一种智能电网敏感图像可检索的属性基加密方案

赵博[1]，李春亮[1]，孙碧颖[2]，徐盼[3]，杨波[2]，卫祥[2]，桂小林[3]
（1. 国网甘肃省电力公司，730000，兰州；2. 国网甘肃省电力公司信息通信公司，730050，兰州；
3. 西安交通大学电子与信息工程学部，710049，西安）

摘要：针对智能电网环境中敏感图像数据面临的被窃取、被滥用等安全威胁，及现有密文图像检索方案难以兼具高检索速度与高检索精度的问题，提出了一种面向智能电网环境的基于多特征融合的敏感图像可检索属性基加密方案。利用神经网络模型提取深度图像特征，并分别提取包括词袋特征、颜色空间特征及方向梯度直方图特征在内的 3 种传统图像特征，采用主成分分析法将这 4 种特征进行融合与降维，以增强特征向量的表示信息；基于 ρ 稳定局部敏感哈希函数和改进后的安全近邻算法对敏感图像集和查询图像分别构建安全索引和搜索陷门，通过对安全索引与搜索陷门进行相似度计算得到密文图像检索结果；将密文策略属性基加密机制与可搜索加密进行结合，只有符合访问策略的用户才能得到明文图像检索结果。理论分析和实验分析表明，在真实的电力数据集中，所提方案的密钥生成时间、索引生成时间及检索时间均为毫秒级，检索精度为 88.6%。3 种对比方案的实验结果表明，所提方案在检索效率与检索精度中的平衡性最好，深度特征的融合能够将 3 种对比方案的检索精度提高 5.7%～11.4%。安全性分析表明，所提方案可抵抗已知密文攻击、已知背景攻击和抗串谋攻击。

关键词：智能电网；密文图像检索；多特征融合；属性基加密；局部敏感哈希
中图分类号：TP309 **文献标志码**：A
DOI：10.7652/xjtuxb202111015 **文章编号**：0253-987X(2021)11-0136-11

OSID 码

Retrievable Attribute-Based Encryption Scheme for Sensitive Images of Smart Grid

ZHAO Bo[1], LI Chunliang[1], SUN Biying[2], XU Pan[3], YANG Bo[2], WEI Xiang[2], GUI Xiaolin[3]
(1. State Grid Gansu Electric Power Company, Lanzhou 730000, China;
2. State Grid Gansu Information & Telecommunications Company, Lanzhou 730050, China;
3. Faculty of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Aiming at the threats to security that sensitive images of smart grid are stolen and abused, and at the problem that the existing ciphertext image retrieval solutions are difficult to achieve both high retrieval rate and accuracy, a retrievable attribute-based encryption scheme for sensitive images of smart grid with multi-feature fusion is proposed. The scheme extracts deep image features through the Dense convolutional network, and respectively extracts three traditional image features including bag-of-words features, colorspace features, and histogram of oriented gradients features. The principal component analysis algorithm is used to fuse and reduce the dimensions of four features to enhance the representative information of the feature vector. Based on the ρ stable locality-sensitive hashing function and the improved secure nearest neighbor

收稿日期：2021-04-19。 作者简介：赵博(1981—)，男，硕士，高级工程师。 基金项目：国家重点基础研究发展规划资助项目(2018YFB1800304)；陕西省重点研发计划资助项目(2019GY-005)。
网络出版时间：2021-07-02 网络出版地址：http://kns.cnki.net/kcms/detail/61.1069.T.20210629.1632.002.html

algorithm, the security index and the trapdoor are constructed for the sensitive image dataset and the query image, respectively. The ciphertext image retrieval result is obtained by calculating the similarity between the security index and the trapdoor. The ciphertext-policy attribute-based encryption mechanism is combined with searchable encryption, and only the users who meet the access strategy can obtain the plaintext image retrieval results. Theoretical analysis and experimental analysis indicate that in the real power dataset, the key generation time, index generation time and retrieval time of the proposed scheme are all in milliseconds, and the retrieval accuracy is 88.6%. Experimental results of the three comparative schemes indicate that the proposed scheme has the best balance between retrieval efficiency and retrieval accuracy. The fusion of deep features can increase the retrieval accuracy of the three comparative schemes by 5.7%－11.4%. The security analysis shows that the proposed solution can resist known ciphertext attacks, known background attacks and collusion attacks.

