云计算环境下开放式网络数据安全存储方法

王 戈[1] 付成龙[1]
WANG Ge FU Chenglong

摘 要 开放式网络的拓扑结构、节点状态以及传输路径频繁变动，难以维持稳定且安全的存储环境，增加了数据风险。为此，文章设计了一种云计算环境下开放式网络数据安全存储方法。首先，划分数据片段并将其存储至节点中，通过主成分分析完成聚类处理，筛选出存储代价最小的聚类簇用于存储数据，减少数据存储的冗余，从根本上降低数据在存储过程中被攻击和泄漏的风险；然后，引入 AES 算法，根据密钥的长度确定迭代轮次，通过多轮加密以及掩码保护对数据进行加密。当用户请求存储数据时，验证用户签名，若符合条件将数据进行存储，否则对数据进行修正或更换存储节点，以保障数据的安全性。实验表明：应用该方法后，在数据量和节点密度不同时，数据存储有效计算比可以稳定提升至 0.73，数据泄漏率始终未超过 0.2%，说明该方法可以为云计算环境下开放式网络数据提供安全存储方案。

关键词 云计算；数据存储；网络数据；安全存储；AES 算法

doi：10.3969/j.issn.1672-9528.2026.03.020

0 引言

在云计算环境中，用户将数据上传云端后，此时数据处于一个不可控域 [1]，导致用户对云服务提供商的信任不足，同时，在数据传输的过程中，开放式的网络环境极易造成数据丢失或泄漏的情况 [2]，破坏数据的完整性以及保密性。

因此，如何在开放式网络下保障数据的安全，已成为当前阶段的重要研究课题，受到了业内的广泛关注。为此，王卓瑜等 [3] 提出了一种基于逆向运算法的安全存储方法，该算法通过 AES 对数据进行加密处理，构建分层存储模型存储加密数据，并通过逆向运算法在读取时恢复数据，实现数据的安全存储。但该方法依赖单一 AES 算法，若密钥被破解，易存在数据传输中被拦截的风险。赵子晨等 [4] 通过数据抽取、转换和加载流程整合数据，随后采用 HBase 存储归档数据，实现数据的安全存储。但在该方法的数据抽取与转换环节中，数据暂存缺乏加密处理，易遭受伪装节点攻击，影响数据的安全。张利华等 [5] 提出一种基于区块链的安全存储方案，通过边缘计算预处理数据，并基于 K-Raft 构建安全存储链，实现数据的安全存储。但该方法的数据传输未明确加密机制，若存在恶意节点伪造路由信息，可能导致数据存储错误。

基于上述分析，本文提出了一种云计算环境下开放式网络数据安全存储方法，通过对数据进行分片聚类，数据加密等操作，实现数据的安全存储。

1 云计算环境下分块存储开放式网络数据

在云计算环境下，开放式网络数据的存储分布与个人设备的存储方式并不相同 [6]。为保证服务的可靠性，云计算往往不会将数据集中存储 [7]。以开放式网络数据为例，在进行云存储的过程中，首先需将数据划分成片段，对片段进行分组随后储存至节点。为了在符合条件的节点中进行选择，本文在分配某一数据时，选择优先度最高的节点，分配后实时更新节点属性，计算当前数据存储的优越性，用公式表示为：

$$\eta=\sum_{i=0}^{n-1} r_i \times M_i \tag{1}$$

式中：η 表示数据存储优越性；r_i 表示存储位置的优先级；M_i 表示基础数据单元；n 表示数据分片数量。

确定开放式网络数据的切片数量后，采用主成分分析方法对数据进行降维，随后经过标准化与正规化处理，作为计算环境下实值形态空间的基础数据单元，每个数据单元的计算公式为：

$$M_i \in \{M_1, M_2, \cdots, M_N\} \tag{2}$$

式中：N 表示基础数据单元总个数。

针对开放式网络中用户对信息的随机访问，本文将高密度信息节点划分为多个区域，并以单个基础数据单元为单位，计算用户访问距离，用公式表示为：

$$D=\sqrt{\frac{(M_a-M_b)^2}{P}} \tag{3}$$

1. 开封文化艺术职业学院计算机学院 河南开封 475000

式中：D 表示基础数据单元节点间的距离；M_a 与 M_b 表示不同的基础数据单元；P 表示聚类中心点。

筛选距离范围在合理区间内的聚类簇集合，并选取存储代价最小的聚类簇，用公式表示为：

$$\mu_k = \frac{\sum_{i=1}^{m} 1\{D_i = j\} M_i}{\sum_{i=1}^{m} 1\{D_i = j\}} \tag{4}$$

