

区域性开放背景下文档中台 数据加密和共享方法研究

李强^{1†}, 赵峰¹, 张均成², 彭永进², 白海滨²

(1. 国网信息通信产业集团有限公司, 北京 102211; 2. 福建亿榕信息技术有限公司, 福建 福州 350001)

摘要:常规的文档中台数据加密共享,主要采用编写数据共享密文的方法实现数据的加密共享过程,忽略了区域性数据传输过程中的数据密文分类对共享结果的影响,导致数据加密共享的数据丢失量较大。因此,提出了区域性开放背景下文档中台数据加密共享方法研究。为文档中台的数据信息设置对应的属性密文,将其进行汇总,提取加密数据的加密因子,上传至区域性开放文档中台,设定数据密钥保护更新逻辑,在客户端解密实现文档中台数据的加密共享。实验结果表明:所提方法应用后得出的文档中台数据加密共享结果,表现出的平均数据丢失量仅为0.09 MB,数据加密共享的安全性较高,满足区域性开放背景下文档中台数据加密共享的现实需求。

关键词:数据加密;数据共享;数据中台;文档中台;区域性开放;数据加密共享

中图分类号:TM769

文献标识码:A

Research on Data Encryption and Sharing Methods for Documents in the Background of Regional Openness

LI Qiang^{1†}, ZHAO Feng¹, ZHANG Juncheng², PENG Yongjin², BAI Haibin²

(1. State Grid Information & Telecommunication Co., Ltd., Beijing 102211, China;

2. Fujian Yirong Information Technology Co., Ltd., Fuzhou, Fujian 350001, China)

Abstract: The conventional encryption and sharing of data in documents mainly adopts the method of writing data sharing ciphertext to achieve the encryption and sharing process of data, ignoring the impact of data ciphertext classification on the sharing results during regional data transmission, resulting in a large amount of data loss in data encryption and sharing. Therefore, a study is proposed on the encryption and sharing method of document data in the context of regional openness. Set the corresponding attribute ciphertext for the data information in the document center, summarize it, extract the encryption factor of the encrypted data, upload it to the regional open document center, set the data key protection update logic, and decrypt it on the client to achieve encryption and sharing of the data in the document center. The experimental results show that after the application of the proposed method, the average data loss of the encrypted and shared data in the document is only 0.09 MB, and the security of the encrypted and shared data is high, meeting the practical needs of the encrypted and shared data in the document in the context of regional openness.

Key words: data encryption; data sharing; data center; document center; regional openness; data encryption sharing

随着互联网领域技术的发展与应用,云端技术在日常工作学习生活中得到了广泛的青睐。在云

端技术的应用实践过程中,文档中台成为数据存储的重要平台。文档中台是一款功能强大、易用性高

的文档管理平台,拥有云端存储、文档协作、管理与权限控制等多项功能。

在区域性开放的背景下,任何人均能通过文档中台获得数据。在这样的背景下,如何实现文档中台中数据加密共享,保障数据安全,避免重要数据信息泄露成为人们关注的重点问题。对此,很多研究人员提出了数据加密共享方法。

文献[1]针对传统数据共享中数据存储中心化的问题,对数据存储的路径链添加密文,基于区块链为每条密文进行属性代理划分,形成对称加密的数据存储形式,根据数据原文,添加索引链,将对称的密钥作为索引关键词,通过索引链将属性代理与对称密文进行连接,实现安全共享。同时,设置数据密文进行周期性更新,授权用户不同层级的解密权限,实现数据信息的加密共享。实验结果表明该方法数据共享安全性不足。文献[2]针对数据加密共享中密文搜索效率较低的问题,将传统的数据共享索引结构进行优化处理,转变成为倒排索引结构,以应对文档中台中的敏感数据被攻击,为存储的数据添加一条私有链和一条联盟链,形成双链式的数据存储结构,根据这一形态,编写对应的密文搜索算法,将密文中的限门信息进行传输,在倒排索引结构实现数据信息的加密共享。实验结果表明该方法数据共享效率较低。文献[3]使用了云技术将数据信息进行统一存储,利用云存储过程中产生的属性基对数据信息进行搜索,在编写数据共享的密文后,消除重复性的数据信息,引入密文验证算法较少数据的占用,将属性基于数据密文进行映射,对数据信息进行细粒度的访问控制,并隐藏访问策略,实现可搜索的数据加密共享。实验结果表明该方法适用领域范围较小。

