

面向气象数据的高效安全流通与共享机制研究

宋甫祺¹, 张汉胜², 饶 华²

(1. 江西科技学院 工程训练中心; 2. 江西科技学院 信息工程学院, 江西 南昌 330000)

摘要: 气象数据具备场景先导性、时效双重性与价值即用性, 要实现“到手即用”, 需依赖高效流通共享。然而, 当前气象数据面临安全流通体系架构缺失、价值安全定价困难及实时性要求高等问题。为此, 提出面向气象数据前后向安全的动态可搜索加密方法(RCFB), 支持联合关键字搜索与安全动态更新, 实现高效安全的数据流通。首先, 设计了基于区块链的“存—算—通”一体可信安全流通体系架构, 保障数据可信存储、可信计算与可信流通; 其次, 构建基于成本途径的多因子修正模型, 实现安全资产定价以量化数据价值。RCFB采用聚合方式以提升联合查询效率, 与现有方法(ODXT)相比, 能实现高效安全的联合关键字查询, 满足气象数据资产定价与可信流通需求。

关键词: 气象数据; 数据流通; 隐私保护; 资产定价

DOI: 10.11907/rjdk.261104

中图分类号: TP18

文献标识码: A

文章编号: 1672-7800(2026)005-0134-09

扫描二维码阅读全文:



Research on Efficient and Secure Circulation and Sharing Mechanism for Meteorological Data

SONG Fuqi¹, ZHANG Hansheng², RAO Hua²

(1. Engineering Training Center, Jiangxi Science and Technology University;

2. School of Information Engineering, Jiangxi Science and Technology University, Nanchang 330000, China)

Abstract: Meteorological data has the characteristics of scene leading, timeliness dual, and value ready to use, and achieving "ready to use" relies on efficient circulation and sharing. However, we are currently facing issues such as the lack of a secure circulation system architecture for meteorological data, difficulties in pricing value security, and high real-time requirements. To this end, this article proposes a dynamic searchable encryption method for forward and backward security of meteorological data (RCFB), which supports joint keyword search and secure dynamic updates, achieving efficient and secure data circulation. Firstly, a blockchain based "storage computing communication" integrated trusted and secure circulation system architecture was designed to ensure the trusted storage, computation, and circulation of data. Secondly, a multi factor correction model based on cost approach is constructed to achieve pricing of secure assets and quantify the value of data. RCFB uses aggregation to improve the efficiency of joint queries. Compared with existing methods (ODXT), it can achieve efficient and secure joint keyword queries, meeting the needs of meteorological data asset pricing and trustworthy circulation.

Key Words: meteorological data; data circulation; privacy protection; asset pricing

0 引言

随着云计算技术的快速发展, 云存储已成为高效灵活的数据存储方式。越来越多的企业和用户将大规模数据外包至云服务器, 利用其强大的计算与存储资源实现高质量服务。气象数据广泛应用于气象预测、监控、预警等

工业生产和学术研究中, 但其大规模存储与管理面临严峻挑战。企业和政府部门将气象数据外包至云服务器, 这将导致用户失去对数据的直接物理控制权。云服务器提供商可能出于经济利益窃取高价值密度的气象数据, 造成严重的隐私泄露风险^[1]。因此, 如何解决云外包气象数据在共享过程中面临的安全存储和检索问题迫在眉睫。本文深入研究气象数据要素价值的安全流通体系架

收稿日期: 2026-03-31

基金项目: 江西科技学院自然科学项目(24ZRYB20)

作者简介: 宋甫祺(1996-), 男, 硕士, 江西科技学院工程训练中心助教, 研究方向为优化算法与空间加密; 张汉胜(1990-), 男, 硕士, 江西科技学院信息工程学院助教, 研究方向为同态加密算法; 饶华(1986-), 男, 硕士, 江西科技学院信息工程学院助教, 研究方向为气象数据分析。本文通讯作者: 宋甫祺。

构及安全共享方法,为构建高效安全的气象数据流通提供理论支撑与技术保障,从而促进气象数据要素市场的健康发展,加速气象数据价值的释放,同时,保障气象数据的隐私与安全,为数字经济的可持续发展奠定坚实的基础。

为此,本文提出一种面向气象数据的动态可搜索优化方法(DSSE with Rapid Conjective Queries, RCFB),支持联合关键字搜索和前后向安全动态更新,构建基于成本途径的多因子修正模型,以实现安全的气象数据资产定价。本文设计了基于区块链的可信安全数据要素流通体系架构,提出“存—算—通”一体的气象数据要素可信安全流通体系架构模型,支持气象数据要素在系统中可信存储、计算与流通。本文采用聚合方式,将所有关键字融合到一个条目中,极大地减少了联合关键字搜索的陷门数量。RCFB利用轻量级对称优化算法,降低了密文检索的计算开销。此外,通过存储最新更新的条目信息,RCFB可以极大地降低客户端的存储开销和计算开销。

1 相关工作

目前,数据安全流通机制的研究主要集中在复杂数据空间构造技术方面,但现有研究仍缺乏统一的数据安全流通机制作为参考,且数据基础设施的研究也较为分散^[2]。近年来,对称可搜索优化(Symmetric Searchable Encryption, SSE)技术受到了学术界广泛关注。Song等^[3]首次提出了SSE技术,SSE支持用户在不解密密文前提下执行关键字搜索。然而,现有的SSE方法仅支持静态数据下的关键字搜索^[4]。气象数据具有实时性和时效性特征,静态SSE方法无法支持气象数据更新状态下的密文检索,限制了气象数据的实用性^[5]。为了解决气象数据实时更新问题,Kamara等^[6]提出了一种动态可搜索优化(Dynamic Symmetric Searchable Encryption, DSSE)方法,能够支持数据动态更新下安全的关键字搜索。在DSSE中,用户首先优化气象数据,随后将优化的气象数据上传至云服务器。当气象数据实时更新时,用户可以对优化的气象数据进行增加、删除和修改操作,同时,保证云服务器无法从更新前后的气象数据中获取隐私信息。然而,在大规模优化的气象数据中进行动态更新,需要耗费极大的计算资源和通信开销。因此,DSSE通常允许泄露访问模式或者搜索模式,以保证云服务器能够实现动态的关键字搜索。然而,访问模式和搜索模式的隐私泄露可能会带来安全隐患,如通过文件注入攻击获取气象数据的机密信息等^[7]。因此,需要研究可信安全数据要素流通体系架构,为深入研究气象数据要素中的高效安全流通与查询方法提供理论支撑依据。