式中：μ_k 表示算术平均值；j 表示聚类簇。

通过上述分块存储，待存储开放式网络数据文件会被拆分为一系列不连续的子文件，由于子文件的顺序与数量由分块粒度决定，窃密者若不掌握具体分块方法，难以通过碎片重组还原原始数据文件。

2 云计算环境下加密开放式网络数据

在云计算环境下开放式网络数据安全存储过程中，本文通过 AES 算法重复执行多轮加密操作，实现数据的加密和解密过程。在 AES 算法中，数据的解密轮数由密钥长度决定 [8]，密钥越长，迭代轮数越多，加密强度越高，具体的对应关系如表 1 所示。

表 1 对应迭代轮数

密钥长度 / 位	迭代轮数 / 轮
128	10
192	12
256	14

密钥长度与轮数的可配置性，可根据开放式网络数据的敏感等级灵活选择，对核心用户数据采用 256 位密钥，对普通公开数据采用 128 位密钥，并通过引入随机数隐藏加密过程中的中间数据，本文以 128 位密钥为例进行计算。

AES 算法将初始密钥扩展为 4 个 32 位子密钥，并通过密钥编排函数扩展为 44 个字的序列，最终生成 10 个与初始密钥长度相同的轮密钥 [9]。将明文与密钥进行混合，前九轮执行字节替换、行移位、列混合和轮密钥加操作，列混合中的状态矩阵变化情况用公式表示为：

$$\begin{pmatrix} y_0' \\ y_1' \\ y_2' \\ y_3' \end{pmatrix} = \begin{pmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix} \tag{5}$$

式中：y 表示待存储的开放式网络数据。

在线性操作部分通过 16 个不同掩码保护 AES 状态，非线性操作部分则对每个状态元素使用统一的输出掩码，第十轮操作时，除了不执行列混合，其余操作相同，AES 算法中按位运算的次数如表 2 所示。

表 2 AES 算法按位操作次数

	XOR	AND	Randomness
密钥加	8	0	0
字节代换	0	0	0
行移位	0	0	0
列混合	56	64	0
密钥扩展	1 144	0	0
标准 AES 算法	10 488	9 216	0

开放式网络数据解密过程与加密过程对称，且使用与加密对应的逆轮密钥，确保解密操作能够正确还原加密操作的结果。

3 开放式网络数据安全存储

在云计算环境下开放式网络数据安全存储中，本文提出基于安全性验证的数据存储与访问机制，通过存储前验证与访问时解密的双重保障，实现数据全生命周期安全管控。当用户向云服务器请求存储数据时，系统首先对用户签名进行安全性验证 [10]，安全性验证的计算公式为：

$$F\left(A, K+H\right) = F\left(Y^{-1}E, K+H\right) \tag{6}$$

式中：F 表示仿射变换函数；A 表示用户的签名；K 表示公开密钥；H 表示哈希编码数据；Y 表示用户私钥；E 表示待存储的数据块。

若等式成立，证明用户身份合法且数据未被篡改，可将开放式网络数据存入对应云服务器；若等式不成立，则判定该云服务器存在信任风险，需执行数据修正、重传或更换存储节点等操作，从源头确保存入云端的数据完整性与合法性。

4 实验验证与分析设置

4.1 实验环境

通过开展如下实验探究云计算环境下开放式网络数据安全存储方法的可行性。

本次实验选择 Hadoop 作为仿真实验环境，仿真平台的操作系统为 Ubuntu-12.04.0，具体配置如表 3 所示。

表 3 计算机配置

节点类型	CPU	内存 /GB
Master	Intel Core i3-2.3 GHz	6
Slave1	Intel Core i5-3.1 GHz	8
Slave2	Intel Core i5-3.1 GHz	8
Slave3	Intel Core i5-3.1 GHz	4

实验选择存储的数据大小分别为 1 TB、5 TB、10 TB、50 TB，数据节点密度分别设为 10 个 /TB、20 个 /TB、40 个 /TB、100 个 /TB，实验时长 24 h，数据周期为 500 s。

4.2 基础性能测试

在云计算环境下开放式网络数据安全存储中，对所设计方法的数据分块重组进行性能测试，针对不同大小的数据，在相同实验环境下进行 10 次重复测试，取其平均值作为最终的实验结果，得到的不同数据量大小下的数据分块与数据合并时间如图 1 所示。