Keywords: smart grid; ciphertext image retrieval; multi-feature fusion; ciphertext-policy attribute encryption; locality-sensitive hash

在全球能源互联网大力推进的背景之下，各供电企业纷纷建设智能电网，投入使用了大量的智能电表，所产生的非结构化数据比传统电网高出 4 个数量级，其中包括了工程建设中的设计图纸、设备运维中的设备细节照片、事故现场照片、人力资源的证件照、劳动合同等敏感图像[1]。随着数据的爆发式增长和信息技术的快速发展，海量的智能电网图像数据被存储在云服务器上，但现有的云存储服务器大多是“诚实却好奇”的，能够对存储数据进行窥探。因此，为了避免电网企业和个人用户的隐私泄露，图像文件需加密后再上传至云端[2]，但密文图像文件不再具备明文可检索特性。

可搜索加密技术允许在密文状态下进行信息检索。因此，基于可搜索加密的信息安全检索问题成为了国内外科研工作者的研究热点[3]。温蜜等分析了智能电网中的数据安全威胁，提出了两种基于带关键词和隐向量加密的可搜索公钥加密方案的设计思路，但不适用于图像检索[4]；Li 等构建了一种适用于智能电网环境的可检索对称加密方案，通过允许少量的信息泄露实现数据的易于更新[5]；Eltayieb 等提出了一种基于属性的在线/离线可搜索加密方案，将加密算法和陷门算法分为两个阶段，在云智能电网中取得了较好的效果[6]。对于密文图像检索：朱旭东等提出了一种基于安全相似度运算的密文图像检索方案，并证明了方案的安全性和有效性，但检索效率低[7]；Yuan 等通过 K-means 聚类算法对图像集进行分类，构建基于树形结构的图像索引，检索速度有所优化，但检索结果不够精确[8]；李颖莹等针对不同密钥加密图像集的场景，在边缘计算环境中构建了基于安全近邻、局部敏感哈希与代理重加密技术的的加密图像检索方案[9]；Xia 等提出了一种基于加密图像的内容检索而不向云服务器泄露敏感信息的方案，利用局部敏感哈希构建预过滤器表以提高搜索效率[10]；Shen 等支持同时保护多个图像所有者的隐私，利用安全多方计算技术对图像特征进行加密，并提出一种新的相似度计算方法以避免泄露图像相似度信息[11]。

在信息检索的过程中，相似度计算是其中一个必不可少的环节，但不太适用于图像本身之间进行比较，故首先需对能够表示图像的特征进行提取。传统方法一般对颜色直方图[12]、纹理[13]、形状[14]等多种特征进行提取与混合，虽简单易实现，但不能有效应用于所有类型图像的相似度匹配中且精确度较低。近些年来，许多研究者认为采用卷积神经网络[15]等深度学习方法可以提取出更可靠有效的图像特征。Liu 等结合了卷积神经网络模型的高级特征和点扩散块截断编码的低级特征，提出了一种有效的图像检索方法[16]；陈享等利用 Faster R-CNN 模型[17]提取图像集的特征向量和关键词集合，采用先粗后细的分类检索模式，但计算消耗较大[18]。

然而，这些检索方案因检索效率、检索精度或检索模式问题，并不能完全适用于智能电网中的图像检索。智能电网环境中的敏感信息访问是分类别分等级的，密文策略属性基加密机制（CP-ABE）[19]允许数据拥有者制定访问策略，与可搜索加密机制相结合，可实现对密文检索结果的细粒度访问控制，进一步地降低隐私泄露的风险。

