基于可搜索加密技术的智慧校园数据安全模型设计

王泽贤 [1]
WANG Zexian

摘 要 随着物联网技术的发展，智慧校园中各种各样的传感器产生的数据与日俱增，云计算技术凭借海量存储和超强的计算能力被广泛应用。现有的智慧校园数据安全方案大多以中心化服务器为核心构建，随着用户数量的增加，网络延迟和用户数据安全问题愈发严重。基于可搜索加密技术和边缘计算技术提出了一种在边缘服务器端完成数据处理的去中心化的数据安全模型，有效提高了数据处理速度和用户数据安全性。

关键词 可搜索加密；边缘计算；去中心化；数据安全

doi：10.3969/j.issn.1672-9528.2024.10.053

0 引言

互联网技术的发展在带来便利的同时也带来很多安全隐患。保障用户的数据隐私安全是互联网应用领域的基本要求。为保障用户数据安全我国于 2021 年 6 月颁布了《中华人民共和国数据安全法》[1]。随着互联网技术的发展智慧校园建设进程不断加快，智能终端设备产生的数据量激增，保护校园数据安全十分重要。

近年来，为了保证校园数据安全，现有方案 [2-3] 大都通过设计复杂的网络安全模型和严格的管理体系来保障数据的安全性。上述方案都把数据以明文的形式在云服务器中进行处理，然而云服务器不能保证数据的安全性。

在这种情况下，可搜索加密技术（SE）能够实现对密文数据的关键词搜索。为保证可搜索加密方案的安全性 [4]、搜索性能 [5]、存储消耗 [6-7] 国内外学者做了广泛研究。随着物联网技术的发展，可搜索加密技术（SE）在物联网（IOT）场景中的应用成为研究的热点。2022 年，陈琪 [8] 提出了一种支持多关键词模糊检索的可搜索加密 SHMP 方案。张明润 [9] 提出了一种支持结果完整性验证查询以及文件动态更新的 VSPSS 方案。张美玲 [10] 提出了轻量级的多授权中心属性基可搜索加密方案。Niu 等人 [11] 提出了一种支持关键字搜索的边缘计算下的高效属性加密方案。Yang 等人 [12] 提出一种基于区块链构建的动态共识可搜索加密方案，实现了密钥和密文可更新的多用户关键词搜索。上述方案虽然解决了云端和边缘端数据传输的安全问题，但数据共享模式上仅支持“一对一”和“一对多”的共享模式，然而智慧校园实际的应用场景是“多对多”共享模式。本文提出了一种去中心化的高效智慧校园数据安全模型，通过属性基加密（SE）方法、边缘计算技术实现密文高效搜索。

首先，本文设计了一种去中心化的模型架构，通过属性基加密（SE）方法、边缘计算功能，支持多对多共享模式。模型架构包括云服务器、5G 基站、部署在网络边缘的边缘服务器、数据所有者和数据用户。数据所有者通过基于属性的 SE 加密方法生成加密数据，并将其发布到最近的边缘服务器。数据用户搜索关键词组并根据属性和密钥进行加密生成加密请求（即加密查询，trapdoors）并发布到所有可用的边缘服务器。边缘服务器验证用户身份权限后把关键词组相关密文返回给授权数据用户。数据用户根据密钥进行解密得到明文。本文设计的模型极大降低了网络传输的延迟，并有效保护了数据用户的隐私。

1 关键词注释

M：明文

C：密文

$W(w_1, w_2, \cdots, w_n)$：关键词组

$X(x_1, x_2, \cdots, x_n)$：属性集

$A(a_1, a_2, \cdots, a_n)$：访问权限集

Q：搜索请求

PK：公钥

SK：私钥

H：哈希函数

e：双线性映射函数

1. 江苏航空职业技术学院文理学院 江苏镇江 212000

[基金项目] 江苏航空职业技术学院院级重点课题（JATC22010108）

2 基于可搜索加密技术的智慧校园数据安全模型设计

2.1 系统模型

本文方案的系统模型如图 1 所示，主要包括 6 个实体，分别是云服务器（cloud server，CS）、边缘服务器（edge server，ES）、数据产生者（data producer，DP）、用户（data user，DU）、属性权威（attribute authority，AA）和可靠的搜索引擎（reliable search engine，RSE）。每个实体的功能具体如下。

