = H(v_i)^{\frac{ktz\theta}{H_1(w')}} \right\}_{i=1}^{h} \right)$。由于 Td_i 中所包含的 z 元素是系统主密钥之一，攻击者无法获取。所以，攻击者无法通过判别式 $e(g^d, \text{Td}_i) = e\left(\text{Td}_1, H(v_i)^{\frac{1}{H(w')}} \right)$ 是否成立猜测出搜索令牌中的属性值和关键字。证毕。

5 性能分析

为了评估本文方案的性能，分别从功能、计算开销和存储开销 3 个方面与均支持策略隐藏的 Zhang 等[19]、Wang 等[17]和 Miao 等[18]的方案进行了对比。

5.1 功能比较

本节从功能方面进行对比，如表 2 所示。其中，F_1、F_2、F_3、F_4、F_5、F_6 分别代表关键字搜索、大属性域、访问结构、解密验证、抗离线字典猜测攻击、恒定解密开销。

表 2 功能比较

方案	F1	F2	F3	F4	F5	F6
Zhang 等方案	×	√	LSSS	√	√	×
Wang 等方案	√	×	与门	×	×	×
Miao 等方案	√	×	与门	×	×	×
本文方案	√	√	LSSS	√	√	√

从表 2 中可以看出，Wang 等[17]和 Miao 等[18]的方案支持关键字搜索，但这 2 种方案并不具备其他功能而且采用“与门”结构。本文方案与 Zhang 等[19]方案都采用大属性域和 LSSS 结构，支持解密验证和抗离线字典猜测攻击，但是 Zhang 等[19]和 Wang 等[17]方案的解密开销与用户属性数量相关，而 Miao 等[18]方案的解密开销与数据所有者合作人数相关，相比之下本文方案支持恒定解密开销，更加适用于计算资源受限的用户设备。

5.2 计算开销

本节将分别从理论和实验方面分析本文方案的计算开销。本文方案中主要涉及一些计算操作，E_1, E_2, E_T 分别代表群 G_1, G_2, G_T 上的指数运算，P 代表双线性配对运算。此外，n_i 代表每个属性的候选值个数，n 代表属性个数，I 代表访问矩阵行数，d 代表数据所有的合作人数。此外，经过 1 000 次测试取平均值，在单位计算开销上 $E_T < P < E_1 < E_2$，具体单位时间开销如表 3 所示。

表 3 单位时间开销

类型	时间/ms
E_1	11.03
E_2	11.23
E_T	0.83
P	6.49

1) 表 4 分别统计了本文方案与其他方案在系统初始化、加密、重加密、密钥生成、搜索令牌生成、关键字搜索和解密 7 个阶段的计算开销。在系统初始化阶段，Wang 等[17]和 Miao 等[18]方案的计算开销是与系统属性个数 n 和每个属性候选值个数 n_i 相关的，而本文方案与 Zhang 等[19]方案均是基于大属性域的。所以，此阶段的计算开销是恒定的常数级，远小于前 2 种方案。此外，本文方案在此阶段的计算开销也略小于 Zhang 等[19]方案。在密钥生成阶段，本文方案的计算开销相比其他对比方案至少降低了 nE_1。本文方案加密阶段包括数据加密和索引加密。在此阶段中，由于 Miao 等[18]方案的计算开销与数据所有者的合作人数 d 相关，与本文方案侧重点不同，所以不进行对比。与 Zhang 等[19]和 Wang 等[17]方案相比，本文方案的加密开销比 Wang 等[17]方案小 IE_1 左右，比 Zhang 等[19]方案小 $2IE_1$ 左右。在搜索令牌生成阶段，本文方案的计算开销同

