《工程科学学报》编辑部

可搜索加密及其驱动的 SQL 隐私数据库设计

陆海 [1)]，薛显斌 [2)✉]，朱岩 [1)✉]，陈娥 [1)]，韩皓庭 [1)]，孟疏桐 [1)]，林鸿杰 [1)]

1) 北京科技大学，计算机与通信工程学院，北京 100083

2) 北方导航控制技术股份有限公司，北京，102600

✉ 通信作者，E-mail: 3032028407@qq.com (薛显斌); zhuyan@ustb.edu.cn (朱岩)

摘　要　隐私数据库是促进国家大数据战略与数据要素市场发展中构建数据开放、共享及治理体系的重要手段，而可搜索加密作为实现隐私数据库的重要密码技术，仍存在缺乏灵活检索机制及抗量子安全等问题，也难以适配关系数据库中的 SQL 查询机制。在对可搜索加密技术现状分析基础上，本文设计了可适配关系数据库 SQL 查询语言的隐私数据库架构，客户端引入隐私 SQL 引擎将索引和数据字段转变为密文状态；用户发起查询请求时，该引擎可依据查询策略生成查询凭证，隐私数据库进而依据凭证对密态索引进行密码化检索，匹配成功的密态数据字段可由用户私钥进行解密。进一步，本文在格密码体制下利用理想格上短整数解（R-SIS）和带误差学习（R-LWE）困难问题，设计了检索策略的属性基可搜索加密（RP-ABSE）方案用以支持上述隐私数据库密码系统的构建。该方案将查询策略与查询凭证相绑定，确保密文数据的索引可依据查询策略进行细粒度密码化检索；同时，引入小策略矩阵（SPM）来优化安全查询策略生成，降低索引匹配过程中累积误差。由安全性证明可知，查询凭证满足在选择策略攻击下的不可伪造性（EU-CPA），所提系统满足在带有策略和标识查询的选择明文攻击下的语义安全性（IND-PIQ-CPA）。

关键词　隐私数据库；安全查询策略；属性基可搜索加密；理想格；小策略矩阵

分类号　TP39

Privacy-Preserving SQL Database Driven by Searchable Encryption

Hai Lu [1)], Xianbin Xue[2)✉], *Yan Zhu*[1)✉], *Cecilia E Chen*[1)], *Haoting Han*[1)], *Shutong Meng*[1)], *Hongjie Lin*[1)]

1) School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China

2) North Navigation Control Technology Co. Ltd., Beijing, 102600

✉ Corresponding author, E-mail: 3032028407@qq.com (Xianbin Xue); zhuyan@ustb.edu.cn (Yan Zhu)

ABSTRACT　Privacy-preserving database plays a crucial role in establishing an open, shared, and governed system amidst the evolution of national big data strategies and data market. Searchable encryption (SE), as a fundamental cryptographic technology for building privacy-preserving database, facilitates efficient searching of encrypted data without the need for decryption. Among various SE schemes, Attribute-Based Searchable Encryption (ABSE) provides advantages in access control, data authenticity, and retrieval efficiency. However, most existing ABSEs could not support the flexible query methods of SQL in relational databases, as well as fine-grained query policies. Moreover, most existing ABSEs are constructed under traditional algebraic structures, such as Bilinear Pairing, which are susceptible to quantum computing attacks. To address these challenges, this paper presents a novel privacy-preserving database architecture that can be adaptable to the SQL query language of relational databases. The architecture is divided into two parts: clients and cloud outsourcing services. All data are in a ciphertext form outside the clients, and data table in cloud privacy-preserving database consists of four kinds of fileds: public fields, encrypted index fields, encrypted data fields, and confidential fields. When a user initiates an SQL query, the Privacy-Preserving SQL Engine (PP-SQL Engine) can transform it into a private SQL

language to perform the cryptographic retrieval on the encrypted index fields. More specifically, the query policy in SQL is transformed into several query credentials that are bound with the policy. Theses credentials are further used to retrieve the encrypted data fields in the database whose encrypted index fields match with the policy. The retrieved encrypted data fields can be decrypted at the client through the user's private key for verifying the user's identity. To provide the cryptographic support for the privacy-preserving database architecture, the Retrieval-Policy Attribute-Based Searchable Encryption (RP-ABSE) scheme is designed on the Key-Policy Attribute-Based Encryption (KP-ABE) framework. The security of this scheme is based on the hard problem over ideal lattice, i.e., the Short Integer Solution (SIS) and the Learning With Error (LWE). Moreover, the secure query policy is bound to the query credentials rather than the encrypted index fileds. It ensures that encrypted data can be cryptographically retrieved by different query policies, eliminating the need for updating the encrypted data even when the query policy undergoes changes. Simultaneously, the Small Policy Matrix (SPM) is introduced to optimize the generation of secure query policies and reduce cumulative errors in the process of index matching. Finally, this paper proves that the query credential satisfies unforgeability under the chosen policy attack, and the RP-ABSE satisfies semantic security under the chosen plaintext attack with policy and identity queries. Therefore, the privacy-preserving database architecture could provide an important technique support for the establishment of data market mechanisms and data governance systems.

KEY WORDS Privacy-preserving Database; Secure Query Policy; Attribute-Based Searchable Encryption; Ideal Lattice; Small Policy Matrix

隐私数据库是一种能够存储敏感或隐私数据的数据库技术，可在不泄露数据内容的情况下支持对数据的查询、分析与利用。隐私数据库是伴随着云计算等第三方数据库外包服务[1]（Database Outsourcing Services, DOS）发展而来的，将数据库外包给云可以带来如降低成本、可扩展性和可用性（即点即用）等优势，但云服务提供商可能不完全可信，或者可能被恶意攻击者入侵等安全风险，也给确保数据的机密性和完整性带来了挑战。近年来，随着国家大数据战略的实施以及数据泄露、数据滥用、数据歧视等安全事件频发，隐私数据库技术由于能够通过技术手段保护数据的隐私性、完整性和可用性，防止数据的恶意攻击和非法利用而日益受到学术界和产业界的重视。特别是隐私数据库技术在保护数据要素开放、共享和交易等方面，由于数据库仍然是目前大规模数据存储的首选方式，因此它对于建立数据市场化机制和数据的治理体系都具有重要技术支撑作用。

目前已经存在多种隐私数据库实现技术，包括：差分隐私（Differential Privacy, DP）[2]、安全多方计算（Secure Multi-Party Computation, SMPC）[3]、可搜索加密（Searchable Encryption, SE）[4]、属性基加密（Attribute-Based Encryption, ABE）[5]、同态加密（Homomorphic Encryption, HE）[6]等。这些技术各具特色，可根据不同应用需要加以选择。例如，差分隐私是一种通过添加噪声但保留数据整体分布特征的数据隐私保护技术，但因为噪声的存在破坏了数据完整性和真实性，因此限制了它在一些数据质量管控严格领域中被应用。又如安全多方计算是一种在分布式环境下通过多个参与方之间共享秘密的方式，保障隐私数据的存储与复杂计算的实现，但该技术只适合一定参与方数目下的隐私数据保护，且通信开销巨大，因此并不适合通常单节点的隐私数据库构建。

相比较而言，采用加密技术（如 SE、ABE、HE）构建的隐私数据库及查询协议则可实现数据库中隐私数据字段的加密或封装等密码处理，并通过密码学方法对密态数据字段进行检索和分析或解密。其优势在于在“字段”级保留了数据真实性和完整性，避免了数据库文件整体加密所带来的粒度过大而无法做到精细化检索和分析，并支持现有数据库范式（用于设计关系型数据库的规范化过程中的一组规则），是对数据保护由粗粒度向细粒度的过度，且可保证数据本身和查询过程的隐私。从密码学发展来说，隐私数据库及查询是将数据加密、数据认证、密态计算等密码技术综合运用形成的一门新兴密码应用领域，对提高外包等开放空间下的敏感数据应用具有重要研究价值。

可搜索加密是隐私数据库检索中的重要技术分支之一，它允许在不解密密文的情况下高效地搜索一组加密数据，可用于将文件或数据外包给不受信任的云存储服务器而不会以明文形式泄露文

件，同时保留服务器对文件的搜索能力。类似于密码加密系统分为对称和非对称两大类，目前可搜索加密主要分为可搜索对称加密（Searchable Symmetric Encryption, SSE）[7]和可搜索非对称加密（Searchable Asymmetric Encryption, SAE）[4]两大类。其中，SAE 中研究最广泛的是基于 ABE 的方案构造，常被称为属性基可搜索加密（Attribute-Based Searchable Encryption, ABSE），它将密码化的数据字段或关键字作为安全索引，在实现数据库密文检索基础上对查询用户的身份进行认证并解密。

然而，就目前数据库系统而言，结构化查询语言（Structured Query Language, SQL）语言作为一种标准的关系数据库管理系统（RDMS）查询语言，可支持存储、检索、修改、删除等各种数据库事务处理和复杂查询的执行，在关系数据库与 NoSQL 数据库中都已得到了极其广泛的应用。但是，已有方案中支持关系数据表格和 SQL 语言进行隐私数据存储与检索仍然鲜有研究。此外，近年来随着量子计算（Quantum Computing）技术的迅猛发展已经对传统密码体制[8]安全产生巨大冲击，然而目前隐私数据库相关密码技术仍然主要依赖于传统密码系统（双线性群下密码构造等），导致缺乏抗量子攻击下的安全性。因此，研制关系数据表下支持 SQL 查询语言且抗量子的隐私数据库密码系统构造是亟待解决的问题。

有鉴于此，本文聚焦于数据库隐私保护的架构和可行密码学构造，对可搜索加密等安全技术的国内外研究现状进行了阐述与分析，并提出了属性基可搜索加密技术等方法在与隐私数据库结合方面存在的挑战性问题。针对所提问题，本文设计了可适配关系数据库 SQL 查询语言的隐私数据库架构，该架构划分为用户端和数据外包服务两部分，用户端之外数据均处于密文状态，SQL 语言可转化为隐私 SQL 语言对密文数据进行密码化检索处理，可在关系数据表中支持多样化隐私字段存储，最终检索到的密文通过用户端解密实现用户身份验证。

进一步，本文在格密码体制下以密钥策略 ABE（Key-Policy ABE, KP-ABE）框架为基础设计了检索策略的属性基可搜索加密（Retrieval-Policy Attribute-Based Searchable Encryption, RP-ABSE）用以支持上述隐私数据库密码系统的构建。在该方案中，安全查询策略与查询凭证相绑定而非索引，以确保密文数据的索引可依据安全策略进行灵活的密码化检索；同时，引入小策略矩阵（Small Policy Matrix, SPM）来优化安全查询策略生成以降低索引匹配过程中累积误差，降低密文存储开销和检索计算开销。通过安全性证明，查询凭证满足在选择策略攻击下的不可伪造性，所提系统满足在带有策略和标识查询的选择明文攻击下的语义安全性。