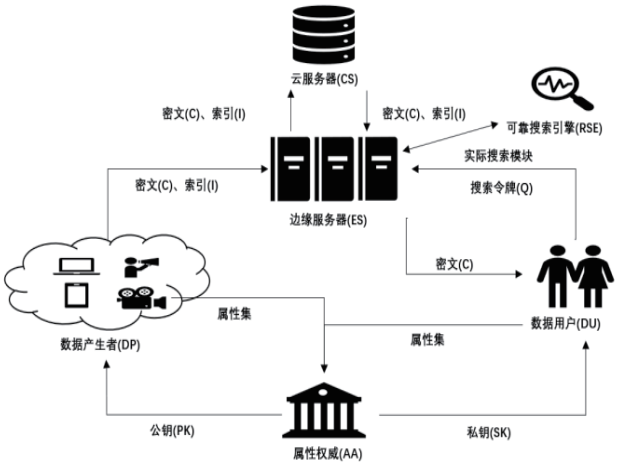

图 1 系统模型图

（1）云服务器（CS）：CS 用于存储密文和索引。CS 收到边缘节服务器上传的密文和索引，把陷密文和索引发送到其他有相应关键词搜索请求的边缘服务器。

（2）边缘服务器（ES）：ES 存储本服务器范围内密文和索引，完成边缘服务器范围内关键词搜索并返回密文。

（3）数据所有者（DP）：DP 对产生的数据行加密，生成密文和索引，然后将密文上传到云服务器 CS 中，索引上传到边缘服务器 ES 中。

（4）用户（DU）：DU 生成搜索令牌，并上传到 ES，当满足查询要求后，ES 返回搜索密文给 DU，DU 使用自己的私钥进行解密。

（5）属性权威（AA）：AA 是可信的第三方实体，AA 用于生成主密钥和系统公共参数。

（6）可靠的搜索引擎（RSE）：RSE 是存放在边缘服务器端用于保存索引和完成关键词搜索的模块。

2.2 方案形式化定义

在智慧校园场景下，实现多用户共享模式的可搜索加密方案由以下 6 个算法组成。

（1）Setup($1^{\lambda,\alpha}$)→(MSK, P_a)：由完全可信的第三方机构 AA 执行该算法，选择随机安全参数输入 λ,α，输出主密钥 MSK 和系统公共参数 P_a。

（2）KeyGen(λ, A, X)→(SK_{du})：由完全可信的第三方机构 AA 执行该算法，输入安全参数 λ，属性集 A，访问权限集 X，输出所有数据用户密钥 (SK_{du})。

（3）Encrypt($P_a, M, W, A, \varepsilon$)→($C$)：数据产生者（DP）输入系统公共参数 P_a、数据明文 M、关键词组 W、访问权限集 A 和随机数 ε，执行该算法得到密文 C。

（4）Req(W, SK_{du}, z)→(Q)：搜索信息生成阶段数据用户（DU）输入搜索关键词 w、用户密钥 SK_{du}、随机数 z，执行该算法得到用户搜索信息。

（5）Search(Q, C)→(C_w)：搜索阶段边缘服务器根据收到的搜索相信进行验证，符合搜索权限且包含该关键词密文，则返回包含关键词 w 的密文 C_w。

（6）Decrypt(C_w, SK_{du})→(m_k)：解密阶段数据用户 (DU) 收到返回的密文后执行该算法得到包含关键词 w 的明文 m_k。

2.3 方案安全性定义

定义 1：如果敌手 D 在多项式时间内赢得下列游戏的概率是忽略不计的，则认为本方案在选择明文攻击下是安全的。

本文方案通过在概率多项式时间内，敌手 D 和挑战者 T 之间的博弈来定义选择明文攻击下的安全性，具体定义如下。

初始化阶段：挑战者 T 执行 Setup 算法生成系统公共参数 P_a 和主密钥 MSK，并将 P_a 发送给敌手 D。

密钥生成阶段：敌手 D 将身份属性集 A 发送给挑战者 T，挑战者 T 执行 KeyGen 算法，并将生成的密钥对 SK_{du} 发送给敌手 D。

加密阶段：敌手 D 选择两个特征相同且访问策略相同的两组明文 m_0 和 m_1 发送给挑战者 T，挑战者 T 随机选择 $b \in (0,1)$ 一组信息进行加密得到密文 C_* 并发送给敌手 D。

猜测阶段：敌手 D 对密文信息进行猜测输出结果 $b' \in (0,1)$，如果 $b'=b$，则敌手 D 赢得游戏，否则失败。敌手 D 赢得游戏的概率定义为：Adv=($p[b'=b]$-1/2)。