表 4　　计算开销比较

方案	Setup	Encrypt	Recrypt	KeyGen	TokenGen	Search	Decrypt
Zhang 等方案	$5E_1+E_T$	$(5I+5)E_1+E_T$	—	$(2n+4)E_1$	—	—	$(2n+1)P+nE_T$
Wang 等方案	$\left(\sum_{i=1}^{n}n_i+2\right)E_1+E_2+2E_T$	$(4I+1)E_1+2E_T$	—	$(2n)E_2+E_T$	$(2n+1)E_2$	$(2n+1)P+E_T$	$(2n+1)P+E_T$
Miao 等方案	$\left(\sum_{i=1}^{n}n_i+1\right)E_1+E_T$	$(2d+2I+2)E_1+3E_T$	—	$(2n+d+4)E_1+E_T$	$(2n+1)E_1$	$(2n+1)P+E_T$	$3P+dE_1+dE_T$
本文方案	$3E_1+E_T$	$(3I+3)E_1+E_T+P$	IE_1	$(n+4)E_1$	$(n+2)E_1$	$(2n+1)P+nE_T$	$4P+E_1+E_T$

样比其他方案小 nE_1 左右。由于本文方案采用 LSSS 结构，在重构秘密值的过程中涉及指数运算，所以本文方案在搜索阶段的计算开销比其他 2 种方案高 $(n-1)E_T$。不过，从表 3 中可以看出，E_T 的计算开销较小。所以，在搜索阶段所增加的开销是可以接受的。最后，Zhang 等[19]和 Wang 等[17]方案在解密阶段计算开销与用户属性个数 n 相关，Miao 等[18]方案与数据所有者合作人数 d 相关，而本文方案的开销是恒定的常数级，具有明显的优势。

2) 为了更直观地了解本文方案的性能，本文在 2.5 GHz 主频、Intel(R) i5-7300HQ 处理器、16 GB 内存、Win10 操作系统的笔记本上基于 Java 配对加密（JPBC, Java pairing-based cryptography）库，采用 A 型奇异曲线 $y^2=x^3+x$ 进行仿真实验。在实验中，属性个数 n 和矩阵行数 I 取值范围为[0,50]，属性候选值个数 $n_i=5$。由于 Miao 等[18]方案在加密、密钥生成和解密阶段均涉及数据所有者合作人数 d，该变量在其他方案中并不包含，在实验中不便对比。所以，这 3 个阶段的实验仅与其他 2 种方案进行对比。为了便于更精确地分析各种方案的开销，表 5 给出了在 $n=50$、$I=50$、$d=1$ 时本文方案和对比方案的各阶段和整体时间开销。

系统初始化时间开销如图 2 所示，在系统初始化阶段，Miao 等[18]和 Wang 等[17]方案的时间开销与属性个数呈线性增长趋势，而本文方案和 Zhang 等[19]方案的时间开销一直处于常数状态，并且开销是非常低的。结合表 5 可以看出，当 $n=50$ 时，Miao 等[18]和 Wang 等[17]方案的时间开销分别为 2 776.9 ms 和 2 885 ms，而本文方案和 Zhang 等[19]方案的时间开销分别为 35 ms 和 57.6 ms。所以，本文方案在系统初始化阶段具有明显的优势。

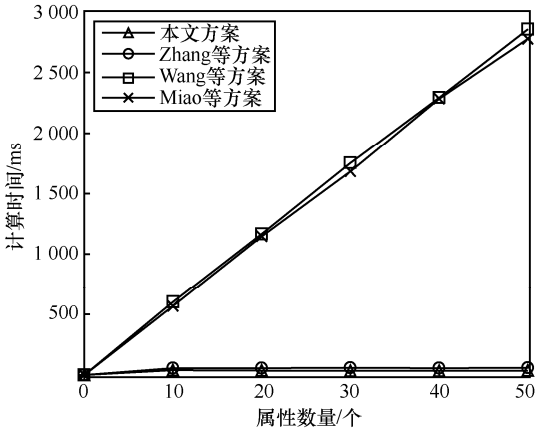


图 2　系统初始化时间开销

加密时间开销如图 3 所示，在加密阶段，本文方案与 Zhang 等[19]和 Wang 等[17]方案的时间开销都随着共享矩阵行数的增加而增加。相比之下，本文方案的增长速度是最低的，明显低于其他方案。结合表 5 可以看出，当 $I=50$ 时，本文方案时间开销为 1 738.4 ms，