在数据要素价值定价与安全共享方面,国家数据局等部门发布的《“数据要素×”三年行动计划(2024—2026年)》同样聚焦于数据价值的定价及数据安全共享,要求持续深

度利用隐私计算、可信数据空间以及区块链等技术,提高多主体间数据流通与共享效率,从而促进数据合规高效流通使用^[8]。其中,学术界和产业界都在积极探索基于IDS的数据要素价值利用机制,以确保数据的合法性与合规性^[9]。国内在数据要素利用中的安全评估和风险防控技术方面也取得了显著进展,通过基于区块链和隐私计算技术,构建基于区块链的数据共享平台,采用同态优化和差分隐私技术,实现了数据的去中心化存储和传输,保证了数据共享的安全性和效率,保护数据隐私的同时实现了数据共享和利用^[10]。此外,为了增强DSSE的安全性,两种安全属性被提出,分别是前向安全和后向安全。其中,前向安全能够确保在服务器不能利用已有的搜索陷门的前提下,推测出新插入的文件的任何信息^[11];后向安全能够确保服务器无法访问已经被删除的文件^[12]。现有的DSSE方法通常仅关注如何权衡可搜索优化的效率和安全性,且已有的DSSE方法实现了前向安全与后向安全的动态更新,但仅支持单关键字搜索^[13]。如果用户需要执行精确的多关键字搜索,现有的DSSE必须重复执行单关键字搜索,且每次单关键字搜索后需要对查询结果进行过滤。这无疑会增加优化文检索的计算开销,且可能会泄露搜索模式或者访问模式。为了实现灵活和精确的联合关键字搜索,Patranabis等^[14]提出一种联合关键字搜索方法ODXT,但ODXT依赖于大量的群运算和指数操作,导致密文搜索和动态更新的计算开销较大。部分学者尝试优化DSSE的效率与安全性,郭宇韬等^[15]提出局域优化的前向安全DSSE方案,有效降低了搜索延迟;黄一才等^[16]构造的基于SH-VE的方案,实现了低开销的联合查询。

尽管国内外在数据要素流通与共享方面取得了一定进展,但在气象数据安全流通机制和架构上仍未形成统一标准和范式,容易导致气象数据在跨域、跨行业、跨部门流通时面临诸多障碍,难以确保气象数据流通的安全性和隐私性。此外,在气象数据要素价值定价和共享方面,仍缺乏对气象数据要素价值定价模型与动态更新下安全查询技术,限制了气象数据在合法合规前提下的高效利用和共享流通。

2 前置知识

前置知识具体包括动态可搜索优化算法、前向安全与后向安全,以及气象数据价值方程。

2.1 动态可搜索优化

动态可搜索优化通常包含3个算法:初始化算法、更新算法和密文搜索算法。

(1)初始化算法。输入安全参数 λ 和气象数据开发价值 C ,该算法输出 (K, σ, EDB, V_q) 。其中, K 表示用户的对称密钥, σ 表示用户当前的状态, EDB 表示云服务器存储的优化气象数据, V_q 表示气象数据调整价值。

(2)更新算法。该算法支持优化数据库中的更新操作。用户输入对称密钥 K 和更新操作 op ,如增加或者删除,包含文件标识符 ind 和关键字 w 。该算法输出更新状态 σ 和更新后的优化数据库 EDB 。

(3)密文搜索算法。输入用户的查询请求、更新状态以及调整价值 (q, σ, V_q) ,云服务器收到用户的查询请求 q 和气象数据调整价值。在优化数据库中,执行密文搜索,输出符合气象数据调整价值的查询结果。用户利用对称密钥解密气象数据,获得明文气象数据 $DB(q)$ 。

在动态可搜索优化中,通常定义一个泄露函数 $\mathcal{L} = (\mathcal{L}^{Stp}, \mathcal{L}^{Upd}, \mathcal{L}^{Srch})$ 。其中, $\mathcal{L}^{Stp}$ 、 $\mathcal{L}^{Upd}$ 、 $\mathcal{L}^{Srch}$ 分别表示在初始化、更新、搜索3个阶段中泄露的信息。

2.2 前向安全与后向安全

前向安全是确保数据更新后,不会泄露之前搜索过的关键字等隐私信息。DSSE中的前向安全通常存在一个泄露函数 $\mathcal{L}^{Upd}(ind, W) = \mathcal{L}'(op, ind, |W|)$,其中, $\mathcal{L}'$ 是更新后的泄露函数, op 是增加或者删除操作, ind 表示文件标识符, W 表示文件中的关键字集合, $|W|$ 表示关键字集合中的数量。

后向安全是确保云服务器无法从删除的数据中获取隐私信息。尤其是云服务器不能从已删除的查询陷门中匹配到之后的查询结果。DSSE中的泄露函数 $\mathcal{L}$ 会记录一个序列 Q ,用于表示每个查询请求所对应的不同条目。对于查询请求 (i, w) 和更新 (i, op, in) ,其中, in 表示更新的输入, i 表示查询请求当前的时间戳。

时间戳函数 $TimeDB(w)$ 的定义包含对应于每个文件的注册表关键字 w ,且每个文件均包含该关键字 w 。在数据库中的时间戳函数表达式如式(1)所示:

$$TimeDB(w) = \{(i, ind) | (i, add, (w, ind)) \in Q \text{ 且 } (i', del, (w, ind)) \notin Q\} \quad (1)$$

此外,更新函数 $Updates(w)$ 的定义表示基于关键字的时间戳下的更新动作,如式(2)所示:

$$Updates(w) = \{i | (i, add, (w, ind)) \text{ or } (i, del, (w, ind)) \in Q\} \quad (2)$$

删除函数 $DelHist(w)$ 的定义表示关于关键字 w 的所有删除记录,关联与之对应时间戳下的增加操作,如式(3)所示:

$$DelHist(w) = \{(i^{add}, i^{del}) | \exists ind, s.t., (i^{add}, add, (w, ind)) \in Q \text{ 且 } (i^{del}, del, (w, ind)) \in Q\} \quad (3)$$