1 当前技术现状

在基于密码学的隐私数据处理中，可搜索加密作为一项支持密码化关键字检索的重要技术，可在确保数据隐私性的同时，允许使用者对密文形式下的数据进行检索，有助于减轻数据泄露的风险，并保持数据库系统的功能。相比于传统加密方案，SE 方案构造通常涉及以下三种算法：

（1）索引建立算法：用于将密文数据的若干关键字密码封装为安全索引；

（2）查询生成算法：可依据待查询索引生成查询凭证（也称查询陷门）用于匹配安全索引；

（3）密文检索算法：根据查询凭证搜索出与之匹配的安全索引，并检索出对应的密文数据。

SE 方案通过上述算法可保证数据关键字或字段的隐私（称为数据隐私）以及数据查询内容的隐私性（称为查询隐私）。下面将按照可搜索对称加密和可搜索非对称加密两类讨论 SE 技术现状。

1.1 对称可搜索加密

SSE 是一种采用对称密码构建的可搜索加密技术，其特征是系统各参与方使用的密钥是相同且共享的。自从文献[7]提出了第一个 SSE 方案以来，多种多样的 SSE 方案被提出用以满足不同的应用要求。例如，文献[7]利用伪随机函数来构造了基础方案，并依据不同需求扩展出具备可控检索、隐藏搜索等功能的方案，然而这些方案需要扫描整个文档，计算开销较大，且同时顺序扫描会向服务器泄露隐私信息。为避免对整个加密文档进行检索，文献[9]提出了基于索引检索的 SSE，并分析

安全索引可能遭受的攻击以及多用户共享同一索引的所带来的困难。之后较多方案[10],[11],[12]都采用了索引检索来设计方案。

已有一些 SSE 方案被应用于隐私数据库构建中，例如，文献[13]利用数据库关系表来实现索引查询和模糊范围查询；文献[14]提出了 CryptDB，该文献通过对数据进行嵌套加密来设计了洋葱模型以支持更多查询方式；文献[15]将高维特征向量作为检索依据而非关键字，并依据局部敏感散列来编码关联文件标识符和特征向量索引来实现动态 SSE；文献[16]利用非确定性加密和保序加密来构建安全索引，并提出了支持诸如等值查询和布尔查询等丰富 SQL 查询方式的 SSE。

由于对称加密体制的限制，SSE 方案中数据加密、搜索及解密均采用相同密钥不利于数据外包和大范围用户群下的数据共享场景，因而并不是目前隐私数据库研究的主体。

1.2 可搜索非对称加密

SAE 是一种采用非对称密码或公钥密码构建的可搜索加密技术，也称为可搜索公钥加密，其特征是系统中密钥包括公钥和私钥两种，通常公钥实现数据加密，而私钥实现数据检索或解密。现有 SAE 方案以基于 ABE 的 SAE 构造为主，其它方案则多种多样，统称为非 ABE 的 SAE 方案。

属性基加密[17],[18]也是具有密文检索特征的公钥密码技术之一，是一种依据数据及其使用者等实体的属性进行访问权限判决的密码技术。它的出现得益于诸如物联网、云计算、人工智能等新技术的快速发展，使信息共享和交互的效率获得极大提升，有助于解决数据安全所面临的严峻挑战[19]。该密码系统的显著特征是：

(1) 可支持大量用户共用同一密码系统，而非每名用户独享系统和参数；

(2) 每名用户拥有各自的标识及属性和唯一性的密钥；

(3) 可对指定用户群进行授权解密，而不仅仅是对某一用户。

这些特征使得属性基密码系统不再局限于保护数据隐私性，还具有认证数据使用者身份的功能。可根据不同数据使用者的属性或角色赋予用户不同的访问级别和权限，能实现更精确和个性化的数据保护，满足各种应用的多样化和动态的安全需求。

结合上述优势，越来越多的研究者将 ABE 与 SE 相结合，提出了基于属性的可搜索加密方案 ABSE，相比于传统 SAE，ABSE 在访问控制、数据真实性和检索效率等方面具有优势[20]。当前，较多 ABSE 方案的设计思路是：利用公共参数将关键字集合封装为安全索引，而利用属性或策略将访问数据加密为密文数据[21]；在检索时，可依据安全索引寻找到匹配的密文数据；若该用户属性集满足数据的访问策略，这些数据可进一步被用户属性密钥解密。例如，文献[22]设计了一个具有接收者匿名性并支持隐藏访问策略的 ABSE 方案，密文数据检索条件为至少存在一个索引属于授权用户的查询范围中；文献[4]进一步将该设计思路引入至区块链系统中；文献[23]将 0-1 编码理论引入安全索引构建和查询中，减少了数字类型的关键字数量，从而降低计算和存储开销。作为可搜索加密的重要分支，ABSE 在保持原有安全性基础上，支持对检索文件的细粒度访问控制，更适用于多用户操作的隐私数据库场景下。目前已有一些 ABSE 方案应用在数据库中用于数据共享。文献[24]以传统 ABE 算法为基础，设计了动态可搜索加密方案；文献[25]提出了一种可验证的属性可搜索加密方案，该方案可通过二维线性表进行属性识别，并结合关键词检索该属性权限下可以访问的数据集。不同于上述方案，文献[26]利用策略来表示查询范围，并将该策略通过 ABE 的思想（类似于密文策略 ABE，即 CP-ABE）转化为安全策略嵌入至安全索引构建中，使得密文数据可被细粒度地检索。

非 ABE 的 SAE 方案依据特殊场景下不同要求分为多种类别[27],[28]，例如，文献[29]提出一种具有模糊关键字搜索的公钥方案，该方案生成了精确关键字搜索陷门和模糊关键字搜索陷门，其中模糊关键字搜索陷门被提供给云服务器以检索匹配的文档，而用户可以通过在本地发出精确关键字搜索陷门来进一步过滤结果；文献[30]提出一种实现搜索包含所有给定关键字（合取关系）方案，文献[31]分析该方案容易受到离线关键字猜测攻击，并提出了对该攻击实现语义安全的改进方案。然而，该方案在生成陷门时需要索引中的关键字完整列表，导致信息泄露和查询隐私不足。文献[32]

提出了一种改善索引大小和陷门生成的方案，保证关键字操作在存储和性能上都更高效。文献[33]依据 LWE 困难假设提出了抵抗量子威胁的 SAE 方案。总之，目前非 ABE 的 SAE 方案研究多种多样并较适合特殊场景下的应用。

综上所述，当前 SAE 方案在与隐私数据库结合方面仍存在以下挑战性问题：

（1）SAE 方案（特别是 ABSE）大多是针对非结构化密文集合上搜索而设计的，难以满足结构化存储的关系数据库。一方面，较多的 ABSE 方案[4],[22]关键字检索机制未引入细粒度的查询机制（安全策略），难以支持 SQL 语言灵活的查询方式，例如，文献[22]仅支持多个索引的“或”逻辑；另一方面，另一类 ABSE[26]将查询作为“安全策略”嵌入至索引集合中，用以控制拥有特定属性集的用户进行检索，然而这种思路使得数据的访问范围固化，难以支撑授权用户对数据的灵活检索。例如，记录具有三个关键字：性别、机构和民族，并以此组建查询 $\Pi := (\text{"性别=男"} \wedge \text{"机构科大北"} \wedge \text{"民族=汉"})$，而用户利用凭证“男”和“汉”查询时是无法检索到该数据的，因为用户没有提供“机构”属性。因此，构建适合关系数据库的 SAE 架构是必须解决的问题。

（2）当前较多 ABSE 方案是构造在双线性群下，例如文献[24]和文献[25]。随着量子计算机的发展[34]，基于传统密码系统的 ABSE 方案难以抵抗量子威胁，而目前几类“后量子密码”中格密码是研究最为广泛和最具实用化的抗量子密码系统。因此有必要开展可抵抗量子威胁的密码算法来提升抗量子安全性，特别是格密码体制下面向隐私数据库的密码构造。

2 系统模型

在关系数据库中数据是存储在由行与列组成的数据表中，其中，行用于存储整条记录，每一列则存储记录的某种属性。作为数据库标准语言，SQL 语言被用于数据的存储、检索和管理。针对 SQL 语言的特点，本文设计了一种针对数据外包存储的隐私数据库模型，用以适配关系数据库中隐私数据的存储、查询等功能。如图 1 所示，本系统共包含三种实体：

（1）隐私 SQL 引擎：负责将用户发起的 SQL 操作（如 INSERT、SELECT 等）进行密码化处理以及用户密钥生成，其位于用户端且为第三方可信的。

（2）隐私数据库系统：在现有关系数据库基础上以密文形式存储敏感数据，并可对隐私 SQL 引擎发来的用户查询进行密码学检索处理。

（3）数据用户：分为提供者和使用者两类，仅需采用标准 SQL 语言即可对服务器端存储的隐私数据进行操作，而无需参与隐私数据的密码化管理操作，即客户的隐私处理“透明性”。

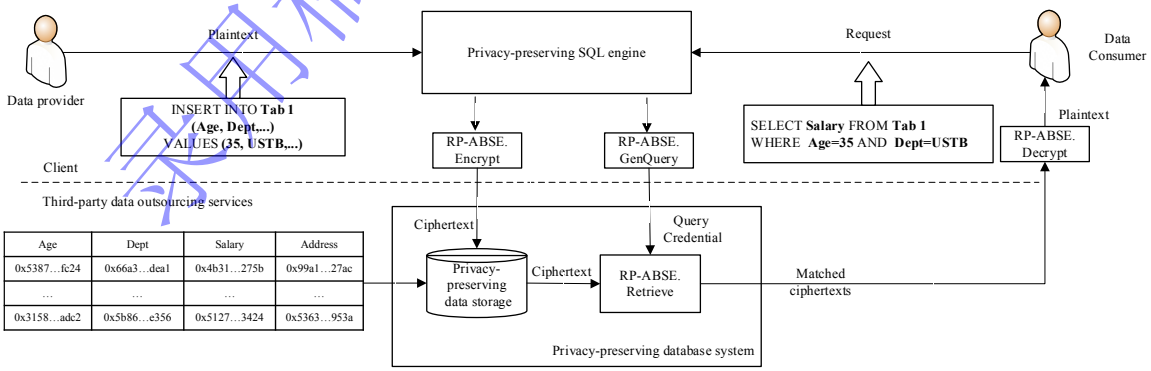

图 1 隐私数据库系统模型图

Fig.1 The architecture of the privacy-preserving database

图 1 包含数据记录的添加和查询流程两个主要流程，细节如下：

（1）数据添加流程：与通常数据添加处理一样，数据提供者调用 INSERT 语句向隐私数据库中添加数据记录，由隐私 SQL 引擎将该 INSERT 语句转化为对应的隐私 SQL 语句，包括对语句包