表 5　　各阶段和整体时间开销

方案	Setup/ms	Encrypt/ms	ReEncrypt/ms	KeyGen/ms	TokenGen/ms	Search/ms	Decrypt/ms	整体开销/ms
Zhang 等方案	57.6	2 905.2	—	1 173.9	—	—	694	—
Wang 等方案	2 885	2 264.9	—	1 118.8	1 133.6	669.5	652	8 676.7
Miao 等方案	2 776.9	1 186.4	—	1 186.8	1 151.2	652.4	31.6	7 001.5
本文方案	35	1 738.4	560.1	618.5	589.3	700.8	38.1	4 280.2

而 Zhang 等[19]和 Wang 等[17]方案的时间开销分别为 2 905.2 ms 和 2 264.9 ms，明显高于本文方案。因此，本文方案在加密时间开销上是优于这 2 种方案的。

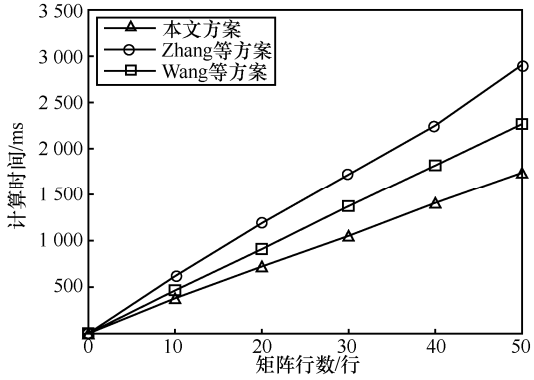


图 3 加密时间开销

密钥生成时间开销如图 4 所示，在密钥生成阶段，Zhang 等[19]和 Wang 等[17]方案的时间开销相近，本文方案相对较低。结合表 5 可以看出，当 $n=50$ 时，Zhang 等[19]和 Wang 等[17]方案的时间开销分别是 1 173.9 ms 和 1 118.8 ms，本文方案的时间开销为 618.5 ms，相比之下，本文方案的时间开销几乎是对比方案的 $\frac{1}{2}$。

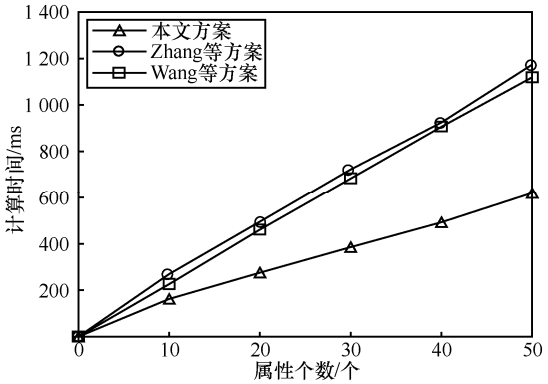


图 4 密钥生成时间开销

搜索令牌生成时间开销如图 5 所示，在搜索令牌生成阶段，Wang 等[17]与 Miao 等[18]方案的时间开销相近，本文方案相对较低。结合表 5 可以看出，当 $n=50$ 时，本文方案的时间开销为 589.3 ms，其他方案的时间开销分别是 1 133.6 ms 和 1 151.2 ms，几乎是本文方案的 2 倍。

搜索时间开销如图 6 所示，在搜索阶段，本文方案的时间开销与 Wang 等[17]和 Miao 等[18]方案相近。结合表 5 可以看出，当 $n=50$ 时，Wang 等[17]和 Miao 等[18]方案的时间开销分别是 669.5 ms 和 652.4 ms，本文方案是 700.8 ms，仅相差不到 50 ms。