2.3 气象数据价值方程

在气象数据资产市场体系中,流通是核心环节。大范围、可持续、高效的数据资产流通,依赖于数据资产市场交易机制的建立。在市场交易机制中,构建公平高效且安全的气象数据资产估值定价体系尤为重要。评估气象数据的价值方法通常包含3种:成本途径、收益途径和市场途径。当未来市场足够活跃并具备足够多的买方场景

效用评价及收益预期后,可以结合收益途径及市场途径交叉验证成本途径估值的合理性。本文采用成本途径下的多因子成本修正模型,以气象数据的开发成本为基础,根据气象数据的特征,引入多重价值隐私对其进行修正。气象数据在成本途径下的多因子修正模型公式如式(4)所示:

$$V_q = C \times f_1 \times f_2 \quad (4)$$

其中, C 表示气象数据产品的开发价值, f_1 表示价值贡献因子, f_2 表示多场景增速因子。在气象数据产品开发过程中,需综合考虑数据产品在整个生命周期内的开发成本,并结合气象数据复用因子、质量因子和时效因子,对开发成本进行调整。气象数据产品的开发价值公式如式(5)所示:

$$C = \sum_{i=1}^n S_i \times \gamma \times p \times q \quad (5)$$

其中, S_i 表示第 i 项气象数据集重置成本, γ 表示气象数据复用因子, p 表示气象数据质量因子, q 表示气象数据时效因子。

3 系统模型与安全模型

3.1 设计目标

结合气象数据动态更新频繁、查询需求精准、传输与计算资源受限等实际应用场景,本文所设计的可搜索加密方案需在安全能力、查询功能与运行效率三方面达成平衡,为此确立以下核心设计目标:

(1)前向安全与后向安全。本文提出的RCFB方法,面向气象数据动态更新环境,能够有效实现前向安全与后向安全的可搜索加密优化,在保障数据动态更新的同时,抵御历史数据泄露与新增数据关联推断等安全威胁,充分满足气象数据在高时效性场景下的安全存储与安全检索需求。

(2)联合关键字查询。针对气象数据多维度、多属性的检索特点,本文提出的RCFB方法支持多关键字组合的联合查询,能够在密文域内完成灵活、精准的匹配检索,避免单一关键字检索带来的结果冗余与定位低效问题。

(3)高效性。本文提出的RCFB方法在数据动态更新与关键字搜索阶段均具备更优的计算效率,在保证安全强度的前提下,有效降低客户端与服务器端的计算开销,同时减少网络交互带来的通信开销,适配气象数据高频访问与实时处理的应用需求。

3.2 系统模型

系统模型主要包含用户、云服务器与区块链3个核心实体。其中,用户负责气象数据的预处理、加密、资产定价及更新检索请求;云服务器提供密文存储与密文检索服务;区块链作为可信存储媒介记录数据的元数据,以确保链下数据与链上记录的一致性。图1展示了本文提出的面向气象数据的动态可搜索加密优化系统模型。

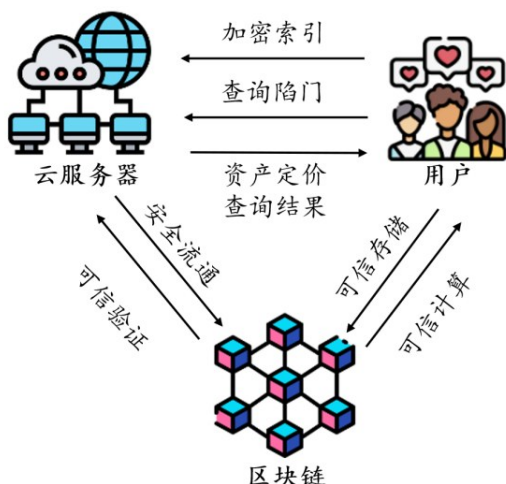


Fig. 1 System model

图1 系统模型

(1)用户。用户收集了大量的气象数据,用于天气预测和灾害预警。对于高价值密度的气象数据,用户根据气象数据价值方程,对气象数据进行预处理和计算,保证在气象数据隐私保护前提下,满足气象数据资产估值定价的需求。为便于存储和管理,用户将大规模气象数据优化后,外包至云服务器。由于气象数据的实时性和时效性,用户通常会对气象数据进行增加和删除操作,以及对气象数据进行估值。因此,用户会向云服务器发送更新、检索和资产定价等请求。

(2)云服务器。云服务器具备强大的计算资源和丰富的存储空间,负责存储优化的气象数据。此外,云服务器为用户提供密文检索和动态更新服务,并对优化的气象数据进行资产定价。

(3)区块链。区块链作为可信存储的媒介,每一个用户所售气象数据都将在上链前进行气象数据分片,并让每个气象数据分片经TEE环境优化存储在本地,在区块链上记录元数据(优化哈希、存储时间戳、复制情况、位置、数据签名等)。

3.3 安全模型

在安全模型方面,本文将云服务器设定为半可信实体,即其会诚实执行协议但可能尝试推断隐私信息,并采用真实世界与理想世界模拟的形式化安全定义。通过泄露函数刻画初始化、更新和搜索3个阶段可能泄露的信息,并将前向安全(确保更新不泄露历史搜索信息)与后向安全(确保删除数据无法被后续查询恢复)作为核心安全目标。

(1)真实世界 $Real^{Adv}(\lambda, q)$ 。在真实世界中,攻击者可以随机访问DSSE的3个算法。系统初始化之后,攻击者自适应地提交更新请求和搜索请求,并通过其返回的结果,计算并输出单个字符 b' 。

(2)理想世界 $Ideal^{Adv}(\lambda, q)$ 。在理想世界中,攻击者可以和模拟器进行交互,提交与真实世界相同的更新请求和

搜索请求。模拟器利用泄露函数 $\mathcal{L} = (\mathcal{L}^{Stp}, \mathcal{L}^{Upd}, \mathcal{L}^{Srch})$ 重复执行更新算法和搜索算法。同样地,攻击者计算并输出单个字符 b' 。