含的若干敏感数据字段加密为密文数据，之后将密文数据替换 SQL 语句中对应原数据，并把处理与好的 SQL 语句发送给服务器端隐私数据库系统，由其将密文数据存入隐私数据库中（细节见图 2）；

（2）数据查询流程：与通常数据查询一样，数据使用者调用 SELECT 语句从数据库中的密文数据进行查询，隐私 SQL 引擎将该 SELECT 语句转化为对应的隐私 SQL 语句，包括依据查询策略为用户生成查询凭证，之后隐私数据库系统依据凭证通过密码化检索处理从数据表中寻找符合要求的密文数据，客户端对密文数据进行脱密处理（细节见图 3）。

在隐私数据库设计方面，本文以关系数据库为基础构建隐私数据表。首先，对于明文形态的数据记录，将其分为公开字段、索引字段和数据字段，其中，公开字段是用于标识指定字段类别，索引字段是数据记录的检索关键词，而数据字段则是记录可查询内容。对于密文形态的数据记录，本文将其分为四类字段：公开字段、密态索引字段、密态数据字段、以及保留字段，这四种数据字段说明如下（见图 2 中示例）：

（1）**公开字段**：用于存储非敏感数据的字段，如图 2 中的“年龄”或“机构”；

（2）**密态索引字段**：存储索引字段的密文值，用于密码学检索过程中的索引匹配，例如“0x7a8d…cd50”是“性别”下索引值“男”的加密索引字段（注意，不同记录的相同索引的密文值是不同的）；

（3）**密态数据字段**：其不参与生成检索但可被解密获取原数据，通常由对称会话密钥所加密，如图 1 中“0x5382…9102”对应于公开字段“薪水”下的加密数据字段；

（4）**保留字段**：为数据记录所对应密码学参数等非公开信息的存储字段，其包括会话密钥的密文形式以及其他用于解密或认证的参数。

本架构包含以下特征：（1）系统模型划分为两部分，即用户端和数据外包服务，除用户端外数据均处于密文状态；（2）SQL 查询语言由隐私 SQL 引擎转化为隐私的 SQL 语言来执行对密文数据的操作；（3）数据库系统支持包含四种类型的隐私数据字段存储及密码化检索处理；（4）实现了用户端解密，从而验证了用户身份。

2.1 系统定义

本文构造了支持检索策略的属性基可搜索加密 RP-ABSE 方案可为所提隐私数据库系统提供密码学支撑，该方案包含以下 6 个密码算法：

（1）系统生成算法 $\mathrm{Setup}(n)\to(\mathrm{pk},\mathrm{msk})$：为系统建立算法，以安全参数 n 为输入，输出系统公钥 pk 和主私钥 msk；

（2）密钥生成算法 $\mathrm{GenKey}(\mathrm{msk},\mathrm{ID})\to \mathrm{sk}_{\mathrm{ID}}$：为密钥生成算法，以 msk 和用户标识 ID 为输入，输出该用户的私钥 $\mathrm{sk}_{\mathrm{ID}}$；

（3）加密算法 $\mathrm{Encrypt}\left(\mathrm{msk},\{\mathrm{att}_i\}_{i=1}^{N},\tau,\mathrm{ek}\right)\to\left((c_i)_{i=1}^{N},c'\right)$：为数据加密算法，以 msk、索引字段集合 $\{\mathrm{att}_i\}_{i=1}^{N}$、数据表标识 τ 和会话密钥 ek 为输入，输出加密索引字段集合 $(c_i)_{i=1}^{N}$ 和会话密钥密文 c'；

（4）查询生成算法 $\mathrm{GenQuery}(\mathrm{msk},\Pi,\mathrm{ID})\to\left(\{T_i\}_{\mathrm{att}_i\in\Pi},\omega_{\mathrm{ID}},\xi\right)$：为查询算法，以 msk、策略 Π 和用户标识 ID 为输入，输出索引查询凭证集合 $\{T_i\}_{\mathrm{att}_i\in\Pi}$、解密参数 ω_{ID} 和加密参数 ξ；

（5）检索算法 $\mathrm{Retrieve}\left(\Pi,\mathbf{M},\Sigma,\{c_i\}_{\mathrm{att}_i\in\Sigma},\{T_i\}_{\mathrm{att}_i\in\Pi},\xi,c'\right)\to(\varphi,\tilde{c})$：为查询策略匹配函数，以 Π 及对应的小策略矩阵 $\mathbf{M}$ 和索引查询凭证集合 $\{T_i\}_{\mathrm{att}_i\in\Pi}$ 与索引字段集合 Σ 及其对应的加密索引字段集合 $\{c_i\}_{\mathrm{att}_i\in\Sigma}$ 以及 ξ 和 c' 为输入，输出匹配归约值 φ 和更新后的会话密钥密文 $\tilde{c}$；

（6）解密算法 $\mathrm{Decrypt}(\mathrm{sk}_{\mathrm{ID}},\omega_{\mathrm{ID}},\varphi,\tilde{c})\to\mathrm{ek}$：为解密函数，以 $\mathrm{sk}_{\mathrm{ID}}$、$\omega_{\mathrm{ID}}$、$\varphi$ 和 $\tilde{c}$ 为输入，输出会话密钥 ek，该 ek 可用于实际密态数据字段解密。

正确性：满足上面密码系统的数据库查询，该用户以必然概率恢复出会话密钥 ek，即

$$\Pr\left[\mathrm{Decrypt}\left(\mathrm{sk}_{\mathrm{ID}},\omega_{\mathrm{ID}},\varphi,\tilde{c}\right)=\mathrm{ek}:\left(\varphi,\tilde{c}\right)\leftarrow\mathrm{Retrieve}\left(\Pi,\mathbf{M},\Sigma,\{c_i\}_{\mathrm{att}_i\in\Sigma},\{T_i\}_{\mathrm{att}_i\in\Pi},\xi,c'\right)\wedge\mathrm{Match}\left(\Pi,\Sigma\right)=\mathrm{True}\right]=1$$

其中， $\mathrm{Match}\left(\Pi,\Sigma\right)=\mathrm{True}$ 表示索引字段集合 Σ 满足查询策略 Π 。

2.2 系统流程

本节主要介绍系统中两个重要的隐私 SQL 语句的密码化实现，即 INSERT 和 SELECT 语句。在此之前，约定隐私 SQL 引擎已调用所提密码系统中 Setup 算法生成系统公私钥对 $(\mathrm{pk},\mathrm{msk})$，并且调用 GenKey 算法为所有用户生成其私钥 $\mathrm{sk}_{\mathrm{ID}}$ 。

当授权用户需向隐私数据库中添加记录时，数据记录的添加采用 SQL 语言中 INSERT 命令的形式，如图 2 所示，授权用户采用标准 INSERT 命令将数据记录以明文形式添加到 VALUES 语句后部（可添加 Token 类认证信息鉴别用户身份）。隐私 SQL 引擎将该命令转化为隐私 SQL 语句，并发送至隐私数据库中用于存储其中的数据记录，其具体执行步骤如下所示：

（1）选取会话密钥 ek 将数据记录中若干数据字段转化为密态数据字段；

（2）调用 Encrypt 算法将若干索引字段和前述会话密钥 ek 分别转化为对应的密态索引字段集合和会话密钥密文，必要时可将会话密钥密文保存至保密字段中。

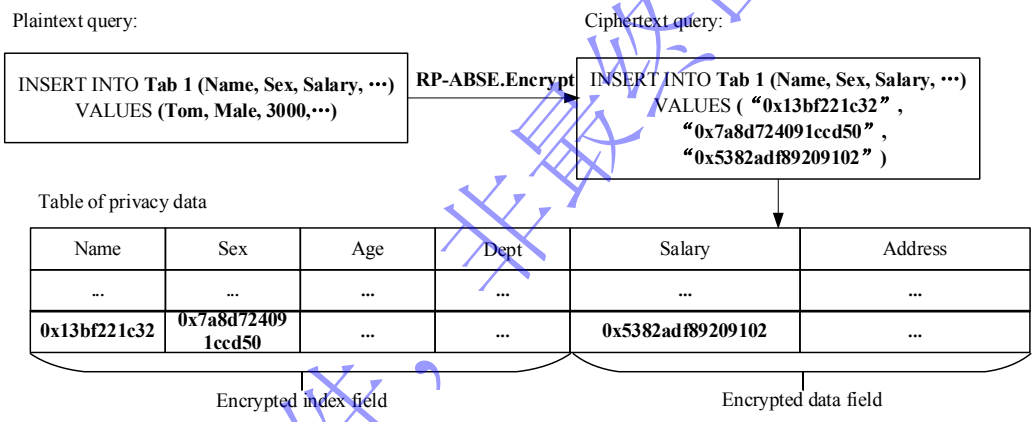

图 2 隐私数据库中 INSERT 语句的处理流程示意图

Fig.2 The workflow of INSERT statement in the privacy-preserving database

对于上述构建完成的隐私数据库，用户可通过标准的 SELECT 语句在隐私 SQL 引擎帮助下完成数据查询任务。图 3 给出了隐私数据库查询过程，为简化描述，引入符号“[]”表示给定字段的密码学封装，例如，“[北科大]”为索引字段“北科大”的密码学封装，索引查询凭证“机构=[北科大]”则为索引判定谓词“机构=北科大”的密码学封装（机构为字段名，属于可公开信息，“[北科大]”则对应为某个密码值，如“机构=0x3277…CBCD”）。隐私数据库查询过程的具体执行步骤如下所示：

（1）SQL 隐私预处理：隐私 SQL 引擎检测该用户的查询范围是否可被授权，若可被授权，将该 SQL 查询转化为 Privacy-SQL，其具体执行步骤为：将 SELECT 语句中 WHERE 所对应的查询范围作为查询策略，调用所提出密码系统中 GenQuery 算法生成“索引查询凭证”集合 $\{T_i\}_{\mathrm{att}_i\in\Pi}$ 与“解密参数” ω_{ID} 、加密参数 ξ 。以图 3 为例，查询策略"年龄=35 AND 机构=北科大"中的索引判定谓词由隐私 SQL 引擎转化为索引查询凭证，即分别转化为查询凭证“年龄=[35]”和“机构=[北科大]”，并将加密参数 ξ 作为查询凭证添加到 FROM 中，作为数据表的辅助参数。最后，将隐私 SQL 语句发送给服务器端，同时保留好“解密参数”用于后续处理；