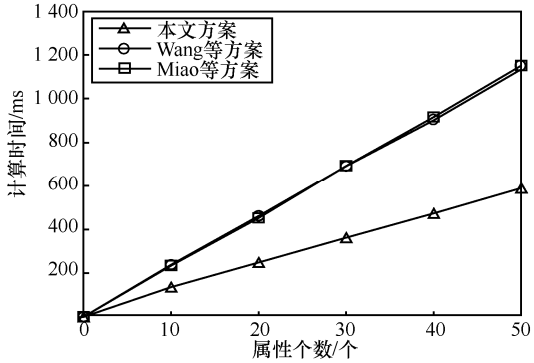


图 5 搜索令牌生成时间开销

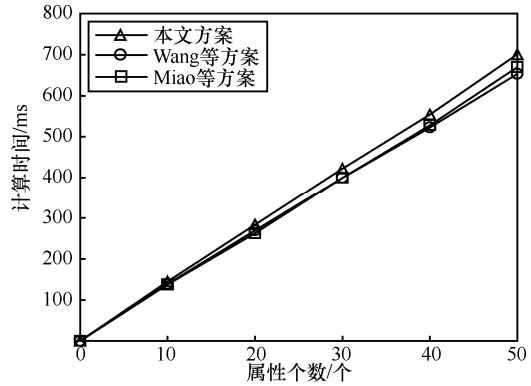


图 6 搜索时间开销

解密时间开销如图 7 所示，在解密阶段，Zhang 等[19]和 Wang 等[17]方案的时间开销随着属性个数增加明显增加。结合表 5 可知，当 $n=50$ 时，Zhang 等[19]方案的时间开销是 694 ms，Wang 等[17]方案的时间开销是 652 ms，而本文方案是 38.1 ms 且恒定不变，相比之下本文方案在此阶段的计算上更加高效。

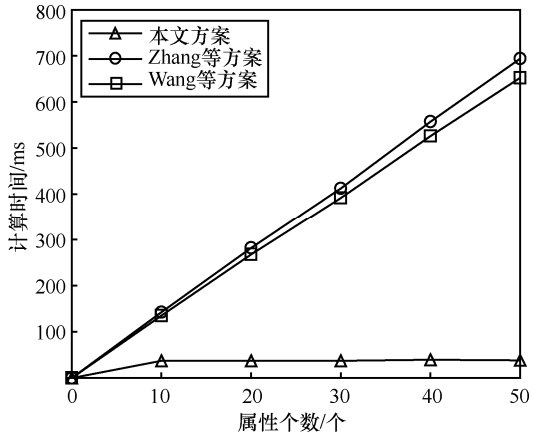


图 7 解密时间开销

最后，从表 5 中可以看出，当 $I=n=50$ 时，本文方案的整体时间开销是 4 280.2 ms，而 Wang 等[17]方案是 8 676.7 ms，是本文方案的 2 倍以上。此外，

即使当$d=1$时，Miao 等[18]方案开销的总和也达到了 7 001.5 ms，明显高于本文方案，并且随着数据所有者合作人数d的增加，Miao 等[18]方案部分阶段和整体的开销还会增加。所以，本文方案在整体开销上是占优的。

综上所述，本文方案仅在搜索阶段的开销略高于其他对比方案，但在其他阶段具有一定的优势，整体开销也低于其他方案。特别是在初始化阶段和解密阶段，本文方案的开销都是恒定的常数级，极大地减轻了授权中心和用户的计算负担，适用于轻量级的医疗用户设备。

5.3　存储开销

本节将分析方案在存储开销上的表现。其中$|G_1|,|G_2|,|G_T|,|Z_p|$分别表示群G_1,G_2,G_T,Z_p上元素的大小，并且在 A 型奇异曲线$y^2=x^3+x$下，$|G_1|=|G_2|=|G_T|$。表 6 分别给出了本文方案与 Zhang 等[19]、Wang 等[17]和 Miao 等[18]方案在系统公钥、系统主密钥、私钥、搜索令牌、密文与索引上的存储开销。