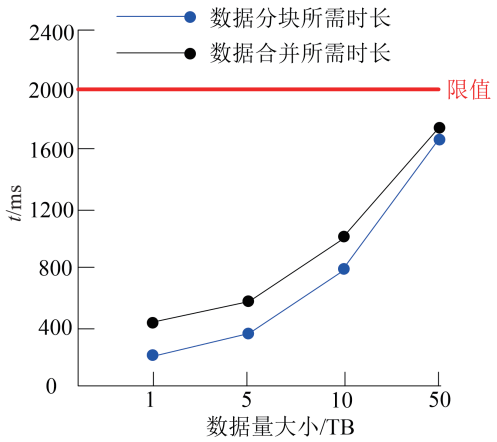


图 1 数据分块与合并所需时长

根据图 1 可以看出，尽管随着数据量的增加，数据分块与合并所需的时间也随之增加，但所需时间均未超过限值。

然后，考虑到本文方法在确定开放式网络数据的切片数量后，采用主成分分析方法对数据进行降维，而降维的效果通常由主成分分析的累积贡献率所决定，因此，统计在数据量大小由 1 TB 增加至 50 TB 的过程中，主成分分析的累积贡献率结果如表 4 所示。

表 4 主成分分析的累积贡献率

数据量 /TB	节点密度 /（个 ·TB^{-1}）			
	10	20	40	100
1	89.5	89.2	88.8	88.3
5	91.9	91.5	91.1	90.7
10	92.4	92.1	91.7	91.3
50	91.2	90.8	90.4	89.9

根据表 4 中的数据可以看出，当节点密度从 10 个 /TB 增至 100 个 /TB 时，相同数据量下的累积贡献率平均下降 0.5%~0.8%，例如当数据量为 1 TB 时，主成分分析的累积贡献率由 89.5% 下降至 88.3%。这是由于节点密度增加导致数据分布更分散，轻微降低了主成分对全局特征的集中表征能力，但降幅控制在 1% 以内，证明方法对节点分布变化具有强适应性。

为了测试开放式数据网络数据的存储性能，对应用本方法前后的数据存储有效比进行测试，数据有效计算比越高，则表明数据的存储性能越好，得到的对比结果如表 5 所示。

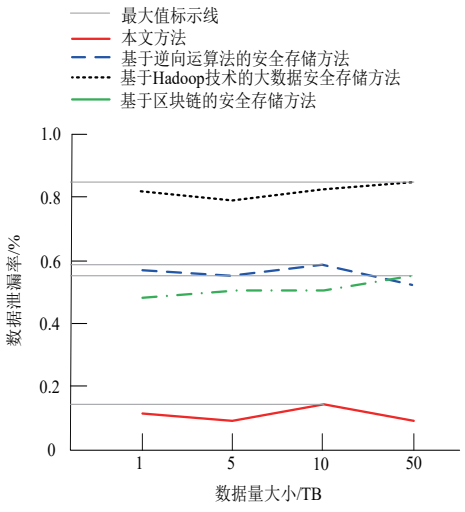
表 5 有效计算比对比结果

数据量 /TB	节点密度 /（个 ·TB^{-1}）	应用前	应用后	提升值
1	10	0.51	0.73	0.22
	20	0.51	0.73	0.22
	40	0.51	0.73	0.22
	100	0.51	0.73	0.22
5	10	0.51	0.73	0.22
	20	0.51	0.73	0.22
	40	0.51	0.73	0.22
	100	0.51	0.73	0.22
10	10	0.51	0.73	0.22
	20	0.51	0.73	0.22
	40	0.51	0.73	0.22
	100	0.51	0.73	0.22
50	10	0.51	0.73	0.22
	20	0.51	0.73	0.22
	40	0.51	0.73	0.22
	100	0.51	0.73	0.22

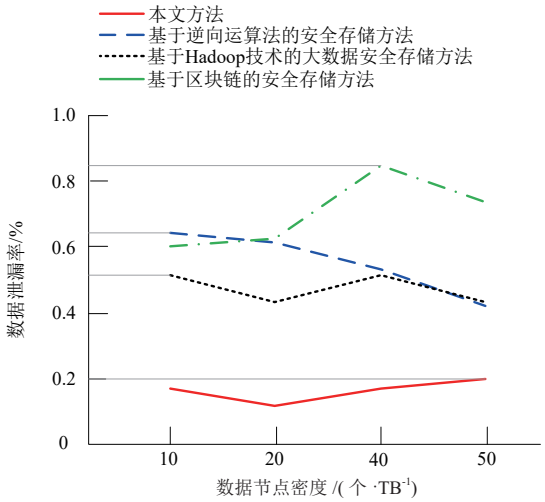
由表 5 可以看出，在所有实验条件组合下，本文方法的数据存储有效计算比均稳定提升至 0.73，较原方法的 0.51 提升 0.22。这表明所设计方法在不同数据规模和节点分布下均能显著优化存储性能。