4 面向气象数据的高效安全流通与共享机制

4.1 “存—算—通”一体的数据要素可信安全流通体系架构模型

在数据要素可信安全流通体系架构层面,本文设计了“存—算—通”一体的可信安全流通模型,其技术架构如图2所示。其中,“存”依托区块链与可信执行环境实现数据分片存储与元数据上链,并通过周期性证明确保数据完整性与真实性;“算”基于“数据不动程序动”的思想,将计算程序迁移至可信执行环境或数据沙箱中对全量数据进行处理,计算结果经签名后由智能合约验证;“通”通过链上合约对可信执行环境节点进行统一调度与验证,确保链下计算结果与链上所见结果一致,从而支撑气象数据的可信存储、可信计算与可信流通。

由图可知,“存”关键在于确保链上表述气象数据与链下存储气象数据的一致性。本文借助区块链作为可信存储机制的基础。每一个供应商所售气象数据都将在上链前进行气象数据分片,并让每个气象数据分片经TEE环境优化存储在本地,同时,在区块链上记录元数据(优化哈希、存储时间戳、复制情况、位置、数据签名等)。此外,每个复制副本都会受到链上合约周期性发起的证明挑战和复制证明挑战,以确保链下气象数据完整性和真实性,从而保证用户访问链上获取相应的链下气象数据,实现所见即所得。

“算”关键在于确保气象数据安全前提下,最大限度挖掘气象数据价值,构造轻量级运行环境和算法。该轻量级运行环境和算法既可以基于TEE硬件隔离,也可以由数据沙箱等软件隔离构成。当气象数据交易完成,气象数据提供方将需求数据安全托管至该运行环境。气象数据需求方利用样本数据调试数据分析程序,当程序顺利运行后,基于安全程序迁移引擎,将程序单向摆渡至轻量级运行环境中的全量真实气象数据上计算。气象数据需求方获得的分析结果会经过数据衍生品智能审查,以确保分析结果的隐私保护。此外,输出结果在受控运行环境中生成签名。该签名能够被智能合约验证。智能合约的不可篡改性能够向外界证实该链下计算结果是在TEE被完整执行,未经修改。通过“数据不动程序动”,既克服了气象数据流通的困境,又实现了气象数据的高效利用。

“通”关键在于确保链下计算结果与用户在链上所见结果的一致性保障。本文基于“存”中TEE可信执行环境与链上智能合约的协同工作机制。具体而言,每一个加入网络的TEE都必须在链上合约中注册,也就是由链上合约管理调度和验证。用户提供所需的算法和模型,进行气象

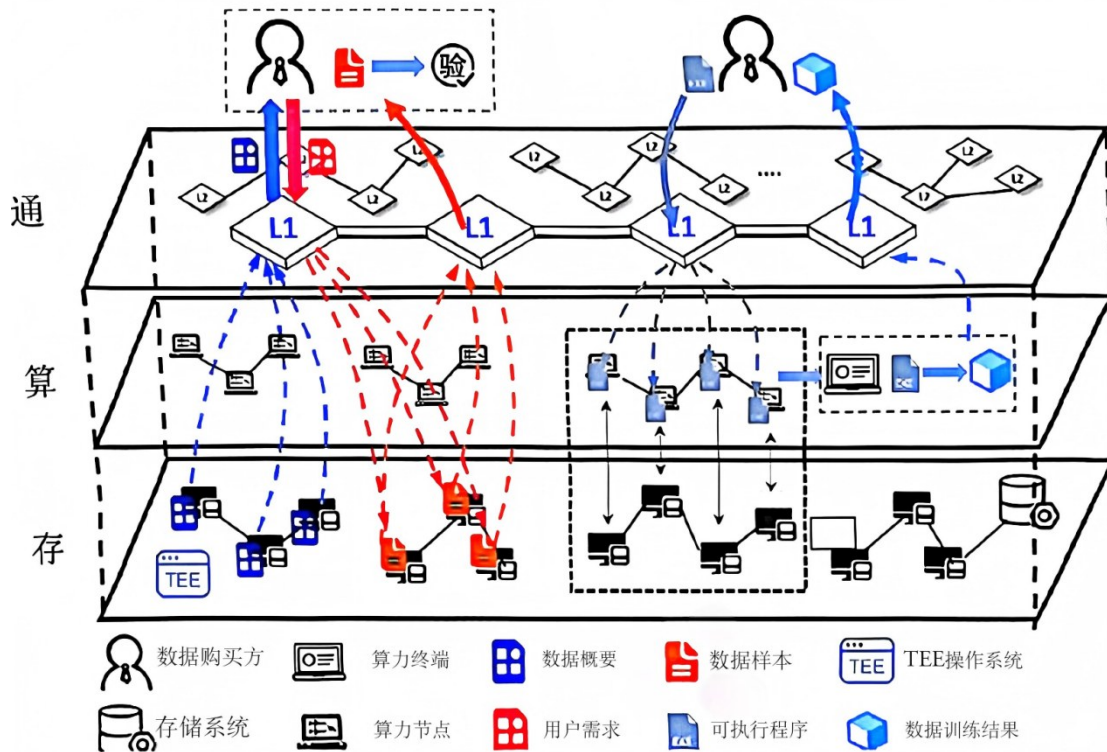


Fig. 2 A model for secure circulation of data in an integrated "storage-compute-network" system

图2 “存—算—通”一体的数据要素可信安全流通体系架构模型

数据资源发现,链上合约将其调度给合约节点并加载到节点的TEE程序中。算法和模型在TEE运行完成后,输出结果将在TEE环境中生成签名,并被智能合约验证。智能合约的不可篡改性可以向用户证实链下计算结果是在TEE中被完整执行,没有被修改。

“存—算—通”一体的数据要素可信安全流通体系架构模型基于区块链和TEE,支持气象数据要素在系统中的可信存储、计算和流通。总体流程概括如下:首先,气象数据供应商对敏感的气象数据进行优化分片,存储于本地,并在区块链上记录分片元数据,实现可信存储;其次,气象数据需求方提供计算程序,链上合约将程序迁移至受控运行区域,对全量气象数据进行可信计算,同时,链上合约对返回的结果进行可信验证;最后,通过构建“存—算—通”一体气象数据要素流通与交易传递技术,支持大规模、分布式和可扩展的气象数据要素流通。