针对智能电网中敏感图像数据面临的安全威胁

及现有密文图像检索方案在智能电网场景下的不足，本文提出一种面向智能电网环境的基于多特征融合的敏感图像可检索属性基加密方案(RAESSG)，该方案对图像集同时提取深度特征与传统特征，然后利用主成分分析(PCA)算法对特征向量集合进行降维与融合，基于稳定局部敏感哈希函数和安全近邻函数对融合后的特征向量集构建安全索引，高效安全地对图像进行相似度计算。采用密文策略属性基加密机制对对称共享密钥进行加密，实现对密文图片检索结果的精确访问。安全性分析表明，本文方案可抵抗已知密文攻击、已知背景攻击和抗串谋攻击。实验比较表明，本文方案在智能电网应用场景中可行。

1 预备知识

1.1 双线性映射

设 G_0 和 G_1 是两个在有限域 Z_p 上的乘法循环群，阶 p 是大安全素数，则双线性映射 $f:G_0\times G_0\rightarrow G_1$ 具有以下性质。

(1)双线性。$\forall m,n\in Z_p$ 及 $\forall x,y\in G_0$，$f(x^m,y^n)=f^{mn}(x,y)$；$\forall x_1,x_2,y\in G_0$，$f(x_1x_2,y)=f(x_1,y)f(x_2,y)$。

(2)非退化性。$\forall x,y\in G_0$，使 $f(x,y)\neq 1$，其中 1 是 G_1 的单位元。

(3)可计算性。$\forall x,y\in G_0$，都存在一种有效多项式时间算法计算 $f(x,y)$。

1.2 访问控制树

假设参与者为集合 $U=\{U_1,U_2,\cdots,U_n\}$，若存在单调访问结构 $S\subseteq 2^{\{U_1,U_2,\cdots,U_n\}}$，对任意的 K_1、K_2，若 $K_1\subseteq S$ 且 $K_2\subseteq K_1$，即有 $K_2\subseteq S$。单调的访问结构 S 是一个非空的单调集合 S，集合 S 被称为授权集合，否则为非授权集合。

本文方案所利用的数据访问策略是通过构建由一组非叶子节点(逻辑阈值门)和叶子节点(用户属性)组成的访问控制树 Γ 进行实现，当某个用户的属性集合满足访问控制树 Γ 的逻辑规则，则该用户可以访问相应的内容。本文方案通过构建访问控制树实现细粒度的访问控制策略。逻辑阈值门中最常用的为 AND 门和 OR 门。给定一个访问控制树 Γ，假设 n_o 是节点 o 的子节点数，$k_o(0<k_o<n_o)$ 为节点 o 的门限值，则当节点 o 的值为真时，至少有 k_o 个子节点的值也为真。即：当节点 o 为 AND 门时，门限值 $k_o=n_o-1$；当节点 o 是叶子节点或 OR 门时，门限值 $k_o=0$。

1.3 ρ 稳定局部敏感哈希函数

ρ 稳定局部敏感哈希函数(ρ stable LSH)是局部敏感哈希函数中的一种，可被用在欧式空间中。ρ stable LSH 可将 l 维特征向量 $\boldsymbol{V}$ 映射为一个数，即

$$h_{a,b}(\boldsymbol{V})=\left\lfloor\frac{\boldsymbol{aV}+b}{W}\right\rfloor \tag{1}$$

式中：$\boldsymbol{a}$ 是由 ρ 稳态分布生成的随机向量；W 是常数，代表着哈希桶宽；b 是由均匀分布$[0,W]$生成的实数；$\lfloor\rfloor$是向下取整运算符号。ρ stable LSH 算法能对维度较高的特征向量构建倒排索引，有效地降维特征并减小相关的计算量。

2 问题定义

2.1 智能电网敏感图像可检索加密方案框架

本文方案包括 4 个实体：① 可信授权机构(TA)，负责系统初始化和用户相关密钥的生成与分发；②云服务器(CS)，存储大量图像文件并提供大部分的计算资源，利用搜索陷门、密文图像、安全索引等进行检索；③图像数据拥有者(DO)，将图像文件存储在云服务器上的电网企业或个人用户，对图像集和加密密钥进行加密并构建安全索引；④图像数据查询者(DU)，想要在云服务器中进行查询的电网企业或者个人用户，利用属性私钥对 CS 返回的密文检索结果进行解密。系统框架如图 1 所示，图中①～⑦为方法步骤。