4.3 开放式网络数据存储安全性对比分析

为更直观地突出本文方法的效果，在 Hadoop 仿真实验环境中对本文设计的存储方法与文献 [3] 中提出的基于逆向运算法的安全存储方法、文献 [4] 中提出的基于 Hadoop 技术的大数据安全存储方法、文献 [5] 中提出的基于区块链的安全存储方法展开对比测试，测试指标为数据泄漏率，结果如图 2 所示。



（a）不同数据量下的安全性对比



（b）不同节点密度下的安全性对比

图 2 不同方法的安全性对比分析

根据图 2（a）可以看出，随着数据量的增加，应用不同方法后，数据泄漏率的变化并不明显。这是由于当数据量增加时，开放式网络的分布式储存特性可平衡不同节点的负载压力，避免因节点拥挤导致的安全漏洞。进一步分析可以看出，本文所设计的数据存储方法的安全性最高，数据泄漏率的变化曲线始终处于三种对比方法曲线之下。由图 2（b）可以看出，随着数据节点密度的增加，数据的存储安全性出现了明显波动，这是由于节点密度增加到一定程度后，攻击者需监控更多节点才能获取完整轨迹。在此情形下，应用本文方法后，数据泄漏率也可以控制在 0.2% 以下，明显低于三种对比方法。综上，本文方法在不同数据大小与数据节点密度下，均表现出了更优的安全性能，为云环境数据安全存储提供了可靠方案。

5 结语

本研究针对云计算环境下开放式网络拓扑多变、数据易泄漏的挑战，提出了一种融合分片聚类与动态密钥加密的安全存储方法。本研究通过数据分片结合主成分分析筛选最小存储代价节点集群，从存储结构源头降低冗余与攻击面。然后引入可配置密钥长度的 AES 算法，依据数据敏感度动态调整迭代轮次并采用掩码技术强化中间状态保护。最后建立用户签名验证与数据修正机制，对非法存储请求实时触发节点更换，实现存储全流程主动防护。

实验验证了该方法在数据量与节点密度变化下，有效计算比稳定提升至 0.73 且泄漏率始终低于 0.2%，显著优于主流方案。未来工作将探索抗量子加密算法集成、区块链辅助的分布式验证架构，并进一步优化边缘计算场景下的动态节点调度策略。

参考文献：

[1] 赵骏，戴欢，唐毅，等 . 基于双账本的物联网数据存储与共享方法 [J]. 计算机工程与设计，2023,44(11):3276-3282.

[2] 林耿豪，周子集，唐鑫，等 . 采用随机块附加策略的云数据安全去重方法 [J]. 西安电子科技大学学报 ,2023,50(5):212-228.

[3] 王卓瑜，王磊，陆婷，等 . 基于逆向运算法的海量大数据安全存储方法 [J]. 兵工自动化 ,2024,43(7):23-26.

[4] 赵子晨，杨锋，郭玉辉，等 . 基于 Hadoop 技术的加速器大数据安全存储与高效分析系统设计 [J]. 现代电子技术，2024, 47(8):9-17.

[5] 张利华，曹宇，张赣哲，等 . 基于区块链的微电网数据安全存储与删除验证方案 [J]. 计算机工程与设计，2023, 44(4): 967-976.

[6] 李晓静，杨秀杰 . 云计算环境下多模态异构网络数据安全存储方法 [J]. 现代电子技术，2025, 48(6): 63-67.

[7] 王颖 . 云计算环境下舰船控制系统大数据有效存储方法 [J]. 舰船科学技术，2023, 45(13):170-173.

[8] 翟禹，武黎明，王祥，等 .AES 加密在嵌入式数据库 SQLite 中的应用 [J]. 火力与指挥控制，2023, 48(6):97-101.

[9] 秦榕霞，苏俊琦，韦萌萌 . 基于区块链技术的分布式数据安全存储方案 [J]. 信息记录材料，2025, 26(7):161-163.

[10] 郑真，黄一楠，黄晨宏，等 . 密钥对交换下电网数据区块链分块安全存储 [J]. 电子设计工程，2025, 33(11):132-135.

【作者简介】

王戈（1981—），女，河南开封人，硕士研究生，副教授，研究方向：计算机网络、计算机网络安全。

付成龙（1989—），男，河南开封人，硕士研究生，讲师，研究方向：计算机网络、计算机网络安全。

（收稿日期：2025-08-11 修回日期：2026-03-16）