考虑到上述文献所提出的数据加密共享方法无法满足文档中台中数据加密共享的现实需求,本文提出区域性开放背景下文档中台中数据加密共享方法。该方法通过将设置好密文的共享数据上传至区域性开放背景下建立的文档中台,在客户端进行解密,实现数据丢失量较少的文档中台中数据加密共享的过程。

1 文档中台中数据加密共享方法设计

1.1 设置文档中台中数据属性密文

对文档中台的可共享数据进行初始化处理。得到初始数据集,如式(1)所示。

$$E = \{E_1, E_2, \dots, E_m\} \tag{1}$$

式中, E 表示初始化处理后的文档中台中数据, m 表示文档中台中数据的属性序列。

将得到的数据进行双线性配对,并生成配对后的元数据^[4-7]。按照元数据的属性序列对数据参数进行表达,如式(2)所示。

$$E^2 = \frac{M_1}{M_2}(q, w) \tag{2}$$

式中, E^2 表示双线性配对后的数据参数, M_1 和 M_2 分别表示数据对中两个数据的线性阈值, (q, w) 表示线性对。

对数据对的参数进行随机选择,并计算数据密文参数,生成主密文,如式(3)所示。

$$s = \sum_{E=1}^m \delta^{E^2+1} + m\delta \tag{3}$$

式中, s 表示对应数据属性参数的主密文, δ 表示随机参数。

利用得到的主密文对文档中台中数据进行属性加密处理。输入主密文对应的明文作为数据的公钥,完成文档中台中数据属性密文的设置。

1.2 上传数据至区域性开放文档中台

在区域性开放的背景下,文档中台的数据共享权限是数据加密共享的前提^[8]。将加密处理后的数据信息与文档中台的数据存储节点进行对应,实现数据的上传过程。

如图 1 所示,将区域性开放的文档中台与以太网进行连接,加密数据的上传过程均在以太网的线路上进行。客户端产生的加密数据通过互联网区进行汇总,再传输至区域性文档中台前置机中,在这一过程中,由专门的防火墙对数据信息进行保护^[9-11]。

区域性文档中台前置机

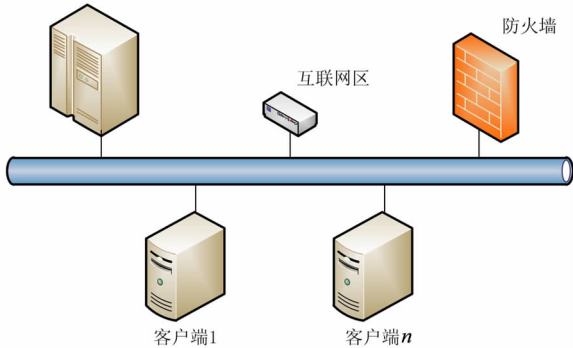


图 1 文档中台中加密数据上传

在互联网区对加密的数据进行汇总,数据汇总的主要内容是对每个加密后的数据信息的加密因子进行提取,根据提取到的加密因子将加密数据按类型划分并排序存储。

加密因子的提取首先要建立专业化访问树,对客户端中的加密数据进行访问。根据随机参数生成的主密文来遍历访问树的根节点,对根节点上的加密因子进行识别,如式(4)所示。

$$d = \frac{s_n K}{e}$$
 (4)

式中, d 表示识别到的加密因子, s_n 表示随机参数对应的主密文, K 表示加密访问树的雾节点, e 表示加密数据类别的分布概率。

通过式(4)得到加密数据的识别结果,将识别到的加密因子发送给访问树的根节点,将对应的加密数据分类存储在区域性文档中台中。

1.3 设定密钥保护更新逻辑

为了保证数据加密共享的长期安全性,需要定期对加密数据对应的密钥保护进行更新,如图 2 所示。

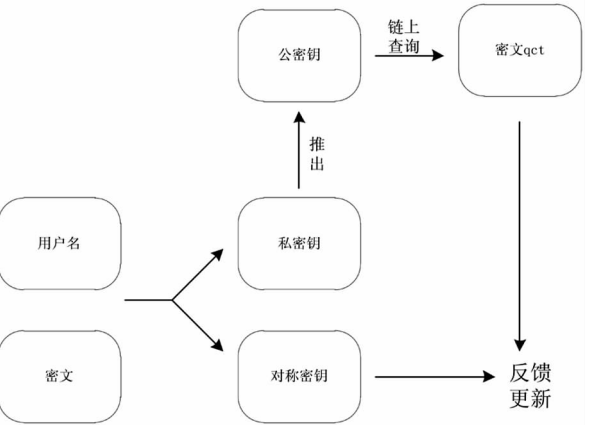


图 2 密钥保护更新

如图 2 所示,数据密钥的保护更新由客户端的用户名及数据密文共同作为认证 Key,由此生成私密钥以及对称密钥,进而推出对应公密钥的值^[12]。同初始数据密文进行比对,将匹配成功的结果作为有效值,生成对应的密文 qct,结合这一结果进行反馈更新^[13,14]。反馈更新的逻辑特性设置如式(5)所示。

$$J = \sum_{i \in s} \lambda t$$
 (5)