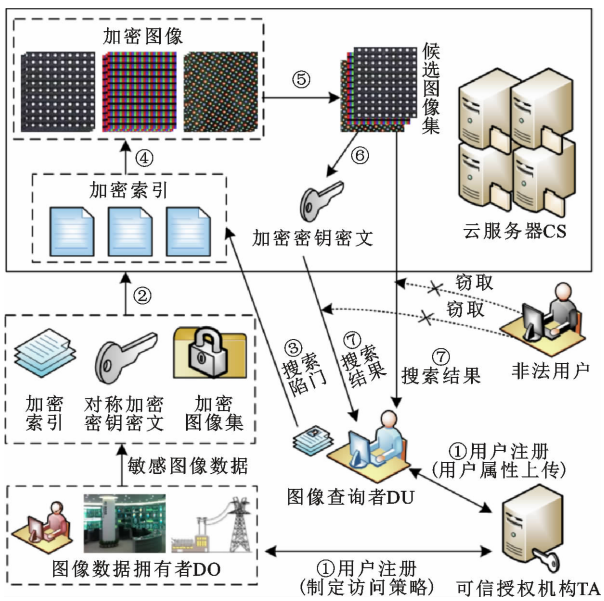

图 1 系统框架

Fig. 1 System structure diagram

2.2 设计目标

为保障智能电网环境中图像数据拥有者的隐私

性和密文图像检索的可行性，本文设计安全和性能目标如下。

(1)密文图像可检索。云服务器 CS 中存储了大量的智能电网密文图像，图像数据查询者 DU 可以利用现有的明文图像在 CS 中检索出相似图像对应的密文图像。

(2)精确地检索及访问。将图像的深度特征与传统特征进行融合与降维，提高检索精度，同时结合密文策略属性加密机制，杜绝非法授权访问。

(3)高效地检索。基于 ρ stable LSH 算法和安全近邻算法构建安全索引，降低相似度计算的复杂度，进而提高检索效率。

(4)数据的隐私保护。云服务器 CS 和其他攻击者无法探测到密文图像集、密文特征向量、安全索引和搜索陷门的明文信息，同时云服务器 CS 无法从密文向量内积计算结果中推断出明文向量内积值。

2.3　安全模型

本文认为 CS 是半可信的，它会诚实而正确地执行 DO 和 DU 发送的请求，但也会主动探测存储内容，并进行关联推理以挖掘更多信息。系统攻击模型可被总结为以下 3 种。

(1)已知密文攻击模型。CS 只知道来自 DO 的加密图像集和安全索引以及来自 DU 的搜索陷门。

(2)已知背景攻击模型。除了密文图像集和搜索陷门，CS 可以获取到更多的背景知识，如一些明文图像、搜索陷门对应的明文等。

(3)用户串谋攻击模型。不符合访问策略的 DU 间相互共享的自己的属性私钥及对称加密密钥密钥，以期能够访问系统的检索结果。

3　多特征融合的智能电网敏感图像可检索属性基加密方案

本文方案总体流程如图 2 所示。首先，TA 初始化系统，为图像数据拥有者 DO 和图像数据查询者 DU 生成属性私钥；其次，DO 利用卷积神经网络(CNN)和传统方法分别提取出图像的 CNN 特征、词袋(BOW)特征、颜色空间(ColorSpace)特征及方向梯度直方图(HOG)特征共 4 类特征，利用 PCA 和 ρ stable LSH 函数对融合特征向量进行降维并生成安全索引，将密文图像集、安全索引和密文加密密钥上传到云服务器 CS；再次，查询用户 DU 输入一张待查询图像，对该图像提取特征并生成搜索陷门以发送给 CS；最后，CS 根据搜索陷门在安全索引中进行相似度匹配以获取候选图像集返回给 DU。当且仅当 DU 的用户属性集合满足 DO 的访问策略时，DU 才能解密密文图像，得到明文图像。