从表 6 中可以看出，本文方案的系统公钥和系统主密钥存储开销是恒定的，分别只需要$5|G_1|+|G_T|$和$6|Z_p|$的存储开销，与 Zhang 等[19]的方案相近，明显小于 Wang 等[17]和 Miao 等[18]的方案。在私钥存储上，本文方案只需$(n+4)|G_1|+|Z_p|$的存储开销，而其他方案均大于$2n|G_1|$，几乎是本文方案存储开销的 2 倍。在搜索令牌存储上，因为，$|G_1|=|G_2|$，所以 Wang 等[17]和 Miao 等[18]的方案存储开销相等，均是$(2n+1)|G_1|+|Z_p|$，而本文方案的存储开销为$(n+2)|G_1|$，比这 2 种方案减少了一半左右。最后，在密文与索引存储方面，本文方案的存储开销为$(2I+2)|G_1|+2|G_T|$，而其他方案存储开销均大于本文方案，特别是 Wang 等[17]的方案几乎是本文方案的 2 倍。

6　结束语

本文提出了一种轻量级可搜索医疗数据共享方案，在保障用户隐私和数据安全的同时提高了效率。在功能上，本文方案支持关键字搜索、部分策略隐藏和数据验证，便于用户在加密数据中检索目标文件，同时保障了数据所有者的隐私和数据安全。在性能上，本文方案采用大属性域、LSSS 结构和 Intel SGX 技术，提高了系统的可扩展性和访问结构的灵活性，加强了系统的安全性。此外，本文方案还实现了常数级的解密计算，适用于轻量级的医疗设备。最后，本文证明了方案具备选择明文攻击不可区分和抗离线字典猜测攻击安全性。在未来的工作中，期望进一步提高方案的搜索效率。

参考文献：

[1] WANG H J, NING J T, HUANG X Y, et al. Secure fine-grained encrypted keyword search for E-healthcare cloud[J]. IEEE Transactions on Dependable and Secure Computing, 2021, 18(3): 1307-1319.

[2] WANG H J, DONG X L, CAO Z F. Multi-value-independent ciphertext-policy attribute based encryption with fast keyword search[J]. IEEE Transactions on Services Computing, 2020, 13(6): 1142-1151.

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

[4] KAMARA S, PAPAMANTHOU C, ROEDER T. Dynamic searchable symmetric encryption[C]//Proceedings of 2012 ACM Conference on Computer and Communications Security. New York: ACM Press, 2012: 965-976.

[5] 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.

[6] JIANG P, MU Y, GUO F C, et al. Secure-channel free keyword search with authorization in manager-centric databases[J]. Computers & Security, 2017, 69: 50-64.

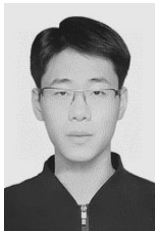
表 6　　**存储开销**

方案	系统公钥	系统主密钥	私钥	搜索令牌	密文与索引
Zhang 等方案	$7\|G_1\|+\|G_T\|$	$5\|Z_p\|+\|G_1\|$	$(2n+4)\|G_1\|$	—	$(3I+4)\|G_1\|+\|G_T\|$
Wang 等方案	$\left(\sum_{i=1}^{n}n_i+4\right)\|G_1\|+3\|G_2\|+2\|G_T\|+\|Z_p\|$	$\left(\sum_{i=1}^{n}n_i+3\right)\|Z_p\|$	$2n\|G_2\|+\|Z_p\|$	$(2n+1)\|G_2\|+\|Z_p\|$	$(4I+1)\|G_1\|+\|G_T\|$
Miao 等方案	$\left(\sum_{i=1}^{n}n_i+2\right)\|G_1\|+\|G_T\|$	$\left(\sum_{i=1}^{n}n_i+2\right)\|Z_p\|$	$(2n+d+3)\|G_1\|+\|G_T\|+(d+2)\|Z_p\|$	$(2n+1)\|G_1\|+\|Z_p\|$	$(2I+d+2)\|G_1\|+2\|G_T\|$
本文方案	$5\|G_1\|+\|G_T\|$	$6\|Z_p\|$	$(n+4)\|G_1\|+\|Z_p\|$	$(n+2)\|G_1\|$	$(2I+2)\|G_1\|+2\|G_T\|$