（2）密码化检索处理：服务器端接收到密码化 SQL 请求后，提取请求中的查询逻辑（如上例年龄=[35] AND 机构=[北科大]），并生成“查询逻辑树”；依据该“查询逻辑树”对指定的数据库中的记录进行匹配，匹配采用“查询逻辑树”自底向上方式，首先完成叶子节点上的“密码属性匹配”，如图 3 中所示，索引查询凭证年龄=[35]与指定记录的加密索引字段[35]进行匹配，从而获取该节点的密码学值；之后依据这些值对中间节点完成“密码逻辑门”计算，最终归约到根节点获取归约值，上述步骤对应于 Retrieve 算法；最后服务器可判定归约值是否正确，以判决该记录是否属于查询查询范围之内，若该记录在查询范围中，隐私数据库利用 ξ 对该记录的会话密钥密文进行更新，用于将会话密钥的解密过程与用户标识进行绑定；

（3）结果数据处理：将所有已判定的查询目标记录聚集形成“加密数据视图 view”，其中每条记录包括密码化检索得到的归约值、查询的加密数据字段以及会话密钥密文，并将其返还客户端；

（4）数据解密：客户端使用“解密参数”和用户私钥对“加密数据视图 view”进行解密，并获得最终的查询结果数据。注意到，隐私数据库仅为数据使用者提供会话密钥密文和归约值，数据的解密则在客户端进行，一方面密文的传输避免了数据在传输过程中的泄露问题，另一发面则是用户需用其密钥解密密文，从而完成对用户的身份认证。

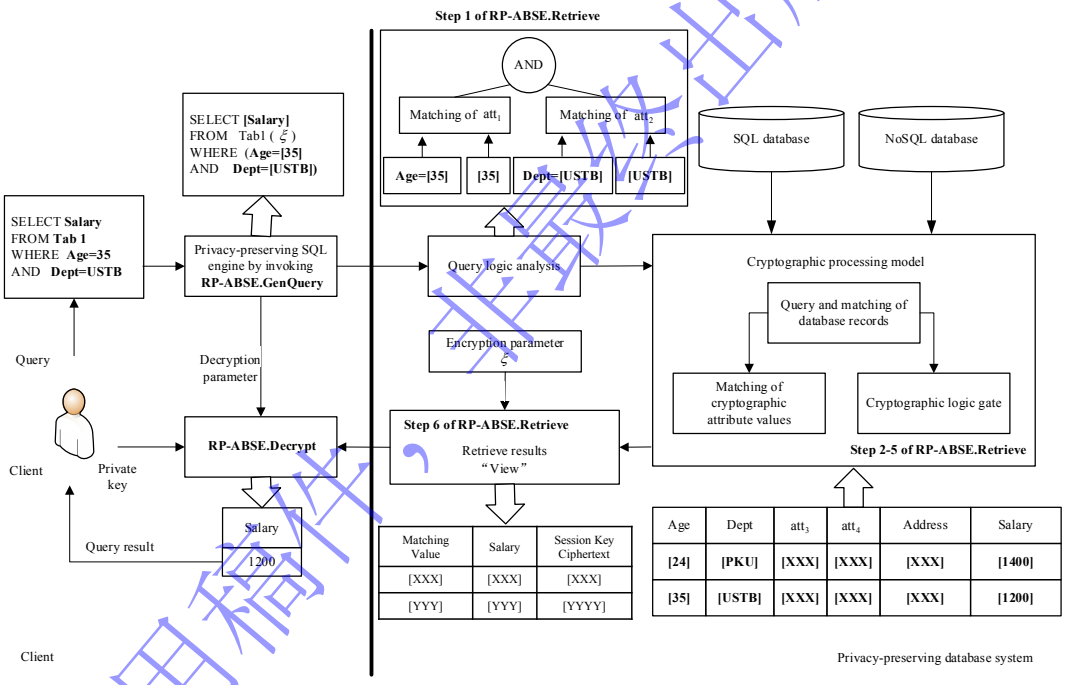

图 3 隐私数据库中 SELECT 语句的处理流程示意图

Fig. 3 The workflow of SELECT statement in the privacy-preserving database

上述仅对增加和查询操作进行介绍，本系统也可实现诸如删除、修改等其他数据库操作。

3 隐私数据库系统密码方案构造

3.1 前提知识

作为 NIST 公布的后量子密码算法的重要候选之一，LBC 被广泛认为可以抵抗量子计算攻击。文献[35]对后量子密码系统安全强度进行了总结，指出 LBC 方案所依赖的假设为 NP 困难问题，目前均无法通过量子算法解决，如最短向量问题（Shortest Vector Problem, SVP）和最接近向量问题（Closest Vector Problem, CVP）。文献[36]则指出 SIS 问题和 LWE 问题在最坏情况下与近似格问题一样困难，并给出了保证安全强度同时的更小参数设计。

此外，LBC 方案中密钥和密文通常为具有较短长度分量的向量，使得其易于编程而无需对大整数进行截断处理，在计算方面更具优势[37]；同时，通过引入特殊结构的理想格[38]，进一步降低了 LBC 方案的密文和密钥存储代价（相比于常规 LWE 降低了 n 倍）[39]。

基于上述分析，本文将所提 RP-ABSE 构造在理想格 R_q 下。令 n 为 2 的幂次，R 表示多项式商环 $\mathbb{Z}[x]/(x^n+1)$，R_q 则表示商环 $\mathsf{R}/q\mathsf{R}=\mathbb{Z}_q[x]/(x^n+1)$。令 $\mathbf{a}^t\in\mathsf{R}_q^m$ 为理想格上 m 维的向量，其中上标 t 代表转置，定义该向量对应的对偶格为 $\Lambda_q^u(\mathbf{a}^t)=\{\mathbf{x}\in\mathsf{R}_q\ s.t.,\mathbf{a}^t\cdot\mathbf{x}=u\}$，其中 $u\in\mathsf{R}_q$ 为任意陪集。最后，令 X 为 R_q 上误差分布。

本系统利用文献[40]所提的原像高斯采样器来构造用户私钥以及查询凭证，该采样器是基于 Gadget-格（G-格）$\Lambda_q^u(\mathbf{g}^t)$ 设计的，其中 $\mathbf{g}^t=(b^0,b^1,\cdots,b^{k-1})\in\mathsf{R}_q$，$u$ 为任意陪集，$b\geq 2$ 为任意基，并且 $k=\lceil\log_b q\rceil$。进一步，以下给出原像高斯采样器中陷门构造函数的定义：

定义 1（采样陷门生成[40]） 对于给定的多项式商环 R_q，存在多项式时间下随机化算法 $\mathrm{TrapGen}(\mathsf{R}_q)$，可输出公共向量 $\mathbf{a}^t\in\mathsf{R}_q^{k+2}$ 和陷门 $\mathbf{R}\in\mathsf{R}_q^{2\times k}$。

本文令 $\mathbf{a}^t=(\bar{\mathbf{a}}^t,-\bar{\mathbf{a}}\cdot\mathbf{R})$，其中 $\bar{\mathbf{a}}^t\in\mathsf{R}_q^2$ 为随机选择的。在进行原像采样过程中，可将用户标识或属性信息通过哈希函数编码为标签元素 $h\in\mathsf{R}_q$，并将公共向量 $\mathbf{a}^t$ 替换为向量 $\mathbf{a}_h^t=\mathbf{a}^t+((0,0),h\cdot\mathbf{g}^t)\in\mathsf{R}_q^{k+2}$，之后可调用原像高斯采样从 $\Lambda_q^u(\mathbf{a}_h^t)$ 的高斯分布中抽取短原像向量。以下给出原像高斯采样器的定义：

定义 2（原像高斯采样[41]） 对于给定公共向量 $\mathbf{a}^t\in\mathsf{R}_q^{k+2}$、陷门 $\mathbf{R}\in\mathsf{R}_q^{2\times k}$、任意陪集 $u\in\mathsf{R}_q$、标签元素 $h\in\mathsf{R}_q$ 和高斯参数 σ，存在多项式时间下随机化算法 $\mathrm{SamplePre}(\mathbf{a}^t,\mathbf{R},u,h,\sigma)$ 可从格 $\Lambda_q^u(\mathbf{a}^t)$ 的以 σ 为标准差和以 0 为中心的高斯分布中抽取原像向量 $\mathbf{z}\in\mathsf{R}_q^{k+2}$，使得 $\mathbf{a}^t\cdot\mathbf{z}=u$。

为简化描述，本文将高斯参数 σ 省略，将原像高斯采样函数设为 $\mathrm{SamplePre}(\mathbf{a}^t,\mathbf{R},u,h,\sigma)\rightarrow\mathbf{z}$。

此外，本文拟采用文献[42]所提的小策略矩阵算法将单调查询策略转化为线性秘密共享矩阵，即 $\mathrm{SPMGen}(\Pi)\rightarrow\mathbf{M}$。其定义如下所示：

定义 3（小策略矩阵[42]） 对于包含 l 个属性谓词的策略 Π 和理想格 R_q，矩阵 $\mathbf{M}$ 在 R_q 上被称为小策略矩阵，若其满足以下性质：

（1）使用算法 $\mathrm{SPMGen}(\cdot)$，可将 Π 转化为矩阵 $\mathbf{M}\in\mathsf{R}_q^{l\times m}$，该矩阵中每个元素均属于 $\{-1,0,1\}$；

（2）存在候选子矩阵 $\mathbf{M}'\subset\mathbf{M}$，其逆矩阵的行列式满足 $\left|(\mathbf{M}')^{-1}\right|=\pm 1$，且 $(\mathbf{M}')^{-1}$ 的第一行所有元素均为 1，即 $(\mathbf{M}')_1^{-1}=(1,1,\cdots 1)$。

由上述定义可知，对于给定的 $\mathbf{M}$，其任意候选可逆子矩阵的逆矩阵中首行分量均为 1，即 $(\mathbf{M}')_1^{-1}=(1,1,\cdots 1)$。该特点可有效降低解密过程中由向量相乘所带来的累积误差过大问题，即给定 R_q 上的误差向量 $\mathbf{e}^t$，其与 $(\mathbf{M}')_1^{-1}$ 相乘的结果为累加误差 $\sum_{\forall e_i\in\mathbf{e}^t}e_i$ 而非累乘。

最后，本系统是基于理想格上的非齐次短整数解（Inhomogeneous Short Integer Solution over Ring, R-ISIS）和带误差学习（Decisional Learning With Error over Ring, R-DLWE）这两个困难假设所构造的，以下给出其定义：

定义 4（R-ISIS 假设[43]） 给定从均匀随机分布中的向量 $\mathbf{a}^t\in\mathsf{R}_q^m$、陪集 $u\in\mathsf{R}_q$ 和正实界参数 $\beta>0$，不存在多项式时间算法以不可忽略的优势检索出 $\mathbf{z}\in\mathsf{R}_q^m$，使得 $\mathbf{a}^t\cdot\mathbf{z}=u$ 且 $0<\|\mathbf{z}\|\leq\beta$。