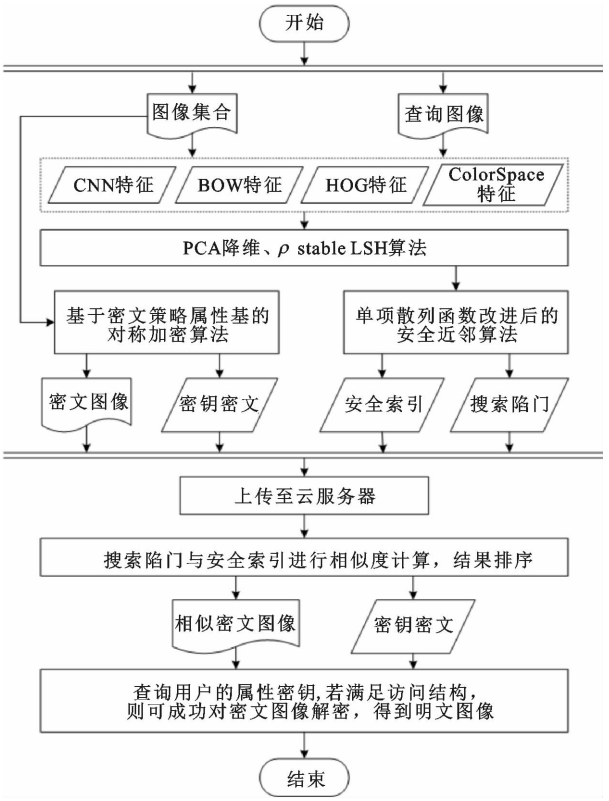

图 2　本文方案总体流程

Fig. 2　Overall flow chart of proposed scheme

3.1　多特征提取与降维

本方案主要提取了 4 类特征，分别是 CNN 特征、Bow 特征、HOG 特征和 ColorSpace 特征。稠密连接网络(DenseNet)[20]的每层输入都与之前每一层的输出有关，可以最大限度地利用各层特征，且省参数与计算量，因此本文通过 DenseNet 网络对图像集进行提取1 024维的 CNN 特征，如图 3 所示。

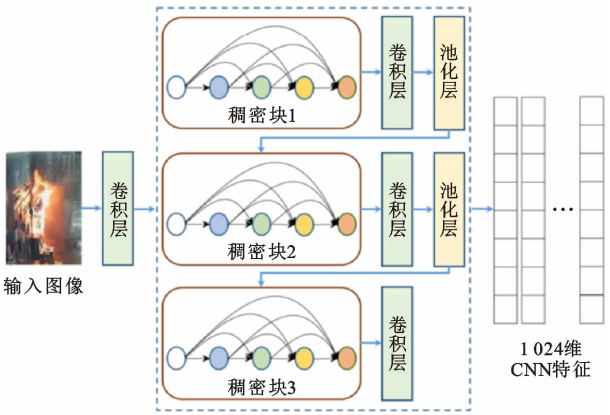

图 3　CNN 特征提取结构

Fig. 3　CNN feature extraction structure diagram

Bow 特征是对图像尺度不变特征变换特征的词袋语义表示，通过 K-means 算法聚类可得到 120 维的 BOW 特征。HOG 特征是一种用来检测图像物体存在的全局图像特征描述子，通过将一副图像区域分割为多个连通的局部区域，计算每个局部区域的边缘或梯度方向直方图，然后进行统计与归一化，最终得到3 780维的 HOG 特征。ColorSpace 特征可对图像颜色的空间分布进行有效地表示，将图像的 RGB 颜色空间映射至 YUV 颜色空间，然后将图像进行分割，对每个局部区域对 YUV 颜色空间的平均值及 DCT 系数矩阵进行计算，最后得到 120 维的 ColorSpace 特征。RGB 颜色空间映射至 YUV 颜色空间的计算式为

$$\left.\begin{aligned} I_{\mathrm{Y}} &= 0.299I_{\mathrm{R}} + 0.587I_{\mathrm{G}} + 0.114I_{\mathrm{B}} \\ I_{\mathrm{U}} &= -0.174I_{\mathrm{R}} - 0.289I_{\mathrm{G}} - 0.436I_{\mathrm{B}} \\ I_{\mathrm{V}} &= 0.615I_{\mathrm{R}} - 0.515I_{\mathrm{G}} - 0.100I_{\mathrm{B}} \end{aligned}\right\} \quad (2)$$