式中, J 表示密钥更新逻辑特性, λ 表示数据密钥的秘密值, t 表示数据密钥的定期更新时间。

通过上述设置,完成对文档中台数据密钥保护更新的逻辑设定过程。

1.4 解密共享文档中台加密数据

数据加密共享的实现需要在终端对加密的数据信息进行解密,实现共享读取。首先设置数据共享传输协议,如图 3 所示。

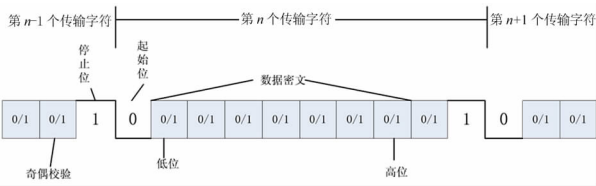


图 3 数据共享传输协议帧结构图

根据数据共享传输协议,将加密数据的密文属性进行撤销处理,即加密数据的解密。数据的解密过程如式(6)所示^[15]。

$$\forall i = (k - k')G$$
 (6)

式中, $\forall i$ 表示数据 i 的密文属性更新因子, k 表示原属性组密钥, k' 表示新属性组密钥, G 表示数据密文属性最小覆盖子集。

通过式(6)的结果,生成文档中台数据的全局公共参数,在客户端进行数据显示,完成数据的解密,最终生成的文档中台数据加密共享结构如图 4 所示。

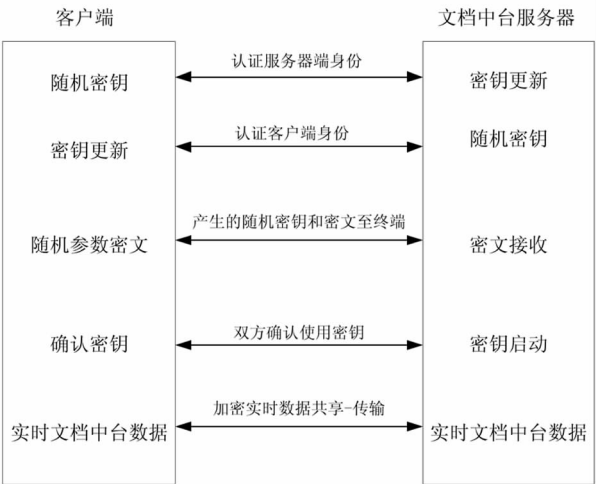


图 4 文档中台数据加密共享架构

按照如图 4 所示的流程架构,在对应的客户端节点进行数据加密共享传输,完成本研究设计的文档中台数据加密共享。

2 实验

2.1 实验准备

对本文所提方法进行验证。本次实验采用数据模拟仿真的方法进行。从实验结果分析本文所提文档中台数据加密共享方法的有效性。

建立实验仿真环境,如图 5 所示。

对如图 5 所示的主控计算机的各项参数进行配置,如表 1 所示。

使用如表 1 所示的配置参数计算机,开展本次模拟仿真实验。

以某单位的文档中台数据为数据模拟的样本集。为保证本次实验与实际应用中的拟合性,选取该单位中不同部门的文档中台数据进行模拟。为不同类型的文档中台数据设置不同的数据长度,如图 6 所示。



图 5 实验仿真环境

表 1 主控计算机配置参数

项目	配置参数
基础操作系统	Windows 10
节点操作系统	Ubuntu 16. 04 LTS
开发语言	OpenJDK 1. 8. 0. 252
主存	16 GB
处理器	AMD Ryzen™ 7 4800H
显卡	RTX 2060 Max-Q
主频	3200 MHz
硬盘	24 TB