4.2 RCFB算法设计

在算法设计上,本文提出一种支持联合关键字搜索与前后向安全的动态可搜索加密方法。其核心在于将每个文件所包含的所有关键字信息聚合为一个条目进行存储,从而减少联合查询所需的陷门数量;采用链状结构组织每个关键字的更新记录,使得搜索时仅需遍历最新更新链而无需扫描全部数据;算法利用伪随机函数与哈希替代复杂的群运算,在降低计算与通信开销的前提下实现前向安全与后向安全。本文提出的RCFB主要包含3个算法:初始化算法(RCFB.Setup)、更新算法(RCFB.Update)和搜索算法(RCFB.Search)。

(1)初始化算法(RCFB.Setup)。输入安全参数 λ 和气象数据开发价值 C ,输出密钥、优化数据、初始化用户状态、优化气象数据,以及气象数据调整价值,即 (K, σ, EDB, V_q) 。初始化阶段,系统对非结构化气象数据进行预处理和资产估值定价,将非结构化气象数据转化为结构化数据,便于在云服务器进行优化气象数据的更新和搜索,以及资产估值定价操作。在多因子成本修正模型下,系统根据气象数据的开发价值 C ,通过气象数据价值方程计算出调整价值 V_q 。此外,系统生成两个伪随机函数 F_1, F_2 , 3个哈希函数 H_1, H_2, H_3 。其中,伪随机函数 F_1 根据关键字 w 和对称密钥 K ,生成关键字密钥 K_w 。伪随机函数 F_2 根据关键字 w 和对称密钥 K ,生成陷门 T_w 。哈希函数 H_1 用于生成文件标识符和关键字对的地址,哈希函数 H_2 用于优化文件标识符,哈希函数 H_3 用于生成优化关键字的对称密钥。 EDB 表示优化的文件标识符和关键字对,存储在云服务器端。 $LastKey$ 和 $UpdateCnt$ 表示用户的状态信息,其中, $LastKey$ 表示存储的对称密钥信息,对应最近更新的文件标识符和关键字对, $UpdateCnt$ 表示每个关键字的更新次数。初始化算法如算法1所示。

算法1 初始化算法 RCFB.Setup(λ)

输入:安全参数 λ 和气象数据开发价值 C

输出: $F_1, F_2, H_1, H_2, H_3, EDB, UpdateCnt, LastKey, \sigma$

1. 初始化对称密钥: $K \leftarrow \text{KeyGen}(1\lambda)$
2. 初始化伪随机函数和开发价值: F_1, F_2, C
3. 初始化哈希函数: H_1, H_2, H_3
4. 初始化对称优化算法:SE

5. 初始化优化数据集、更新次数、更新密钥: $EDB, UpdateCnt, LastKey$

6. 初始化用户状态: $\sigma \leftarrow (UpdateCnt, LastKey)$

用户:

7. 将优化气象数据集 $E(DB)$ 和气象数据调整价值 V_q 发送给云服务器

(2) 更新算法(RCFB.Update)。在更新算法中,当用户需要在优化数据集中插入一个新条目时,首先需确定待插入的关键词是否为新关键词。若该关键词是新关键词,则关键词更新计数器增加1。用户利用伪随机函数 F_1 和 F_2 为该关键词生成关键词密钥 K_w 和陷门 T_w 。此外,用户为更新后的文件标识符和关键词对随机生成新的优化密钥。随后,用户利用关键词密钥 K_w 和陷门 T_w 进行异或操作,并采用哈希函数 H_2 映射得到最终值 val 。同时,用户利用当前状态和陷门,进行哈希函数 H_2 映射,得到关键词条目的地址 $addr$ 。每个关键词条目的值被划分为两个部分: val 和 $flags$ 。其中, val 包含文件标识符、操作类型和当前优化密钥; $flags$ 包含文件中其他关键词的信息。对所有关键词及更新后的关键词进行哈希运算,可确保即使条目中包含相同关键词,其哈希值也不同。更新算法如算法2所示。

算法2 更新算法 RCFB.UpdateK, σ , op , w , W' , ind , $EDB \rightarrow EDB'$

输入: $K, \sigma, op, w, W', ind, EDB$

输出: 更新后的优化数据集 EDB'

用户:

1. if $UpdateCnt = NULL$ then

2. $UpdateCnt \leftarrow 0$

3. end if

4. $UpdateCnt \leftarrow UpdateCnt + 1$

5. $K_w \leftarrow F_1(K, w)$

6. $T_w \leftarrow F_2(K_w, w)$

7. $i \leftarrow UpdateCnt[w]$

8. $sti-1 \leftarrow LastKey[w]$

9. $sti \leftarrow 0, 1 \lambda$

10. $LastKey[w] \leftarrow sti$

11. $addr \leftarrow H_1(Twist_i)$

12. $val \leftarrow H_2(sti \parallel Tw) \quad \{[(ind \parallel op) \quad Kw] \parallel sti-1\}$

13. $tag \leftarrow \emptyset$

14. for $i=1$ to W' do

15. $tag \leftarrow tag \cup H_3(W'[i] \parallel w)$

16. end for

17. $con \leftarrow SE.Enc(sti, flag)$

18. 用户发送 $addr, val, con, op$ 至云服务器

云服务器:

19. 输出: $EDB' \leftarrow (val, con, op)$

(3) 搜索算法(RCFB.Search)。在搜索阶段,用户提交查询请求,包含用户状态和气象数据调整价值 (q, σ, V_q) 。云服务器收到用户的查询请求 q 和气象数据调整价值 V_q 后,在优化气象数据库中,执行密文搜索,输出符合气象数据调整价

值的查询结果。具体而言,用户首先基于关键词 w_1 、气象数据调整价值 V_q 、用户计算关键词密钥 K_{w_1} ,以及基于关键词和调整价值的陷门 T_{w_1} ,生成查询陷门 $TList$ 。其中,查询陷门 $TList$ 便于关键词 w_1 的索引匹配以及气象资产定价的计算。其次,用户将 $(K_{w_1}, T_{w_1}, UpdateCnt[w_1], LastKey[w_1], TList)$ 发送至云服务器。收到用户发送的查询请求后,云服务器从优化数据集 EDB 中检索并计算出与关键词 w_1 和 $(val, flag)$ 最相关的查询结果。云服务器利用对称密钥从优化的文件标识符和 $(val, flag)$ 中解密并判断 $TList$ 是否在 $flag$ 中。若 $TList$ 在 $flag$ 中,则该文件满足联合查询条件,云服务器将优化的文件标识符加入查询结果,并返回给用户。然后,云服务器重复上述计算步骤,直到对应于 w_1 的所有条目都检索完成。最后,云服务器将优化的查询结果集合返回给用户。该查询结果满足查询请求以及气象数据资产定价。用户利用密钥 K_{w_1} 解密查询结果,并基于更新操作增加或删除结果集合中的文件标识符。搜索算法如算法3所示。