定义 5（R-DLWE 假设[44]） 给定理想格 R_q 和 R 上的误差分布 X 以及挑战喻言机 O，其中 O 的类别有以下两种：

- O_s：输出样本 $(a, as+e)\in R_q\times R_q$，其中 $s\in R_q$ 是从均匀分布中随机抽取的恒定不变量，$a\in R_q$ 和 $e\in X$ 为随机选择的；
- O_S：输出真正均匀随机抽取的 $(a,b)\in R_q\times R_q$。

不存在多项式时间算法以不可忽略的优势区分 O 的类别。

3.2 密码方案构造

如第 1.2 节中提出的问题可知，对于结构化存储的关系数据库而言，查询策略应嵌入至查询凭证中，用以检索数据库中与策略相符的数据密文，该思想更类似于 KP-ABE 而非 CP-ABE。因此，本文以 KP-ABE 为基础，在格密码体制下构造了检索策略属性基可搜索加密方案 RP-ABSE，以提供抗量子攻击能力，具体构造如下所示：

（1）算法 1 给出了系统初始化算法的详细步骤，由隐私 SQL 引擎执行，用于生成系统主私钥 msk 和公钥 pk，并将 pk 在系统中公布。由于在后续加密阶段中，索引字段和数据字段一经加密便会在数据库中长期存储，因此将用于加密的随机因子 s 作为 msk 的一部分。

（2）算法 2 给出了用户私钥生成算法的具体步骤，其由隐私 SQL 引擎执行。该算法首先需要依据用户标识 ID 将公共向量 $\mathbf{a}^t$ 转化为对应于 ID 的向量 $\mathbf{a}_{\mathrm{ID}}^t$，并调用原像采样函数 $\mathrm{SamplePre}(\cdot)$ 来为该用户生成短原像向量作为其私钥。

表 1 Setup 算法

Table 1 Algorithm Setup

Algorithm 1 $\mathrm{Setup}(n)\rightarrow(\mathrm{pk},\mathrm{msk})$
1. Generate $R_q=\mathbb{Z}_q[x]/(x^n+1)$ according to n;
2. Generate $(\mathbf{a}^t,\mathbf{R})\leftarrow \mathrm{TrapGen}(R_q)$;
3.Randomly choose $u,s\in R_q$, and select a hash function $h:\{0,1\}^*\rightarrow R_q$;
4. return $\mathrm{msk}=(\mathbf{R},s,h)$ 和 $\mathrm{pk}=(R_q,\mathbf{a}^t,u)$.

表 2 GenKey 算法

Table 2 Algorithm GenKey

Algorithm 2 $\mathrm{GenKey}(\mathrm{msk},\mathrm{ID})\rightarrow \mathrm{sk}_{\mathrm{ID}}$
1. Compute $\mathbf{a}_{ID}=\mathbf{a}+((0,0),h(\mathrm{ID})\mathbf{g})$;
2. Sample $T_{\mathrm{ID}}\leftarrow \mathrm{SamplePre}(\mathbf{a}_{\mathrm{ID}},\mathbf{R},u,h(\mathrm{ID}))$, where T_{ID} satisfies $\mathbf{a}_{\mathrm{ID}}\cdot T_{\mathrm{ID}}=u$;
3. **return** $\mathrm{sk}_{\mathrm{ID}}=T_{\mathrm{ID}}$。

（3）算法 3 给出了明文数据加密的完整步骤，其由隐私 SQL 引擎执行，并将密文数据存储在数据库中。该算法主要分为两部分，其中第一部分对应于第 1-3 步，用于依据公共向量 $\mathbf{a}^t$ 将索引字段集合 $\{\mathrm{att}_i\}_{i=1}^N$ 转化为加密索引字段集合 $\{c_i\}_{i=1}^N$；第二部分对应于第 4 步，用于对会话密钥 ek 进行加密，即将 ek 封装为密文 $c'=\tau\cdot s+\mathrm{ek}\cdot\lfloor q/2\rfloor+e'$，此处的 ek 是用于对数据字段进行对称加密的密钥，只有当查询策略与加密索引字段集合相匹配，才可恢复该密钥从而解密数据字段。

（4）算法 4 给出了查询生成函数的具体执行步骤，其由隐私 SQL 引擎执行来将用户提交的查询请求（即策略）转化为与策略相绑定的查询凭证，以及加密参数和解密参数。该算法分为三部分，第一部分对应于第 1-4 步，用于生成密码学查询策略，该部分依据策略 Π 中索引判定谓词抽取短原像向量用于以线性共享随机选取的秘密值 r_1，在该过程中小策略矩阵被引入以降低加密索引字段与原像向量匹配（对应于算法 5）所带来的累积误差；第二部分对应于第 5 步，用于为该用户生成解密参数，其实际上是对共享值 r_1 和随机选取的秘密值 s' 的密码学封装；第三部分对应第 6-7 步，用于生成加密参数 $\xi=(\xi_1,\xi_2)$，其中 ξ_1 是用于将会话密钥密文与用户标识绑定的参数，其随后将被发送至隐私数据库系统，用于对会话密钥密文 c' 的更新；而 ξ_2 是用于检索匹配的参数。

表 3 Encrypt 算法

Table 3 Algorithm Encrypt

Algorithm 3 $\mathrm{Encrypt}(\mathrm{msk},\{\mathrm{att}_i\}_{i=1}^N,\tau,\mathrm{ek})\rightarrow(\{c_i\}_{i=1}^N,c')$
1. **for** $i\in[1,N]$ **do**;
2. Compute $\mathbf{a}_i^t=\mathbf{a}^t+((0,0),h(\mathrm{att}_i)\mathbf{g}^t)$, and generate $c_i=\mathbf{a}_i^t s+\mathbf{e}_i^t$, where;
3. **end for**
4. Compute $c'=\tau\cdot s+\mathrm{ek}\cdot\lfloor q/2\rfloor+e'$, where $e'\in X$;

表 4 GenQuery 算法

Table 4 Algorithm GenQuery

Algorithm 4 $\mathrm{GenQuery}(\mathrm{msk},\Pi,\mathrm{ID})\rightarrow(\{T_i\}_{\mathrm{att}_i\in\Pi},\omega_{\mathrm{ID}},\xi)$
1. Generate $\mathbf{M}\in\{-1,0,1\}^{l\times m}\leftarrow \mathrm{SPMGen}(\Pi)$, and randomly choose a vector $\mathbf{r}=(r_1+\tau,\cdots,r_m)\in R_q^m$, further compute $\lambda=\mathbf{M}\cdot\mathbf{r}$;
2. **for** $\mathrm{att}_i\in\Pi$ **do**
3. Extract $T_i\leftarrow \mathrm{SamplePre}(\mathbf{a}_i^t,\mathbf{R},\lambda_i,h(\mathrm{att}_i))$ s.t.,

4. **return** $\left(\{c_i\}_{i=1}^{N}, c'\right)$.

$\mathbf{a}_i^t \cdot T_i = \lambda_i$;

4. **end for**
5. Randomly choose $s' \in \mathsf{R}_q$, further compute $c'' = \mathbf{a}_{\mathrm{ID}}^t \cdot s' + (\mathbf{e}'')^t$ and $t = r_1 \cdot s + u \cdot s' + e''$, where $\mathbf{e}'', e'' \in \mathsf{X}$;
6. Compute $\xi_1 = h_{\mathrm{ID}} \cdot s' + e_h$, where $e_h \in \mathsf{X}$;
7. Compute $\xi_2 = (r_1 + \tau) \cdot s + e_v$, where $e_v \in \mathsf{X}$;
8. **return** $\left(\{T_i\}_{\mathrm{att}_i \in \Pi}, \omega_{\mathrm{ID}}, \xi = (\xi_1, \xi_2)\right)$.

（5）算法 5 给出了检索函数的具体构造，其由隐私数据库系统执行。假设待查询数据的索引集合 Σ 满足查询策略 Π，则可依据 Σ 从 $\mathbf{M}$ 中检索出候选可逆子矩阵 $\mathbf{M}'$，其逆矩阵的首行 $(\mathbf{M}')_1^{-1}$ 为全 1 向量。之后针对 $\forall att_i \in \Sigma$，计算 $d_i = c_i \cdot T_i = \left(\mathbf{a}_i^t \cdot s + \mathbf{e}_i^t\right) \cdot T_i = \lambda_i \cdot s + \mathbf{e}_i^t \cdot T_i$；进一步，可由 $(\mathbf{M}')_1^{-1}$ 和向量 $\mathbf{d} = (d_i)_{\mathrm{att}_i \in \Sigma}$ 计算出正确的归约值 φ，即

$$\varphi = (\mathbf{M}')_1^{-1} \cdot \mathbf{d} = (\mathbf{M}')_1^{-1} \cdot \left(\mathbf{M}' \cdot \mathbf{r} \cdot s + \left(\mathbf{e}_i^t T_i\right)_{\mathrm{att}_i \in \Sigma}\right) = (1, 0, \cdots, 0) \cdot \mathbf{r} \cdot s + \sum_{\mathrm{att}_i \in \Sigma} \mathbf{e}_i^t T_i = (r_1 + \tau) \cdot s + \sum_{\mathrm{att}_i \in \Sigma} \mathbf{e}_i^t T_i$$

此外，该归约值可通过隐私 SQL 引擎提供的依据 ξ_2 来判定索引匹配是否成功，即计算 $|\varphi - \xi_2|$，若该归约值 φ 是正确的，则有 $|\varphi - \xi_2| = \left|(r_1 + \tau) \cdot s + \sum_{\mathrm{att}_i \in \Sigma} \mathbf{e}_i^t \cdot T_i - (r_1 + \tau) \cdot s - e_v\right| = \left|\sum_{\mathrm{att}_i \in \Sigma} \mathbf{e}_i^t \cdot T_i - e_v\right|$，因此对于 $i \in [1, n]$，该绝对值的每个多项式系数 $|\varphi - \Delta|_i$ 都应满足 $|\varphi - \Delta|_i \le q/4$。对于每条记录，若索引匹配成功，则将其会话密钥密文 c' 更新为 $\tilde{c} = c' + \xi_1 = \tau \cdot s + \mathrm{ek} \cdot \lfloor q/2 \rfloor + h_{ID} \cdot s' + e' + e_h$，并输出 $(\varphi, \tilde{c})$；反之，则输出无效归约值 $\perp$。

最后，将匹配成功的加密数据字段整合在一起发送给用户。值得注意的是，会话密钥密文更新流程的目的是将待查询密文与用户标识相绑定，用于对用户身份的认证；同时，即使 Π 与多条待查询数据相对应，也仅需为用户生成一个解密参数 ω_{ID}；同时，即使有 N'' 条数据待查询，对这些数据密文进行更新也仅需 N'' 次加法。