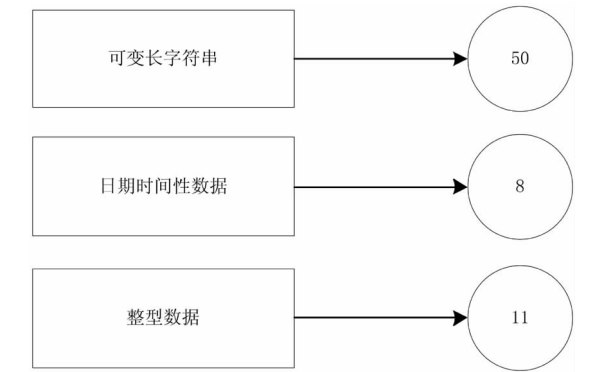


图 6 实验数据类型及长度

按照如图 6 所示的数据类型长度标准模拟出本次实验所需数据信息,并存储到对应的文档中台模拟装置中。模拟所得的数据信息在语义上存在一定歧义,为避免由此带来的实验误差,对模拟所得的数据信息进行消歧处理。基于篇幅有限,仅对部分消歧处理效果进行展现,如图 7 所示。

```
luggage_id:airp2_luggage_id airp2_luggage_idluggage_weight:airp2_luggage_weight airp3_luggage_weightflight_number:airp2_flight_number airp2_luggage.idname:airp2_name airp2_luggage_id airp3_name
```

图 7 数据信息消歧处理

利用消歧处理后的数据信息进行数据加密共享实验。经过模拟,得到文档中台数据 50000 条,将其中的 40000 条作为实验过程中算法训练用数据,其余 10000 条作为实验所用数据。

2.2 文档中台数据加密共享

将本研究所设计的文档中台数据加密共享方法进行实现,可视化处理得到数据加密共享界面如图 8 所示。



图 8 数据加密共享界面

以模拟所得的数据进行实验,设置覆盖实验环境的以太网宽带值为 25 GB。在这样的实验环境下,设备之间的数据共享上传速率为 45 Mbps。有服务器端发送数据共享请求,在数据加密共享的客户端将会得到数据化加密共享结果,界面示意图如图 9 所示。

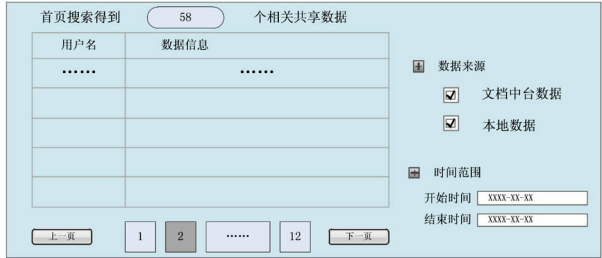


图 9 共享数据获取显示界面

由图 8 和图 9 可知,本文设计的文档中台数据加密共享方法的共享界面较为清晰,可以通过不同方式限定数据共享的范围。在数据加密共享的发送界面中,发送设备与接收设备的服务器信息全面,能够保证数据加密共享的安全性。从这一结果可以初步判断本文所提方法具有一定的可行性,展现了该方法的应用结果,表明了所提方法的应用性。

2.3 结果评价指标

为了更加直观地体现出本研究所提方法的有效性,将本次模拟仿真实验的结果评价指标设置为量化的数据安全测试指标。其计算方法如式(7)所示。

$$\eta = |N_s - N_g| \quad (7)$$

式中, η 表示数据丢失量, N_s 表示数据加密共享上传数据量, N_g 表示接收数据量。

数据丢失量直接反映了数据加密共享过程的安全性,计算所得的结果数值越小,则说明两个端口的数据量差值越小,丢失的数据量更少,表明该方法的数据加密共享安全性越高,在文档中台数据的加密共享过程中,具备更高的实践应用价值。

2.4 结果分析与讨论

为了体现本研究所提方法实验结果的有效性,分别应用文献[1]和文献[2]所提出的加密数据共享方法作为对照组进行文档中台数据加密共享实验。

将模拟所得的数据信息分别应用于不同的数据加密共享方法中,测试不同方法在共享过程中的数据丢失量。为了保证本次模拟仿真实验结果真实性,在每轮实验中为不同数据加密共享方法导入的数据量保持一致。同时,为了减少实验误差,本次实验中,对于每个数据量的共享测试均多次进行,最终结果取多次实验结果的平均值。