算法3 搜索算法 RCFB.SearchK, σ , q , V_q , op , $EDB \rightarrow Rq$

输入: $K, \sigma, q, V_q, op, EDB$

输出: 查询结果 Rq

用户:

1. w_1 是联合查询请求中最近更新的关键词 $q = w_1 \wedge w_2 \wedge \dots \wedge w_n$

2. $K_{w1} \leftarrow F_1(K, w_1)$

3. $T_{w1} \leftarrow F_2(K_{w1}, w_1)$

4. for $i=1$ to $q.size()$ do

5. $TList \leftarrow TList \cup H_3(wi \parallel V_q)$

6. end for

7. $NewsK \leftarrow LastKey[w_1]$

8. $Sent \leftarrow UpdateCnt[w_1]$

9. 将 $TList, Tw1, NewsK, Scent$ 发送至云服务器

服务器:

10. $Fq \leftarrow \emptyset$

11. $i \leftarrow Sent$

12. $sk \leftarrow NewsK$

13. while $i > 0$ do

14. $addr = H_1(Tw1 \parallel i \parallel sk)$

15. $val, con \leftarrow EDB[addr]$

16. $\{[(ind \parallel op) \quad Kw] \parallel sti-1\}, Encid, sklast \leftarrow H_2(ski \parallel Tw1) \oplus val$

17. $flag \leftarrow SE.Dec(sk, con)$

18. if $TList \subseteq flag$ then

19. $Fq \leftarrow Fq \cup Encid$

20. end if

21. $sk = sklast$

22. $i = i - 1$

23. end while

24. 将 Fq 发送给用户

用户:

25. for $i=1$ to $Fq.size$ do

26. $ind, op \leftarrow Kw1 \quad Fq[i]$

27. if $op = Add$ then

```

28. Rq ← Rq ∪ {ind}
29. else if op = Del & Del = Rq then
30. 从 Rq 中去除 ind
31. end if
32. end for
33. 输出 Rq

```

4.3 安全性分析

在更新阶段,用户利用哈希函数和伪随机函数生成(addr, val, flag),并将其发送至云服务器。由于flag的大小取决于关键字的数量,因此flag会泄露文件中关键字的数量。在搜索阶段,云服务器接收到查询请求q后,会遍历最近更新 w_1 的所有条目,但这一过程可能会泄露 w_1 条目的更新时间。基于此,本文给出RCFB安全性证明,如定理1所示。

定理 1 (RCFB 安全性) 假设哈希函数是随机生成的,且伪随机函数是安全的,则RCFB关于泄露函数 $\mathcal{L}_{RCFB}$ 是语义安全的。

本文采用真实世界和理想世界的设定对攻击者和挑战者进行模拟游戏,证明攻击者无法区分真实世界的输出和模拟理想世界的输出。

Game 0 是一款在真实世界运行的安全游戏,如式(6)所示:

$$\Pr[Real^{Adv}(\lambda, q) = 1] = \Pr[G_0 = 1] \quad (6)$$

在Game 1中,伪随机函数 F_1 和 F_2 的输出是随机数,用户初始化映射 M_1 ,用于存储关键字 w 对应的密钥和陷门。当执行搜索时,用户从 M_1 中获取关键字 w 的密钥和陷门。显然地,攻击者无法从随机字符串(由随机函数生成)区分伪随机函数,如式(7)所示:

$$|\Pr[G_1 = 1] - \Pr[G_0 = 1]| \leq Adv_{F_1}^{PRF}(\lambda) + Adv_{F_2}^{PRF}(\lambda) \quad (7)$$

在Game 2中,哈希函数 H_3 的输出值是随机数,用户初始化映射 M_2 ,用于存储关键字 w 和其他关键字的查询陷门。在搜索阶段,用户从 M_2 中获得查询所需的陷门。显然地,攻击者无法通过随机字符串(由随机函数生成)区分伪随机函数,如式(8)所示:

$$|\Pr[G_2 = 1] - \Pr[G_1 = 1]| \leq Adv_{H_3}^{Hash}(\lambda) \quad (8)$$

在模拟阶段,用户为每个时间戳下的更新、关键字 w 、 $UpdateCnt[w]$ 、 add ,初始化映射 P 。此外,当检索联合查询请求 $q = w_1 \wedge w_2 \wedge \dots \wedge w_n$ 时,用户使用映射 P 获取相关信息(发送至云服务器)。该模拟操作可以保证模拟器和Game 2拥有相同的分布,如式(9)所示:

$$\Pr[Ideal^{Adv, S, \mathcal{L}}(\lambda, q) = 1] - \Pr[G_2 = 1] = 0 \quad (9)$$

根据以上模拟游戏,攻击者无法从真实世界中区分模拟器输出值,证明了提出的RCFB是语义安全的,如式(10)所示:

$$\Pr[Real^{Adv}(\lambda, q) = 1] - \Pr[Ideal^{Adv, Sim, \mathcal{L}}(\lambda, q) = 1] \leq Adv_{F_1}^{PRF}(\lambda) + Adv_{F_2}^{PRF}(\lambda) + Adv_{H_3}^{Hash}(\lambda) \quad (10)$$

5 性能评估

5.1 复杂度分析

(1)存储开销。存储开销主要分为服务器端和用户端两部分。其中,服务器端需要存储优化后的数据集EDB。在 N 次更新之后,服务器端的存储复杂度为 $O(N)$ 。在用户端,用户需要存储两个映射函数的输出值,即 $UpdateCnt$ 和 $LastKey$ 。两者在初始化阶段都置为0,增加操作 $AddCnt$ 的存储复杂度为 $O(|W| \cdot \log |D|)$ 。此外,更新密钥($LastKey$)的存储复杂度为 $O(|W|)$ 。其中, $|W|$ 表示增加关键字后所有关键字的数量, $|D|$ 表示更新后所有文件的数量。