（6）算法 6 给出了数据解密算法的具体执行步骤，其在客户端被执行。首先，授权用户可依据其私钥计算出 d'，即 $d' = c'' \cdot T_{\mathrm{ID}} = \left(\mathbf{a}_{\mathrm{ID}}^t \cdot s' + (\mathbf{e}'')^t\right) \cdot T_{\mathrm{ID}} = (u + h_{\mathrm{ID}}) \cdot s' + (\mathbf{e}'')^t \cdot T_{\mathrm{ID}}$；之后，可计算

$$t' = \varphi - t = (r_1 + \tau) \cdot s + \sum_{\mathrm{att}_i \in \Sigma} \mathbf{e}_i^t \cdot T_i - r_1 \cdot s - u \cdot s' - e'' = \tau \cdot s - u \cdot s' + \sum_{\mathrm{att}_i \in \Sigma} \mathbf{e}_i^t \cdot T_i - e''$$

进一步，计算

$$t'' = t' + d' = \tau \cdot s - u \cdot s' + \sum_{\mathrm{att}_i \in \Sigma} \mathbf{e}_i^t \cdot T_i - e'' + (u + h_{\mathrm{ID}}) \cdot s' + (\mathbf{e}'')^t \cdot T_{\mathrm{ID}} = \tau \cdot s + h_{\mathrm{ID}} \cdot s' + \sum_{\mathrm{att}_i \in \Sigma} \mathbf{e}_i^t \cdot T_i - e'' + (\mathbf{e}'')^t \cdot T_{\mathrm{ID}}$$

之后，会话密钥 ek 可依据 t' 从 $\mathrm{Round}(\tilde{c} - t'')$ 中恢复，即 $\mathrm{Round}(\tilde{c} - t'') = \mathrm{Round}\left(\mathrm{ek} \cdot \lfloor \frac{q}{2} \rfloor + \bar{e}\right) = \mathrm{ek}$，其中，$\bar{e} = e' + e_h + \sum_{att_i \in \Sigma} \mathbf{e}_i \cdot T_i - e'' + (\mathbf{e}'')^t \cdot T_{\mathrm{ID}}$，$\mathrm{Round}(\tilde{c} - t'')$ 的执行步骤为：$(c' - t'')$ 的第 i 个系数接近于 0，则 ek 的第 i 个系数为 0；反之，则设为 1。此处，$\bar{e}$ 应满足 $\bar{e} < q/4$ 以保证解密正确性。最后，用户可利用 ek 对加密数据字段进行解密操作。

表 5 Retrieve 算法

Table 5 Algorithm Retrieve

Algorithm 5 Retrieve $\left(\Pi, \mathbf{M}, \{T_{\mathrm{att}_i \in \Pi}\}, \Sigma, \{c_i\}_{\mathrm{att}_i \in \Sigma}, \xi, c'\right) \to (\varphi, \tilde{c})$

1. According to the authorized attribute set Σ, the candidate reversible submatrix $\mathbf{M}' \subseteq \mathbf{M}$ can be retrieved from $\mathbf{M}$;
2. **for** $\mathrm{att}_i \in \Sigma$ **do**
3. Compute $d_i = c_i \cdot T_i$;
4. **end for**
5. Let $\mathbf{d} = (d_i)_{att_i \in \Sigma}$, and compute $\varphi = (\mathbf{M}')_1^{-1} \cdot \mathbf{d}$;
7. **if** $|\varphi - \xi_2|_j < q/4$ for $j \in [1, n]$ **then**
8. Compute the updated ciphertext $\tilde{c} = c' + \xi$;
9. **return** $(\varphi, \tilde{c})$;

表 6 Decrypt 算法

Table 6 Algorithm Decrypt

Algorithm 6 Decrypt $(\mathrm{sk}_{\mathrm{ID}}, \omega_{\mathrm{ID}}, \varphi, \tilde{c}) \to \mathrm{ek}$

1. Compute $d' = c'' \cdot T_{ID}$;
2. Compute $t' = \varphi - t$, and generate $t'' = t' + d'$;
3. Compute $\mathrm{ek} = \mathrm{Round}(\tilde{c} - t'')$;
4. **return** ek.

10. **else return** $\perp$.

4 安全性分析

1.索引查询凭证的选择策略攻击下的不可伪造性（Existential Unforgeability against Chosen-Policy Attack，EU-CPA）：该特征确保了即使敌手学习了与其他访问策略相绑定的索引查询凭证，也无法伪造新的索引查询凭证。其可被描述为敌手和挑战者之间的博弈，如下所示：

（1）初始化：敌手 A 宣称其要攻击的策略 Π^*，该策略至少包含一个授权集合 Σ^*，而 Σ^* 至少包含一个索引 att^*；

（2）设置：挑战者 B 模拟 Setup 算法来生成系统公钥，进一步模拟 Encrypt 算法来为给定索引字段集合生成加密索引字段，最后将系统公钥和加密索引字段发送至 A；

（3）学习：A 针对不同的 Π 向 B 发起索引查询凭证查询，其中 $att^* \notin \Pi$；针对每一次询问，B 模拟 GenQuery 算法生成索引查询凭证集合，并将其发送至 A；

（4）伪造：通过 B 与 A 的多次交互，A 伪造出对应于 Σ^* 的索引查询凭证 $\{T_i^*\}_{att_i \in \Sigma^*}$；若 $\{T_i^*\}_{att_i \in \Sigma^*}$ 有效，则称 A 赢得了博弈。

理论 1（EU-CPA）在 R-ISIS 困难假设下，索引查询凭证在选择策略攻击下存在性不可伪造的。

证明：假设存在一个多项式时间敌手 A 可以不可忽略的优势来伪造与指定策略 Π^* 对应的索引查询凭证，即 $Adv^{\text{EU-CPA}}(\mathsf{A}) \geq \grave{o}$，则本证明的目标是基于 A 的优势来构造模拟器 B 来求解 R-ISIS 问题，即给定 $(\mathbf{a}^t, u) \in \mathsf{R}_q^{k+2} \times \mathsf{R}_q$，输出 $T \in \mathsf{R}_q^{k+2}$ 满足 $\mathbf{a}^t \cdot T = u$ 且 $\|T\| < \beta$。此外，在本证明中，B 可访问 R-ISIS 喻言机来获取除挑战实例 $(\mathbf{a}, u)$ 之外任意 $(\mathbf{a}', u')$ 的短原像向量。B 的构造如下所示：

（1）初始化阶段：敌手 A 宣称要攻击的策略 Π^* 和数据表标识 τ^*，其中 Π^* 至少包含了一个授权集合 Σ^*，该集合至少包含一个索引字段 att^*；

（2）设置阶段：B 向 R-ISIS 喻言机来获取 $(\mathbf{a}^*, r^*)$，设置 $\mathbf{a}^t = \mathbf{a}^* - ((0,0), h(\text{att}^*)\mathbf{g})$，并随机选择 $s \in \mathsf{R}_q$；此外，B 为索引集合 $\{\text{att}_i\}_{i=1}^N$ 生成加密索引字段，即针对 $i \in [1, N]$，计算 $c_i = \mathbf{a}_i^t \cdot s + \mathbf{e}_i$，其中，$\mathbf{a}_i^t$ 满足 $\mathbf{a}_i^t = \mathbf{a}^t + ((0,0), h(\text{att}_i)\mathbf{g})$，$\mathbf{e}_i \in \mathsf{X}$ 为随机选择的；最后，B 将 $\mathbf{a}^t$ 和 $\{c_i\}_{i=1}^N$ 发送给 A；

（3）学习阶段：A 选择不同策略 Π 和数据表标识 τ 并向 B 发起询问，其中 Π 不包含索引 att^*，以及 $\tau \neq \tau^*$；对于每次询问，B 执行以下步骤：

1) 生成 $\mathbf{M} \in \{-1, 0, 1\}^{l \times m} \leftarrow \text{SPMGen}(\Pi_i)$，并生成随机向量 $\mathbf{r} = (r_1 + \tau, \cdots, r_m) \in \mathsf{R}_q^m$；
2) 计算 $\lambda = (\lambda_1, \cdots, \lambda_l) = \mathbf{M} \cdot \mathbf{r}$；
3) 对于 $\forall \text{att}_i \in \Pi_j$，B 访问 R-SIS 喻言机获取 T_i，使得 $\mathbf{a}_i^t \cdot T_i = \lambda_i$，其中 $\mathbf{a}_i^t = \mathbf{a}^t + ((0,0), h(\text{att}_i)\mathbf{g}^t)$；
4) 最后，B 将 $\{T_i\}_{\text{att}_i \in \Pi}$ 发送至 A；

（4）伪造阶段：假设 B 随机选择 r' 作为秘密共享值，A 需伪造一个有效的索引查询凭证集合 $\{T_i^*\}_{\text{att}_i \in \Pi^*}$，其子集 $\{T_i^*\}_{\text{att}_i \in \Sigma^*} \subseteq \{T_i^*\}_{\text{att}_i \in \Pi^*}$ 可与 Σ^* 对应的加密索引集合匹配生成有效的归约值。其伪造阶段由如下交互过程所示：

1) B 设置共享参数 $r' = r^* - \tau^*$；
2) 对于 $\forall \text{att}_i \in \Sigma^* / \text{att}^*$，B 要求 A 伪造对应的 T_i^* 满足 $\|T_i\| \leq \beta$，A 伪造出 T_i^* 后，B 计算 $r' = r' + \mathbf{a}_i^t \cdot T_i$；
3) 对于索引 att^*，B 将 $r' = r^* - \tau^* + \sum_{\text{att}_i \in \Sigma^* / \text{att}^*} \mathbf{a}_i^t \cdot T_i$ 和 τ^* 发送至 A，并要求 A 伪造 T^* 使得

$\sum_{\text{att}_i\in\Sigma^*/\text{att}^*}\mathbf{a}_i^t\cdot T_i+\mathbf{a}^*\cdot T^*=r'+\tau^*$，这要求$T^*$满足$\mathbf{a}^*\cdot T^*=r^*$和$\|T^*\|\le\beta$。

分析：在上述伪造阶段中，共享秘密值$(r'+\tau^*)$是在A与B多次交互后才确定的，其原因是r'本身就是由系统管理员每次运行GenQuery算法时随机选择的而非公开的；同时在密码系统实际运行中，A无法获取r'的值，因此B将$(r'+\tau^*)$发送给A本身就是给A的额外优势。若伪造的T^*是有效的，则其满足$\mathbf{a}^*\cdot T^*=r^*$且$\|T^*\|\le\beta$，因此B可将$T^*$作为R-ISIS挑战实例的解。这意味着B的优势为$\text{Adv}_{\text{R-SIS}}(\text{B})=\text{Adv}^{\text{EU-CPA}}_{\text{RP-ABSE}}(\text{A})\ge$ò。然而当前不存在任意多项式时间算法可求解R-ISIS问题，因此有$\text{Adv}_{\text{R-SIS}}(\text{B})<$ò。