对图像集提取 4 种特征后进行特征融合，然后采用 PCA 算法降低不同特征之间的相关性以对融合特征进行降维，主要步骤如下。

(1)设图像集的特征矩阵为 $\boldsymbol{V}=[\boldsymbol{V}_1,\cdots,\boldsymbol{V}_i,\cdots,\boldsymbol{V}_n]^{\mathrm{T}}$，第 i 幅图像的特征向量为 $\boldsymbol{V}_i=[v_{i1},v_{i2},\cdots,v_{ie}]$，其中，$n$ 为图像集的图像数，e 为特征向量的维度。对特征矩阵 $\boldsymbol{V}$ 进行中心化得到 $\hat{\boldsymbol{V}}$

$$\hat{\boldsymbol{V}}=[\boldsymbol{V}_1-\bar{\boldsymbol{V}}_1,\cdots,\boldsymbol{V}_i-\bar{\boldsymbol{V}}_i,\cdots,\boldsymbol{V}_n-\bar{\boldsymbol{V}}_n)]^{\mathrm{T}} \quad (3)$$

$$\bar{\boldsymbol{V}}_i=\sum_{j=1}^{e}v_{ij}/e,\ 1\leqslant i\leqslant n \quad (4)$$

(2)计算 $\hat{\boldsymbol{V}}$ 的协方差，计算式为

$$\mathrm{cov}(\hat{\boldsymbol{V}})=E([\hat{\boldsymbol{V}}-E(\hat{\boldsymbol{V}})][\hat{\boldsymbol{V}}-E(\hat{\boldsymbol{V}})]^{\mathrm{T}}) \quad (5)$$

(3)对协方差 $\mathrm{cov}(\hat{\boldsymbol{V}})$ 进行奇异值分解，进而得到特征值矩阵 $\boldsymbol{\lambda}$ 和特征向量矩阵 $\boldsymbol{\zeta}$

$$[\boldsymbol{\lambda},\boldsymbol{\zeta}]=\mathrm{SVD}(\mathrm{cov}(\hat{\boldsymbol{V}})) \quad (6)$$

式中：$\boldsymbol{\lambda}=[\lambda_1,\lambda_2,\cdots,\lambda_l]$；$\boldsymbol{\zeta}=[\boldsymbol{\zeta}_1,\boldsymbol{\zeta}_2,\cdots,\boldsymbol{\zeta}_l]$，$l$ 代表特征集的数量，$\boldsymbol{\zeta}_i$ 代表特征值 λ_i 所对应的特征向量；SVD 代表奇异值分解函数。

(4)对特征值矩阵 $\boldsymbol{\lambda}$ 中的元素进行降序处理，然后根据降序处理后的特征值矩阵对特征向量矩阵 $\boldsymbol{\zeta}$ 的元素位置进行调换，选择前 d 个特征向量得到的投影矩阵 $\boldsymbol{\zeta}'=[\boldsymbol{\zeta}_1,\cdots,\boldsymbol{\zeta}_d]$。

(5)根据投影矩阵 $\boldsymbol{\zeta}'$ 和特征矩阵 $\boldsymbol{V}$ 得到降维后的特征矩阵 $\boldsymbol{V}'$

$$\boldsymbol{V}'=\boldsymbol{\zeta}'^{\mathrm{T}}\boldsymbol{V}=[\boldsymbol{V}'_1,\cdots,\boldsymbol{V}'_i,\cdots,\boldsymbol{V}'_n]^{\mathrm{T}} \quad (7)$$

式中 $\boldsymbol{V}'_i=[v'_{i1},v'_{i2},\cdots,v'_{id}]$ 为图像 i 降维后的融合特征向量，d 为特征的维数。