(2)计算开销。在搜索开销方面,用户需要计算关键字 w_1 的查询陷门和密钥,以及其他文件中关键字的查询陷门。陷门生成过程的计算复杂度为 $O(m)$,其中, m 表示查询请求中关键字的数量。在密文检索阶段,与 w_1 相关的所有条目索引都会被检索,检索过程的计算开销为 $O(UpdateCnt[w_1])$ 。最终,用户获得查询结果,并利用密钥解密优化文件索引,解密过程的计算复杂度为 $O(UpdateCnt[w_1])$ 。此外,用户和云服务器的通信开销取决于查询陷门和优化文件索引的大小。因此,用户端的通信开销为 $O(m)$,云服务器端的通信开销为 $O(UpdateCnt[w_1])$ 。

5.2 实验评估

本文实现了RCFB算法,并将其与同样支持前向安全与后向安全的最先进DSSE方法(ODXT)进行了对比实验。本文实验环境和配置如下:64-bit Windows 10操作系统,16 GB内存,Intel(R) Core(TM) i7-12700F CPU。本文所有实验均采用Python 3.9运行实现。本文实验数据集采用南京市气象局提取的300 000文件,每个文件包含10~20个气象关键字,整个文件集中包含200个气象关键字。气象关键字包括气温、相对湿度、气压、风速、降水量、卫星、地图、观测数据、台风、海洋高空气象、地面气象、天气雷达、风云气象、天气预报和天气实况等。本文主要评估更新时间开销、搜索时间开销以及通信开销。

(1)更新时间开销。本文主要测试了气象文件集规模为100 000和300 000两种情况下的更新时间开销,并和ODXT进行了对比。文件集将根据删除比例(0%~50%)进行随机删除。更新时间开销与删除比例关系,如图3所示。实验结果表明,本文提出的RCFB方法更新效率远高于对比方法ODXT。更新时间随气象文件集规模和删除比例的增加而增加,这是由于更新时间与文件-关键字对的数量呈线性增长关系。

(2)搜索时间开销。本文测试了气象文件集规模为100 000和300 000两种情况下的搜索时间开销,同时,本文的删除比例变化范围为10%~50%。搜索时间开销与删除比例关系,如图4所示。在不同的删除比例情况下,

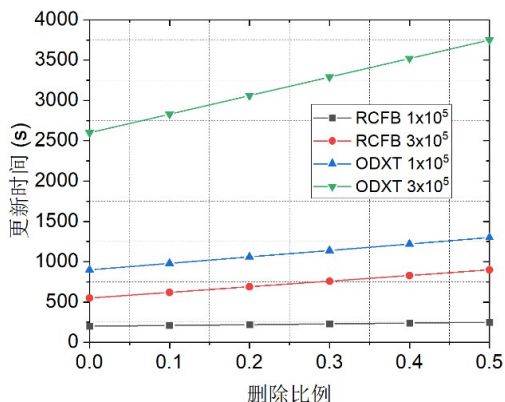


Fig. 3 Relationship between update time overhead and deletion ratio

图3 更新时间开销与删除比例关系

RCFB的搜索时间基本保持不变,维持在较低的时间开销。即使删除比例增加到50%,RCFB的搜索时间开销的增幅也较小。这意味着RCFB在更新请求非常频繁的情况下,也能够保持高效的搜索。其原因是RCFB只需搜索最新更新的关键字,而无需遍历整个文件集。通过上述方式,RCFB可以减少检索气象数据集的规模,从而提高密文检索的效率。

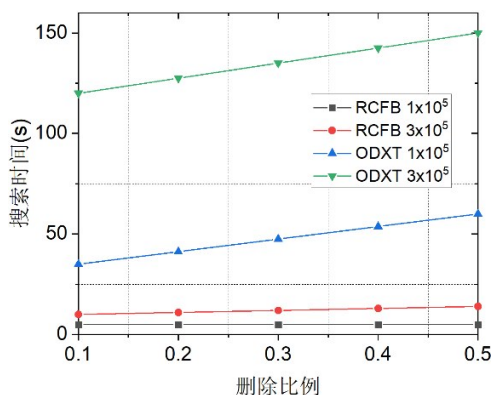


Fig. 4 Relationship between search time overhead and deletion ratio

图4 搜索时间开销与删除比例关系

此外,本文还测试了关键字数量对检索效率的影响,如图5所示。本文随机选择了10个关键字,并计算了10次搜索时间的平均值。结果表明,搜索时间随着关键字数量的增加而增长。其中,RCFB随着关键字数量的增加,其搜索时间始终保持平稳。这是因为RCFB在搜索阶段,云服务器检索的条目不会随着关键字数量的增加而增大,关键字数量的增加仅会影响云服务器判断文件是否包含该关键字的时间。

(3)通信开销。本文测试了气象文件集规模为100 000和300 000两种情况下用户和云服务器之间的通信开销,并和ODXT进行了对比。同样地,文件集的删除比例为0%~50%。云服务器密文搜索的通信开销,如图6所示。本文随机选择了10个关键字集合,每个关键字集合中包含5个关键字。结果表明,RCFB的通信开销远低于对比方案ODXT,主要原因如下:①RCFB仅需要检索最近更新的关键字相关的条目,而不需要遍历其他关键字集合;②RCFB

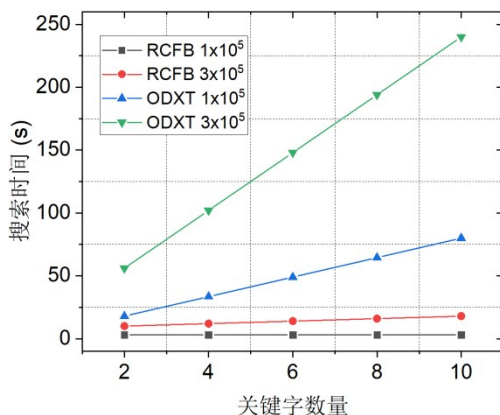


Fig. 5 Relationship between search time overhead and keyword number

ber

图5 搜索时间开销与关键字数量关系

会存储之前检索过的条目,用户仅需向云服务器发送与之前检索请求相关的查询陷门,且这些查询陷门由哈希函数生成,存储开销较小。然而,对比方法ODXT需要为每个条目生成查询陷门,且需要对多个关键字进行逐一检索,其通信开销较大。