2.语义安全性：

进一步，本文证明了所提密码系统满足语义安全对于指定的用户标识，该特征确保了密文不会泄露隐私信息若该用户没有获得授权。本文将所提密码系统的语义安全定义为在带有策略和用户标识查询的选择明文攻击下不可区分性（Indistinguishability with Policy and Identity Queries against Chosen-Plaintext Attacks，IND-PIQ-CPA），其可被定义如下敌手A与挑战者B之间的博弈：

（1）初始化阶段：A宣称其要攻击的用户标识ID^*；

（2）设置阶段：B模拟Setup算法生成公共参数$(\mathbf{a}^t,u^*)$，并将其公布至A；

（3）学习阶段：A针对不同的用户标识ID、会话密钥密文f和策略Π发起询问，对于每次查询，B生成用户私钥T_{ID}、加密索引字段集合$\{c_i\}_{\text{att}_i\in\Pi}$、数据密文$c'$、索引查询凭证集合$\{T_i\}_{\text{att}_i\in\Pi}$和解密参数$\omega_{\text{ID}}=(c'',t)$，其中$\text{ID}\ne\text{ID}^*$；

（4）挑战阶段：A选择两个消息m_0和m_1，并随机指定访问策略Π^*；B随机选择$b^*\in\{0,1\}$，并模拟Encrypt算法和GenQuery算法生成加密索引字段和解密参数，并将其发送给A；

（5）猜测阶段：A输出猜测$b'\in\{0,1\}$，若$b'=b^*$，则A在博弈过程中获胜。

理论2（IND-PIQ-CPA） 在R-DLWE困难假设下，所提密码系统满足在带有策略和标识查询选择明文攻击下的语义安全。

证明：假设存在多项式时间敌手A可以不可忽略的优势在带有策略和标识查询的选择明文攻击下可破坏所提密码系统的语义安全，即$Adv^{\text{IND-PIQ-CPA}}_{\text{RP-ABSE}}(\text{A})\ge$ò。本证明的目标是基于A的优势来构造模拟器B来求解R-DLWE问题，即给定R-DLWE喻言机O，B可分辨O是O_s还是$\text{O}_\$$。在学习阶段，B可访问R-SIS喻言机来获取除指定挑战实例外的任意二元组$(\mathbf{a}',u)$的短原像向量。B的构造如下所示：

（1）初始化阶段：A宣称其要攻击的用户标识ID^*；

（2）设置阶段：B访问O来获取挑战实例$(\mathbf{x},\mathbf{y})\in\text{R}_q^{k+4}\times\text{R}_q^{k+4}$。令$\mathbf{x}'=(x_3,\cdots,x_{k+4})$和$\mathbf{y}'=(y_3,\cdots,y_{k+4})$，B并随机选择$s\in\text{R}_q$，设置公共参数$u^*=x_1$，$\text{ID}^*$的哈希值$h^*=x_2$和公共向量$\mathbf{a}^t=\mathbf{x}'-((0,0),h^*\mathbf{g}^t)$；最后，B将$(\mathbf{a}^t,u^*)$公布给A；

（3）学习阶段：A针对不同的用户标识$\text{ID}\ne ID^*$，数据会话密钥ek和策略Π发起询问。对于每次查询，B首先生成加密索引字段和会话密钥密文，其步骤如下所示：

1) 针对$\forall\text{att}_i\in\Pi$，生成加密索引字段$c_i=\mathbf{a}_i^t\cdot s+\mathbf{e}_i$，其中$\mathbf{e}_i\in\text{X}$；

2) 针对ek，随机选取$\tau,h_{\text{ID}}\in\text{R}_q$和$e'\in\text{X}$，并生成数据密文$\tilde{c}=\tau\cdot s+h(\text{ID})\cdot s'+e'+\text{ek}\cdot\lfloor q/2\rfloor$；

其次，B生成索引查询凭证集合与解密参数，其具体步骤如下所示：

1) B生成$\mathbf{M}\leftarrow\text{SPMGen}(\Pi)$，并生成$\mathbf{r}=(r_1+\tau,r_2,\cdots,r_m)$，计算$\lambda=\mathbf{M}\cdot\mathbf{r}$；

2) 对于$\forall\text{att}_i\in\Pi$，访问R-SIS喻言机获取$T_i$使得$\mathbf{a}_i^t\cdot T_i=\lambda_i$；

3) 生成$c''=\mathbf{a}_{\text{ID}}^t\cdot s'+(\mathbf{e}'')^t$和$t=r_1\cdot s+u^*\cdot s'+e''$，其中$\mathbf{a}_{\text{ID}}^t=\mathbf{a}^t-((0,0),h_{\text{ID}}\mathbf{g}^t)$，$s'\in\text{R}_q$和$(\mathbf{e}'')^t,e''\in\text{X}$是随机选择的；

再次，B针对ID访问R-SIS喻言机获取T_{ID}使得$\mathbf{a}_{\text{ID}}^t\cdot T_{\text{ID}}=u^*h_{\text{ID}}$；最后，B将$\{c_i\}_{\text{att}_i\in\Pi},\tilde{c},\{T_i\}_{\text{att}_i\in\Pi},c'',t$

和 T_{ID} 发送给 A ；

（4）挑战阶段： A 选择两个消息 m_0 和 m_1 ，并随机指定访问策略 Π^* ； B 执行以下步骤：

1) 对于 $\forall \mathrm{att}_i^* \in \Pi^*$ ，计算 $c_i^* = \mathbf{a}_i^* \cdot s + \mathbf{e}_i$ 作为加密索引字段，其中 $\mathbf{a}_i^* = \mathbf{a}^t + \left((0,0), h\left(\mathrm{att}_i^*\right)\mathbf{g}^t\right)$，$\mathbf{e}_i^t \in \mathsf{X}$ 和 $h\left(\mathrm{att}_i^*\right) \in \mathsf{R}_q$ 为随机选择的；

2) 随机选择 $\tau^* \in \mathsf{R}_q$ ，并生成会话密钥密文 $\tilde{c}^* = \tau^* \cdot s + y_2 + m_{b^*} \cdot \lfloor q/2 \rfloor$ ；

3) 计算 $c^{*\prime\prime} = \mathbf{y}'$ ，并生成 $t^* = r_1 \cdot s + y_1$ ；

4) 最后，将 $\left\{c_i^*\right\}_{\mathrm{att}_i^* \in \Pi^*}, \tilde{c}^*, c^{*\prime\prime}$ 和 t^* 发送给 A ；

（5）猜测阶段： A 输出猜测 b' ，若 $b' = b^*$ ，则 B 输出 1；反之则输出 0。

基于上述描述，可分两种情况进行分析：

（1）若 $\mathsf{O} = \mathsf{O}_s$ ，则所获取的 R-DLWE 实例满足 $\mathbf{y} = \mathbf{x} \cdot s^* + \mathbf{e}$ 对于挑战实例中的恒定值 $s^* \in \mathsf{R}_q$ ，这意味着所构造的会话密钥密文和解密参数具有以下正确的形式：

$$\begin{cases} c^{*\prime\prime} = \mathbf{x}' \cdot s^* + \left(e_2, \cdots, e_{k+3}\right) = \mathbf{a}_{ID^*} \cdot s^* + \mathbf{e} \\ t^* = r_1 \cdot s + x_1 \cdot s^* + e_1 = r_1 \cdot s + u^* \cdot s^* + e_1 \\ \tilde{c}' = \tau^* \cdot s + x_2 \cdot s^* + e_2 + m_{b^*} \cdot \lfloor q/2 \rfloor = \tau^* \cdot s + h^* \cdot s^* + e_2 + m_{b^*} \cdot \lfloor q/2 \rfloor \end{cases}$$

若 A 可以不可忽略优势 $\mathrm{Adv}_{\mathrm{RP\text{-}ABSE}}^{\mathrm{IND\text{-}PIQ\text{-}CPA}}(\mathsf{A}) \geq \grave{\mathrm{o}}$ 猜测出 b^* ，即 $b' = b^*$ ，则 B 可判断出 $\mathsf{O} = \mathsf{O}_s$ ；因此， B 求解 R-DLWE 问题的概率为 $\Pr\left[b' = b^* \mid \mathsf{O} \equiv \mathsf{O}_s\right] = \frac{1}{2} + Adv_{\mathrm{RP\text{-}ABSE}}^{\mathrm{IND\text{-}PIQ\text{-}CPA}}(\mathsf{A}) \geq \frac{1}{2} + \grave{\mathrm{o}}$ ；

（2）若 $\mathsf{O} = \mathsf{O}_{\$}$ ，则实例 $(\mathbf{x}, \mathbf{y})$ 为随机选择的，这意味着构造的会话密钥密文和解密参数不具有正确形式，因此 A 正确猜测出 b^* 值的概率仅为 $1/2$ ，所以 B 求解 R-DLWE 问题的概率为 $\Pr\left[b' = b^* \mid \mathsf{O} \equiv \mathsf{O}_{\$}\right] = 1/2$ ；

基于上述两种情况的分析， B 求解 R-DLWE 问题的优势为

$$\begin{aligned} \mathrm{Adv}_{\mathrm{R\text{-}DLWE}}^{\mathrm{IND}}(\mathsf{B}) &= \Pr\left[b' = b^*\right] - \frac{1}{2} \\ &= \Pr\left[b' = b^* \mid \mathsf{O} \equiv \mathsf{O}_s\right] \cdot \Pr\left[\mathsf{O} \equiv \mathsf{O}_s\right] + \Pr\left[b' = b^* \mid \mathsf{O} \equiv \mathsf{O}_{\$}\right] \cdot \Pr\left[\mathsf{O} \equiv \mathsf{O}_{\$}\right] - \frac{1}{2} \\ &= \frac{1}{2}\left(\left(\frac{1}{2} + \grave{\mathrm{o}}\right) + \frac{1}{2}\right) - \frac{1}{2} = \frac{\grave{\mathrm{o}}}{2} \end{aligned}$$

然而，当前不存在有效的多项式时间算法可以不可忽略的优势来求解 R-DLWE 问题，因此不存在多项式时间敌手不具备可忽略优势来破坏所提密码系统的语义安全，即 $\mathrm{Adv}_{\mathrm{RP\text{-}ABSE}}^{\mathrm{IND\text{-}PIQ\text{-}CPA}}(\mathsf{A}) < \grave{\mathrm{o}}$ 。