3 本文具体方案

（1）Setup($1^{\lambda,\alpha}$)→(MSK, P_a)：系统初始化阶段，由完全可信的第三方机构（AA）输入安全参数 $\lambda, \alpha \in Z_p$，两个 P 阶的素数群 G 和 G_T，一个双线性映射函数 e：$G\times G\to G_T$，其中 $g, h_1, h_2 \in G$，三个哈希函数 H_1, H_2, H_3:$\{0,1\}\to\{0,1\}^{\log p}$。输出主密钥 MSK = ($\lambda, \alpha$) 和系统公共参数 P_a=($g, g^{\alpha}, g^{\lambda}, h_1, h_2$)。属性权威 (AA) 生成属性集 ($x_1, x_2, \cdots, x_n$) 和访问权限集 ($a_1, a_2, \cdots, a_n$)。

（2）KeyGen(λ, A, X)→(SK_{du})：密钥生成阶段：属性权威（AA）输入安全参数 λ，属性集 X 和访问权限集 A 为数据用户生成密钥 (SK_{du}) 具体形式为：

$$SK_{du}=(\lambda,\beta_1,\beta_2)$$

$$(\beta_1=\left(h_1g^{-\lambda}\right)^{\frac{1}{\alpha-\sum_{i=1}^{n}H_1(x_i)}},\beta_2=\left(h_2g^{-\lambda}\right)^{\frac{1}{\alpha-\sum_{i=1}^{n}H_1(x_i)}})$$

（3）Encrypt(P_a, M, W, A, ε)→(C)：数据上传阶段数据产生者（DP）根据公钥 PK_a、数据明文 M、随机数 $\varepsilon(\varepsilon\in Z_p)$、关键词组 $W(w_1,w_2,\cdots,w_n)$ 和访问权限集 A，执行该算法得到密文 C 并上传到最近的边缘服务器中。具体形式如下：

$$C=(c_1,c_2,c_3,c_4,c_5,c_6,v)$$

$$(c_1=g^{\alpha\varepsilon}\cdot g^{-\varepsilon\sum_{i=1}^{l}H_1(a_i)}\ ,\ c_2=e(g,g)^{\varepsilon}\ ,$$

$$c_3=M\cdot e(g,h_1)^{-\varepsilon}\ ,c_4=g^{\lambda v}\cdot g^{-v\sum_{i=1}^{l}H_1(a_i)},$$

$$c_5=e(g,g)^{\varepsilon}\ ,\ c_6=g^{v\sum_{i=1}^{z}H_2(w_i)}\ ,v=(H_3(c_1,c_2,c_3))$$

（4）Req(W, SK_{du}, z)→(Q)：令牌生成阶段数据用户根据自己的密钥 SK_{du}、要查找的关键词组 $W(w_1,w_2,\cdots,w_n)$ 和一个随机数 $z(z\in Z_p)$ 生成搜索的信息并上传到边缘服务器 RSE 中，搜索信息 Q 具体形式为：

$$Q=(q_1,q_2,q_3)$$

$$(q_1=\left(\beta_2^{z\cdot\sum_{i=1}^{k}H_2(w_i)}\right),\ \ q_2=\lambda\cdot z\cdot\sum_{i=1}^{k}H_2(w_i)\ ,q_3=h_2^z)$$

（5）Search(Q, C)→(C_w)：搜索阶段边缘服务器 (RSE) 根据收到的搜索请求信息进行计算，RSE 首先计算是否存在搜索关键词 w_i，如果不存在返回结果 0，则搜索结束。否则继续计算数据用户访问权限集 A 是否与数据属性集 X 相匹配，如果匹配成功则输出结果 1，否则输出结果 0。计算公式为：

$$e(c_4,q_1)\cdot c_5^{q_2}=e(c_6,q_3)$$

（6）Decrypt(C_w, SK_{du})→(m_k)：解密阶段数据用户（DU）收到返回的密文用密钥 SK_{du} 解密得到包含关键词组 k_w 的明文 m_k，计算公式为：

$$m_k=c_3\cdot e(c_1,\beta_1)\cdot c_2^{\lambda}$$

4 方案分析

4.1 方案正确性分析

搜索阶段：

$$e(c_4,q_1)\cdot c_5^{q_2}=e(c_6,q_3)$$

$$e\left(g^{\lambda v}\cdot g^{-v\sum_{i=1}^{l}H_1(a_i)},\beta_2^{z\cdot\sum_{i=1}^{k}H_2(w_i)}\right)\cdot e(g,g)^{\varepsilon\lambda\cdot z\cdot\sum_{i=1}^{k}H_2(w_i)}$$