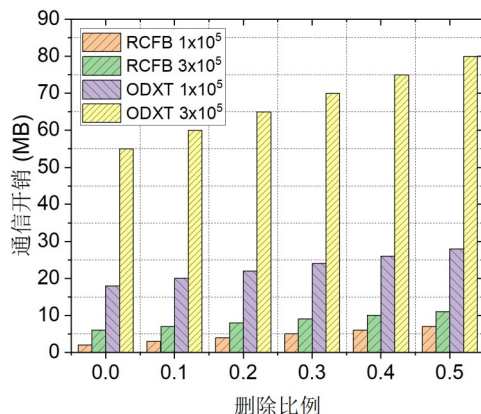


Fig. 6 Communication overhead of search

图6 搜索的通信开销

6 结语

本文提出了一种面向气象数据前后向安全的动态可搜索优化方法(RCFB),支持高效的联合关键字搜索和前后向安全动态更新。为了在高价值密度气象数据动态更新下,实现高效的联合关键字查询以及气象数据的资产定价,本文将所有关键字字典信息存储为一个关键字条目,以提高查询效率。此外,本文设计了基于区块链的可信安全数据要素流通体系架构,提出“存—算—通”一体的气象数据要素可信安全流通体系架构模型,支持气象数据要素在系统中的可信存储、可信计算与可信流通。同时,本文构建了一种基于成本路径的多因子修正模型,用于实现安全的气象数据资产定价,量化气象数据要素价值,以促进数据要素流通。该模型通过综合考虑多种因素,优化数据定价过程,确保了气象数据的安全性、与定价的科学性,为

气象数据安全共享提供了可靠的理论支持。RCFB通过每个文件的相关条目与所有关键字信息保持一致,并采用链状结构存储关键字条目,从而降低了通信与计算开销,同时提高了联合关键字的查询效率。当前,在气象数据领域,关于动态可搜索优化中查询结果的完整性与正确性验证问题仍未解决。下一步,将面向气象数据,研究可验证的动态搜索优化与流通方法,并对其进行深入分析与探索。

参考文献:

- [1] ZHU D, ZHU H, HUANG C, et al. Efficient and accurate cloud-assisted medical pre-diagnosis with privacy preservation [J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(2): 860–875.
- [2] SUN B, WANG H J, HUANG Y Y, et al. Characteristics and causes of the hot-dry climate anomalies in China during summer of 2022[J]. Trans Atmos Sci, 2023, 46(1): 1–8.
孙博, 王会军, 黄艳艳, 等. 2022 年夏季中国高温干旱气候特征及成因探讨[J]. 大气科学学报, 2023, 46(1): 1–8.
- [3] SONG F, LIANG J, ZHANG C, et al. Achieving efficient and privacy-preserving location-based task recommendation in spatial crowdsourcing[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(4): 4006–4023.
- [4] XUE L, NI J, LIU D, et al. Blockchain-based fair and fine-grained data trading with privacy preservation [J]. IEEE Transactions on Computers, 2023, 72(9): 2440–2453.
- [5] LIU D, HUANG C, XUE L, et al. Authenticated and prunable dictionary for blockchain-based VNF management [J]. IEEE Transactions on Wireless Communications, 2022, 21(11): 9312–9324.
- [6] KAMARA S, PAPAMANTHOU C, ROEDER T. Dynamic searchable symmetric encryption[C]//Proceedings of the 2012 ACM conference on Computer and communications security, 2012: 965–976.
- [7] CAO H D, ZHANG P, LIAO M, et al. Quality assessment of the long-term recalibration history dataset of the FY-3 microwave temperature sounder[J]. Trans Atmos SCI, 2024, 47(4): 669–680.
曹皓东, 张鹏, 廖蜜, 等. 风云三号微波温度计长序列再定标历史数据集质量评估[J]. 大气科学学报, 2024, 47(4): 669–680.
- [8] WANG X, MA J, LIU X, et al. Forward/backward and content private DSSE for spatial keyword queries[J]. IEEE Transactions on Dependable and Secure Computing, 2023, 20(4): 3358–3370.
- [9] ZHANG S, RAY S, LU R, et al. Efficient and privacy-preserving spatial keyword similarity query over encrypted data[J]. IEEE Transactions on Dependable and Secure Computing, 2023, 20(5): 3770–3786.
- [10] LIANG J, QIN Z, XIAO S, et al. Efficient and secure decision tree classification for cloud-assisted online diagnosis services[J]. IEEE Transactions on Dependable and Secure Computing, 2021, 18(4): 1632–1644.
- [11] ZHENG Y, ZHU H, LU R, et al. Efficient and privacy-preserving aggregated reverse kNN query over crowd-sensed data [J]. IEEE Transactions on Information Forensics and Security, 2023, 18: 4285–4299.
- [12] DENG H, ZHANG J, QIN Z, et al. Policy-based broadcast access authorization for flexible data sharing in clouds[J]. IEEE Transactions on Dependable and Secure Computing, 2022, 19(5): 3024–3037.
- [13] SHEN Y, GUO B, SHEN Y, et al. A pricing model for big personal data[J]. Tsinghua Science and Technology, 2016, 21: 482–490.
- [14] PATRANABIS S, MUKHOPADHYAY D. Forward and backward private conjunctive searchable symmetric encryption [C]//Proceeding 28th Annual Network and Distributed System Security Symposium, 2021: 1–12.
- [15] GUO Y T, LIU F, WANG F, et al. Optimizing locality for forward secure dynamic symmetric searchable encryption scheme[J]. Journal of Electronics & Information Technology, 2025, 47(8): 2780–2790.
郭宇韬, 柳枫, 王峰, 等. 优化局部性的前向安全动态对称可搜索加密方案[J]. 电子与信息学报, 2025, 47(8): 2780–2790.
- [16] HUANG Y C, YU B. A dynamic symmetric searchable encryption scheme for conjunctive query based on SHVE[J]. Journal of Computer Research and Development, 2024, 61(6): 1545–1558.
黄一才, 郁滨. 一种基于 SHVE 的连接查询动态对称可搜索加密方案[J]. 计算机研究与发展, 2024, 61(6): 1545–1558.

(责任编辑:陈维捷)