[7] WANG H J, DONG X L, CAO Z F, et al. Secure key-aggregation authorized searchable encryption[J]. Science China Information Sciences, 2019, 62(3): 1-3.
[8] SAHAI A, WATERS B. Fuzzy identity-based encryption[C]//Annual International Conference on the Theory and Applications of Cryptographic Techniques. Berlin: Springer, 2005: 457-473.
[9] GOYAL V, PANDEY O, SAHAI A, et al. Attribute-based encryption for fine-grained access control of encrypted data[C]//Proceedings of the 13th ACM Conference on Computer and Communications Security. New York: ACM Press, 2006: 89-98.
[10] BETHENCOURT J, SAHAI A, WATERS B. Ciphertext-policy attribute-based encryption[C]//Proceedings of 2007 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2007: 321-334.
[11] ZHENG Q J, XU S H, ATENIESE G. VABKS: verifiable attribute-based keyword search over outsourced encrypted data[C]//Proceedings of IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2014: 522-530.
[12] SUN W H, YU S C, LOU W J, et al. Protecting your right: verifiable attribute-based keyword search with fine-grained owner-enforced search authorization in the cloud[J]. IEEE Transactions on Parallel and Distributed Systems, 2016, 27(4): 1187-1198.
[13] NISHIDE T, YONEYAMA K, OHTA K. Attribute-based encryption with partially hidden encryptor-specified access structures[C]//International Conference on Applied Cryptography and Network Security. Berlin: Springer, 2008: 111-129.
[14] LAI J Z, DENG R H, LI Y J. Expressive CP-ABE with partially hidden access structures[C]//Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security. New York: ACM Press, 2012: 18-19.
[15] CUI H, DENG R H, WU G W, et al. An efficient and expressive ciphertext-policy attribute-based encryption scheme with partially hidden access structures[C]//International Conference on Provable Security. Berlin: Springer, 2016: 19-38.
[16] QIU S, LIU J Q, SHI Y F, et al. Hidden policy ciphertext-policy attribute-based encryption with keyword search against keyword guessing attack[J]. Science China Information Sciences, 2016, 60(5): 1-12.
[17] WANG S P, GAO T T, ZHANG Y L. Searchable and revocable multi-data owner attribute-based encryption scheme with hidden policy in cloud storage[J]. PLoS One, 2018, 13(11): 1.
[18] MIAO Y, LIU X, CHOO K K R, et al. Privacy-preserving attribute-based keyword search in shared multi-owner setting[J]. IEEE Transactions on Dependable and Secure Computing, 2019, 18(3): 1080-1094.
[19] ZHANG Z S, ZHANG W, QIN Z G. A partially hidden policy CP-ABE scheme against attribute values guessing attacks with online privacy-protective decryption testing in IoT assisted cloud computing[J]. Future Generation Computer Systems, 2021, 123: 181-195.
[20] MIAO Y B, MA J F, LIU X M, et al. VMKDO: verifiable multi-keyword search over encrypted cloud data for dynamic data-owner[J]. Peer-to-Peer Networking and Applications, 2018, 11(2): 287-297.
[21] SHINDE S, CHUA Z L, NARAYANAN V, et al. Preventing page faults from telling your secrets[C]//Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security. New York: ACM Press, 2016: 317-328.
[22] NING J T, HUANG X Y, SUSILO W, et al. Dual access control for cloud-based data storage and sharing[J]. IEEE Transactions on Dependable and Secure Computing, 2022, 19(2): 1036-1048.
[23] FISCH B, VINAYAGAMURTHY D, BONEH D, et al. IRON: functional encryption using intel SGX[C]//Proceedings of 2017 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM Press, 2017: 765-782.

[作者简介]

殷新春（1962− ），男，江苏泰州人，博士，扬州大学教授、博士生导师，主要研究方向为密码学、软件质量保障、高性能计算等。



王梦宇（1997− ），男，江苏邳州人，扬州大学硕士生，主要研究方向为属性基加密、信息安全等。

宁建廷（1988− ），男，浙江龙游人，博士，福建师范大学教授、博士生导师，主要研究方向为应用密码学与数据安全、区块链与机器学习安全、隐私保护技术等。