$$=e(g^{v\sum_{i=1}^{z}H_2(w_i)},h_2^z)$$

$$e\left(g^{v(\lambda-\sum_{i=1}^{l}H_1(a_i))},\left(h_2g^{-\lambda}\right)^{\frac{z\cdot\sum_{i=1}^{k}H_2(w_i)}{\alpha-\sum_{i=1}^{n}H_1(x_i)}}\right)\cdot e(g,g)^{\varepsilon\lambda\cdot z\cdot\sum_{i=1}^{k}H_2(w_i)}$$

$$=e(g^{v\sum_{i=1}^{z}H_2(w_i)},h_2^z)$$

$$e(g,g)^{h_2\cdot v(\lambda-\sum_{i=1}^{l}H_1(a_i))\frac{-\lambda z\cdot\sum_{i=1}^{k}H_2(w_i)}{\alpha-\sum_{i=1}^{n}H_1(x_i)}}\cdot e(g,g)^{\alpha\lambda\cdot z\cdot\sum_{i=1}^{k}H_2(w_i)}$$

$$=e(g^{v\sum_{i=1}^{z}H_2(w_i)},h_2^z)$$

$$e(g,g)^{h_2\cdot v(\lambda-\sum_{i=1}^{l}H_1(a_i))\frac{-\lambda z(\alpha-\sum_{i=1}^{k}H_2(w_i))}{\alpha-\sum_{i=1}^{n}H_1(x_i)}}=e(g^{v\sum_{i=1}^{z}H_2(w_i)},h_2^z)$$

如果访问权限 a_i 和数据集 x_i 相同，则输出结果为 1 返回密文 C_w，否则搜索失败输出结果为 0。

$$e(g,g)^{h_2\cdot v\cdot z\sum_{i=1}^{k}H_2(w_i))}=e(g^{v\sum_{i=1}^{z}H_2(w_i)},h_2^z)$$

解密阶段：

$$c_3\cdot e(c_1,\beta_1)\cdot c_2^{\lambda}=M\cdot e(g,h_1)^{-\varepsilon}\cdot e(g^{\varepsilon(\alpha-1)\sum_{i=1}^{l}H_1(a_i)},$$

$$\left(h_1g^{-\lambda}\right)^{\frac{1}{\alpha-\sum_{i=1}^{n}H_1(x_i)}}))\cdot e(g,g)^{\lambda\varepsilon}$$

$$M\cdot e(g,h_1)^{-\varepsilon}\cdot e(g^{\varepsilon(\alpha-)\sum_{i=1}^{l}H_1(a_i))},$$

$$\left(h_1g^{-\lambda}\right)^{\frac{1}{\alpha-\sum_{i=1}^{n}H_1(x_i)}}))\cdot e(g,g)^{\lambda\varepsilon}$$

$$M\cdot e(g,h_1)^{-\varepsilon}\cdot e(g,h_1^{\frac{-\lambda\varepsilon(\alpha-\sum_{i=1}^{l}H_1(a_i))}{\alpha-\sum_{i=1}^{n}H_1(x_i)}}))\cdot e(g,g)^{\lambda\varepsilon}$$

$$M=M\cdot e(g,g)^{-\lambda\varepsilon}\cdot e(g,g)^{\lambda\varepsilon}$$

4.2 方案安全性分析

定义 1：如果敌手 D 在任意多项式时间内赢得下列游戏的概率是忽略不计的，则认为本方案在选择明文攻击下是安全的。

证明：假设存在一个敌手 D 在任意多项式时间赢得下列游戏的概率是忽略不计的，则称本文方案在选择明文攻击下是安全的。

初始化阶段：挑战者 T 执行 Setup 算法，得到公共参数 P_a 和主密钥 MSK，将 P_a 发送给敌手 D 并保留主密钥 MSK。

密钥生成阶段：敌手 D 将身份属性集 A 发送给挑战者 T 寻求密钥（由三方安全协议生成的密钥是安全的），挑战者 T 选择安全参数 λ^* 执行 KeyGen 算法得到个人密钥 SK_{du} 发送给敌手具体形式为：

$$SK_{duD}=(\lambda^*,\beta_1,\beta_2)$$

$$(\beta_1=\left(h_1g^{-\lambda}\right)^{\frac{1}{\alpha-\sum_{i=1}^{n}H_1(x_i)}},\beta_2=\left(h_2g^{-\lambda}\right)^{\frac{1}{\alpha-\sum_{i=1}^{n}H_1(x_i)}})$$