本文所提 RP-ABSE 方案是与关系数据库 SQL 查询语言相适配的，相比于传统的 ABSE 方案，其在支持对用户身份认证的同时，可支持基于查询策略的检索机制，而非简单的“或”关系。尽管该方案在功能上仍存在不足，缺乏对区间判定等检索功能的支持，但从密码学角度考虑，这些功能是有可能实现的，例如文献[45]可在双线性群下实现比较关系，文献[46]在有限域上实现了隐私计数等。因此在未来工作中，本文希望在原有系统基础上增强对查询策略的表达能力，设计支持除“等值”以外其他属性谓词构造（属于、不属于、比较或区间判定等），从而提升密码化检索效率。

5 结论

隐私数据库技术是保护数据要素开放、共享和交易以及实现外包数据库隐私安全的重要技术，对于建立数据市场化机制和数据的治理体系都具有重要支撑作用。本文构造了基于 RP-ABSE 的隐私数据库架构以适配关系数据库中 SQL 查询机制，该架构在关系数据表中支持多样化隐私字段存储，可通过隐私 SQL 引擎将 SQL 数据库检索转化为隐私 SQL 查询语句对密态数据进行细粒度的密码化检索处理。然而，当前方案对大数据集仍存在着检索时间长、查询能力单一等问题，因此隐私数据库在产业界应用还有较长距离，无论是在原理上、架构上和技术层面上还有待进一步改进。总

之，本文所提隐私数据库作为保护数据要素开放、共享和交易以及实现外包数据库隐私安全的重要技术，对于建立数据市场化机制和数据的治理体系都具有重要支撑作用。

致谢

本研究由以下基金项目支持：国家自然科学基金（61972032）；科技部国家重点技术研发计划（2018YFB1402702）。

参 考 文 献

[1] Yao Z B, Niu W J, Zhang Y, et al. Development and application of a rockburst database management system. *Chinese Journal of Engineering*, 2022, 44(5): 865.
（姚志宾, 牛文静, 张宇, 等. 岩爆数据库管理系统开发及应用. 工程科学学报, 2022, 44(5): 865.）

[2] Zhao Y, Chen J J. A survey on differential privacy for unstructured data content. *ACM Computing Surveys*, 2022, 54(10s): 1.

[3] Guo G L, Zhu Y, Chen C E, et al. Privacy-Preserving Queries Using Multisource Private Data Counting on Real Numbers in IoT. *IEEE Internet of Things Journal*, 2023, Early Access.

[4] Zhang K, Zhang Y, Li Y P, et al. A Blockchain-based Anonymous Attribute-based Searchable Encryption Scheme for Data Sharing. *IEEE Internet of Things Journal*, 2024, 11(1): 1685.

[5] Lu H, Yu R Y, Zhu Y, et al. Policy-driven data sharing over attribute-based encryption supporting dual membership. *Journal of Systems and Software*, 2022, 188: 111271.

[6] Cheon J H, Kim A, Kim M, et al. Homomorphic encryption for arithmetic of approximate numbers // *23rd International Conference on the Theory and Applications of Cryptology and Information Security*. Hong Kong, 2017: 409.

[7] Song D X D, Wagner D, Perrig A. Practical techniques for searches on encrypted data // *Proceeding 2000 IEEE symposium on security and privacy*. Berkeley, 2000: 44.

[8] Yin H J, Zhu Yan, Wang J, et al. Design and implementation of a smart-contract voting system based on zero-knowledge proof. *Chinese Journal of Engineering*, 2023, 45(4): 632.
（殷红建, 朱岩, 王静, 等. 基于零知识证明的智能合约投票系统设计与实现. 工程科学学报, 2023, 45(4): 632.）

[9] Shmueli E, Waisenberg R, Elovici Y, et al. Designing secure indexes for encrypted databases // *Data and Applications Security XIX*. Berlin, 2005: 54.

[10] Minaud B, Reichle, M. Dynamic local searchable symmetric encryption // *Annual International Cryptology Conference*. Cham, 2022: 91.

[11] Molla E, Rizomiliotis P, Gritzalis S. Efficient searchable symmetric encryption supporting range queries. *International Journal of Information Security*, 2023, 22: 785.

[12] Wang Q, Zhang X, Qin J, et al. A verifiable symmetric searchable encryption scheme based on the AVL tree. *The Computer Journal*, 2023, 66(1): 174.

[13] Hore B, Mehrotra S, Tsudik G. A privacy-preserving index for range queries // *Proceedings of the Thirtieth international conference on Very large data bases*. Toronto, 2004: 720.

[14] Popa R A, Redfield C M S, Zeldovich N, et al. CryptDB: processing queries on an encrypted database. *Communications of the ACM*, 2012, 55(9): 103.

[15] Wang Q, He M Q, Du M X, et al. Searchable encryption over feature-rich data. *IEEE Transactions on Dependable and Secure Computing*, 2016, 15(3): 496.

[16] Sun X Z, Zhou F C, Li Y X, et al. A database encryption scheme based on searchable encryption. *Chinese Journal of Computers*, 2021, 44(4): 806.

（孙僖泽, 周福才, 李宇溪, 等. 基于可搜索加密机制的数据库加密方案. 计算机学报, 2021, 44(4): 806.）

[17] Wang Y Y, Pan J X, Chen Y. Fine-grained secure attribute-based encryption. *Journal of Cryptology*, 2023, 36: 33.

[18] Zhong H, Cui J, Zhu W L, et al. Efficient and verifiable multi-authority attribute-based encryption scheme. *Journal of Software*, 2018, 29(7):2006.
（仲红,崔杰,朱文龙,许艳.高效且可验证的多授权机构属性基加密方案.软件学报,2018,29(7):2006.）

[19] Jangjou M, Sohrabi M K. A comprehensive survey on security challenges in different network layers in cloud computing. *Archives of Computational Methods in Engineering*, 2022, 29(6): 3587.

[20] Varri U S, Pasupuleti S K, Kadambari K V. Practical verifiable multi-keyword attribute-based searchable signcryption in cloud storage. *Journal of Ambient Intelligence and Humanized Computing*, 2023, 14(9): 11455.

[21] Varri U S. Privacy-Preserving Ciphertext-Policy Attribute-Based Search over Encrypted Data in Cloud Storage. *Journal of Computer Science and Technology*, 2023, 23(1): 7.

[22] Chaudhari P, Das M L. Keysea: Keyword-based search with receiver anonymity in attribute-based searchable encryption. *IEEE Transactions on Services Computing*, 2020, 15(2): 1036.

[23] Wang H Y, Li Y, Susilo W, et al. A fast and flexible attribute-based searchable encryption scheme supporting multi-search mechanism in cloud computing. *Computer Standards & Interfaces*, 2022, 82: 103635.

[24] Xu L, Xu C G, Liu J K, et al. Building a dynamic searchable encrypted medical database for multi-client. *Information Sciences*, 2020, 527: 394.

[25] Chen L Q, Zhang L Y, Chen Y. An attribute-based searchable encryption with verifiable database. *Journal of Cryptologic Research*, 2022, 9(5): 910.
（陈立全, 张林樾, 陈垚. 一种结合可验证数据库的属性可搜索加密方案. 密码学报, 2022, 9(5): 910.）

[26] Varri U S, Pasupuleti S K, Kadambari K V. CP-ABSEL: ciphertext-policy attribute-based searchable encryption from lattice in cloud storage. *Peer-to-Peer Networking and Applications*, 2021, 14: 1290-1302.

[27] Handa R, Krishna C R, Aggarwal N. Searchable encryption: a survey on privacy-preserving search schemes on encrypted outsourced data. *Concurrency and Computation: Practice and Experience*, 2019, 31(17): e5201.

[28] Andola N, Gahlot R, Yadav V K, et al. Searchable encryption on the cloud: a survey. *The Journal of Supercomputing*, 2022, 78(7): 9952.

[29] Xu P, Jin H, Wu Q H, et al. Public-key encryption with fuzzy keyword search: a provably secure scheme under keyword guessing attack. *IEEE Trans Comput*, 2013, 62(11):2266.

[30] Ding M Z, Gao F, Jin Z P, et al. An efficient public key encryption with conjunctive keyword search scheme based on pairings // *3rd IEEE International Conference on Network Infrastructure and Digital Content*. Beijing, 2012: 526.

[31] Hwang M S, Hsu S T, Lee C C. A new public key encryption with conjunctive field keyword search scheme. *Inf Technol Control*, 2014, 43(3):277.

[32] Farràs O, Ribes-González J. Provably secure public-key encryption with conjunctive and subset keyword search. *Int J Inf Secur*. 2019, 18: 533.

[33] Wang P, Chen B W, Xiang T, et al. Lattice-based public key searchable encryption with fine-grained access control for edge computing. *Future Generation Computer Systems*, 2022, 127: 373.

[34] Shor P W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, 1999, 41(2): 303.

[35] Fernández-Caramés T M, From pre-quantum to post-quantum IoT security: a survey on quantum-resistant cryptosystems for the Internet of Things. *IEEE Internet of Things Journal*, 2020, 7(7): 6457.

[36] Micciancio D, Peikert C. Hardness of SIS and LWE with small parameters// *Annual cryptology conference*. Berlin, 2013: 21.

[37] Lu H, Zhu Y, Chen E C, et al. Efficient key generation on lattice cryptography for privacy protection in mobile IoT crowdsourcing. *IEEE Internet of Things Journal*, 2024, 11(2): 1893.

[38] Lyubashevsky V, Peikert C, Regev O. On ideal lattices and learning with errors over rings// *29th Annual International Conference on the Theory and Applications of Cryptographic Techniques*, Berlin, 2010: 6110.

[39] Asif R. Post-quantum cryptosystems for Internet-of-Things: a survey on lattice-based algorithms. *IoT*, 2021, 2(1): 71.

[40] Micciancio D, Peikert C. Trapdoors for lattices: Simpler, tighter, faster, smaller // *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Berlin, 2012: 700.

[41] Lu H, Zhu Y, Chen E C, et al, efficient key generation on lattice cryptography for privacy protection in mobile IoT crowdsourcing. *IEEE Internet of Things Journal*, 2024, 11(2): 1893.

[42] Chen E, Zhu Y, Liang K, et al. Secure remote cloud file sharing with attribute-based access control and performance optimization. *IEEE transactions on cloud computing*, 2023, 11(1): 579.

[43] Langlois A, Stehlé D. Worst-case to average-case reductions for module lattices. *Designs, Codes and Cryptography*, 2015, 75(3): 565.

[44] Gür K D, Polyakov Y, Rohloff K, et al. Practical applications of improved Gaussian sampling for trapdoor lattices. *IEEE Transactions on Computers*, 2018, 68(4): 570.

[45] Wang Z J, Huang D J, Zhu Y, et al. Efficient Attribute-Based Comparable Data Access Control. *IEEE Transactions on Computers*, 2015, 64(12): 3430.

[46] Guo G L, Zhu Y, Chen E C, et al. Privacy-Preserving Queries Using Multisource Private Data Counting on Real Numbers in IoT. *IEEE Internet of Things Journal*, 2024, 11(7): 11353.