经过实验,得到实验结果如图 10 所示。

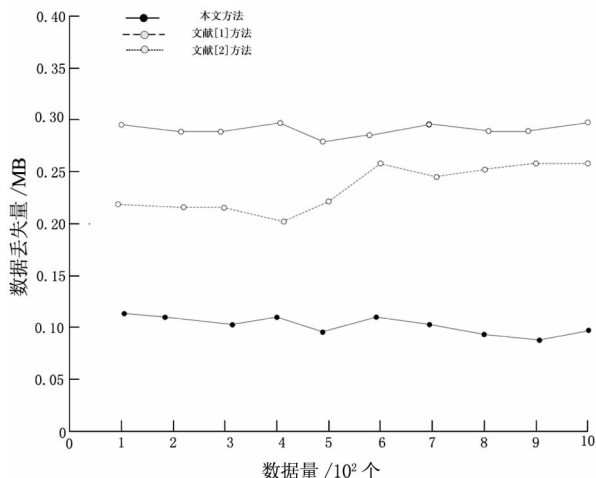


图 10 不同方法数据加密共享丢失量结果

由图 10 可知,本文所提出的文档中台数据加密共享方法,最高的数据丢失量仅为 0.12 MB,而其他两种方法的最高数据丢失量分别达到 0.30 MB 和 0.25 MB。从不同数据量的测试结果来看,本文所提方法的平均数据丢失量为 0.09 MB,而其他两种方法的平均数据丢失量分别达到了 0.28 MB 和 0.24 MB,有着较大的差距。在共享数据量的变化下,本文所提方法的数据丢失量变化幅度不大,文献[1]所提方法的数据丢失量变化幅度也比较小,而文

献[2]所提方法的数据丢失量变化幅度则比较大,与其他两种方法相比超出将近二倍的幅度。综合来看,本文所提方法的数据丢失量较小且数据丢失量的变化幅度较为稳定,说明本文方法能够适应不同数据量下的数据加密共享。

从对比结果可以看出,本文所提出的文档中台数据加密共享方法中,数据丢失量较少,且变化幅度较小,数据加密共享的安全性较高。在文档中台数据的实际应用中,能够实现安全程度较高的数据加密共享过程,具备较高的实践应用价值。

3 结 论

文档中台数据的加密共享是其重要的应用功能之一,为区域性开放背景下的信息交流提供了很大的便利。为了提高文档中台数据加密共享过程的安全性,提出了文档中台数据加密共享方法。从实验结果可以看出,依据所提方法得出的数据加密共享结果的数据丢失量较少,表明本文研究内容能够有效提高文档中台数据加密共享的安全性,为文档中台的数据共享功能提供科学支持,推动文档中台在区域性开放背景下的应用与发展。

参考文献

- [1] 翟社平,童彤,白喜芳. 基于区块链的属性代理重加密数据共享方案[J]. 计算机工程与应用,2023,59(8):270-279.
- [2] 刘炜,白晓丹,余维,等. 基于倒排索引的可搜索加密数据共享方案[J]. 计算机工程与应用,2023,59(10):270-279.
- [3] 牛淑芬,宋蜜,方丽芝,等. 智慧医疗中基于属性加密的云存储数据共享[J]. 电子与信息学报,2022,44(1):107-117.
- [4] 刘雪娇,曹天聪,夏莹杰. 区块链架构下高效的车联网跨域数据安全共享研究[J]. 通信学报,2023,44(3):186-197.
- [5] 薛庆水,孙晨曦,马海峰,等. 基于条件代理重加密的跨链数据共享方案[J]. 计算机应用研究,2023,40(5):1324-1329.
- [6] 薛庆水,时雪磊,王俊华,等. 基于属性加密的个人医疗数据共享方案[J]. 计算机应用研究,2023,40(2):589-594+600.
- [7] 牛淑芬,于斐,陈俐霞,等. 加密电子病历数据共享方案[J]. 计算机工程与科学,2022,44(9):1610-1619.
- [8] 高文静,咸鹤群,程润辉. 基于双层加密和密钥共享的云数据去重方法[J]. 计算机学报,2021,44(11):2203-2215.
- [9] 宋宇翔,杨丽. 基于数据中台的无线传感器网络数据共享系统[J]. 计算机工程与设计,2022,43(11):3022-3029.
- [10] 孙彩萍,王维,何立环,等. 大气环境科学综合数据采集共享平台建设及应用研究[J]. 环境科学研究,2021,34(1):202-212.
- [11] 徐佳丽,罗丽. 基于排序学习的流式文档数据无损隐藏仿真[J]. 计算机仿真,2022,39(3):174-178.
- [12] 先卫红. 大学图书馆“开放教育资源(OER)”建设与服务研究——以加拿大卑诗省大专院校联盟开放教育馆员社区实践为例[J]. 大学图书馆学报,2022,40(4):120-127.
- [13] 黄富才. 区域性开放背景下的医院档案数据共享研究[J]. 档案学研究,2021(5):145-148.
- [14] 聂梓萌,杜劲松,朱建龙,等. 基于仿真区域性数据的服装团体定制归号机制[J]. 纺织学报,2023,44(5):191-197.
- [15] 孟军海,马龙,王金海,等. 重力数据在德令哈地区区域性综合解释中的开发应用研究[J]. 物探与化探,2021,45(2):369-378.