加密阶段：敌手 D 选择两个特征相同且访问策略相同的两组明文 m_0^* 和 m_1^* 发送给挑战者 T，挑战者 T 随机选择 $b\in(0,1)$ 一组信息进行加密得到密文 C_0^* 和 C_1^* 并发送给敌手 D，密文具体形式为：

$$C_0^*=(c_1,c_2,c_3,c_4,c_5,c_6,v)$$

$$C_1^*=(c_1,c_2,c_3,c_4,c_5,c_6,v)$$

猜测阶段：敌手 D 对密文信息进行猜测输出结果 $b' \in (0,1)$，如果 $b'=b$，则敌手 A 赢得游戏，否则失败。敌手 D 赢得游戏的概率为：$Adv=(p[b'=b]-1/2)$。此概率忽略不计，因此本文方案在选择明文攻击是安全的。

4.3 性能分析

本节通过对比分析各方案在加密、解密、搜索三个阶段的计算开销来分析方案的性能。在分析的过程中，主要考虑指数运算 E 计算复杂性，对于哈希函数运算和乘法运算统一归为常数，相对于指数运算，这些运算可忽略不计。

参考文献 [9-11] 均是引用了边缘服务器构造可搜索加密模型的方案。其中方案 [9] 是基于对称可搜索加密构建的单用户方案，因此在加密、解密、搜索阶段的计算开销上很小，但不适用于目前校园数据安全传输“多对多”用户使用场景。方案 [10-11] 都是基于“多对多”用户场景设计的方案。在加密阶段，本文方案和方案 [10] 均采用数据产生者自行加密的方式，相较于方案 [11] 加密效率较高。在解密阶段本文方案只需一次指数运算，用户的计算开销较低。在搜索阶段本文构造的方案只需两次指数运算即可完成目标文件检索，效率较高。如表 1 所示，通过对三个阶段的比较，在满足“多对多”用户使用场景下，本文方案具备较高的计算性能。

表 1 计算复杂度开销比较

方案	加密	解密	搜索
文献 [9]	$O(n)$	$O(n)$	$O(n)$
文献 [10]	$O(KE)$	$O(E)$	$O(4E)$
文献 [11]	$O((k+3)E)$	$O(3E)$	$O(4E)$
本文方案	$O(KE)$	$O(E)$	$O(2E)$

5 总结

本文基于属性基可搜索加密技术和边缘计算技术提出了一种适用于高校智慧校园建设的数据安全模型。通过安全性分析和性能分析，本文提出的模型在理论上具备较高的性能，未来将根据用户数量、通信需求、数据异构、动态更新管理制定优化算法进行进一步的研究。

参考文献：

[1]《中华人民共和国数据安全法》发布 2021 年 9 月 1 日施行 [J]. 信息技术与标准化 ,2021(7):4.

[2] 杨赛 , 焦驰宇 , 赵子彦 , 等 . 大数据背景下双重预防机制的高校实验室安全管理体系研究与实践 [J]. 实验技术与管理 ,2023,40(11):240-245.

[3] 陈辉定 . 基于计算机网络技术的网络信息安全防护体系构建 [J]. 现代雷达 ,2023,45(2):101-103.

[4]HUANG Q, LI H.An efficient public-key searchable encryption scheme secure against inside keyword guessing attacks[J]. Information sciences,2017,403: 1-14.

[5]ZHANG K, LONG J, WANG X, et al.Lightweight searchable encryption protocol for industrial internet of things[J].IEEE transactions on industrial informatics,2020,17:4248-4259.

[6]WU L, CHEN B, CHOO KIM-K R, et al.Efficient and secure searchable encryption protocol for cloud-based Internet of things[J].Journal of parallel and distributed computing,2018,111: 152-161.

[7]JIANG P, GUO F, LIANG K, et al.Searchain: blockchain-based private keyword search in decentralized storage[J]. Future generation computer systems,2020,107: 781-792.

[8] 陈琪 . 基于云边端协同下的可搜索加密方案 [D]. 西安 : 西安电子科技大学 ,2022.

[9] 张明润 . 基于边缘计算的多关键词可搜索加密技术研究 [D]. 西安 : 西安电子科技大学 ,2023.

[10] 张美玲 . 轻量级的属性基可搜索加密算法研究 [D]. 西安 : 西北师范大学 ,2023.

[11]NIU S F, HU Y, ZHOU S W, et al. Attribute-based searchable encryption in edge computing for lightweight devices[J].IEEE systems journal,2023,17:3503-3514.

[12]YANG N B, TANG C M, ZHOU Q, et al.Dynamic consensus committee-based for secure data sharing with authorized multi-receiver searchable encryption[J].IEEE transactions on information forensics and security, 2023,18:5186-5199.

【作者简介】

王泽贤（1995—），男，江苏宿迁人，硕士，助教，研究方向：计算机网络、信息安全、工业互联网。

（收稿日期：2024-07-22）