3.2 基于 CP-ABE 的密文图像检索算法

3.2.1 *系统初始化* $\mathrm{Setup}(1^{\varepsilon})\rightarrow\boldsymbol{K}_{\mathrm{P}},\boldsymbol{K}_{\mathrm{S}}$。输入安全系数 ε，定义 G_0 和 G_1 是两个有线域 Z_p 上的乘法循环群，双线性映射 $f:G_0\times G_0\rightarrow G_1$，$H=\{h_1,h_2,\cdots,h_\gamma\}$ 是 γ 个抗碰撞散列函数集合，g 是 G_0 的生成元，阶 p 是大安全素数。设定图像的最终特征向量维数为 d，μ 为随机正整数，则随机选取 $d+\mu+1$ 维二值向量 $\boldsymbol{U}$、$(d+\mu+1)\times(d+\mu+1)$ 维可逆矩阵 $\boldsymbol{A}_1$ 和 $\boldsymbol{A}_2$。TA 随机选择 $\alpha,\beta\leftarrow Z_p$ 作为系统主私钥 $\boldsymbol{K}_{\mathrm{S}}=[\alpha,\beta]$，主公钥 $\boldsymbol{K}_{\mathrm{P}}=[G_0,G_1,p,g,f^{\alpha}(g,g),g^{\beta},H,\boldsymbol{U},\boldsymbol{A}_1,\boldsymbol{A}_2]$ 为系统公共参数。

3.2.2 *用户注册阶段* $\mathrm{URegist}(\boldsymbol{K}_{\mathrm{S}},\boldsymbol{K}_{\mathrm{P}},S)\rightarrow\boldsymbol{K}_{\mathrm{U}}$。当新用户向系统请求注册时，TA 随机选择 $r\leftarrow Z_p$、$r_s\leftarrow Z_p$，并按照该用户的属性集合 S 计算其相应的属性私钥 $\boldsymbol{K}_{\mathrm{U}}$ 为

$$\left.\boldsymbol{K}_{\mathrm{U}}=\left[L=g^{\frac{\alpha+r}{\beta}},L_s=g^{r}h^{r_s}(s),L'_s=g^{r_s}\right]\right\}_{\forall s\in S} \quad (8)$$

3.2.3 *图像加密阶段* $\mathrm{Enc}(\boldsymbol{K}_{\mathrm{P}},D,\Gamma,\boldsymbol{K})\rightarrow C,\boldsymbol{C}_{\mathrm{K}}$。假设 DO 有明文图片集 $D=\{D_1,D_2,\cdots,D_n\}$ 需进行存储。

(1)$\mathrm{PEnc}(D,\boldsymbol{K})\rightarrow C$，图像集与特征向量加密。DO 随机生成 k 维的二值向量 $\boldsymbol{K}$ 作为明文图像集 D 的对称加密密钥，对其进行加密以得到密文图像集 $C=\{C_1,C_2,\cdots,C_n\}$，并发送给 CS。

(2)$\mathrm{KEnc}(\boldsymbol{K}_{\mathrm{P}},\boldsymbol{K},\Gamma)\rightarrow\boldsymbol{C}_{\mathrm{K}}$，加密密钥加密。DO 根据访问策略定义访问树结构 Γ，以根节点 r 为起点，为每个节点 o 生成一个多项式 q_o，设 k_o 为节点 o 的门限值，多项式 q_o 的阶 $\deg(q_o)=k_o-1$。对于根节点 r，DO 从 Z_p 中随机选择 $\deg(q_r)$ 个系数和一个常数 δ(作为 $q_r(0)$ 的值)以生成多项式 q_r；至于其他任意非根节点 o 的多项式 q_o 的生成，$\deg(q_o)$ 个系数同样是从 Z_p 中随机选取，常数项 $q_o(0)=q_{\mathrm{par}(o)}(\mathrm{id}(o))$，其中，$\mathrm{par}(o)$ 为节点 o 的父节点，$\mathrm{id}(o)$ 为节点 o 的父节点给节点 o 的标识值。为方便理解，假设节点 o 是其父节点的第 2 个子节点，则节点 o 的多项式 q_o 的常数项 $q_o(0)=q_{o_{\mathrm{f}}}(2)$，其中，函数 $q_{o_{\mathrm{f}}}$ 为父节点的多项式。DO 计算加密密钥 $\boldsymbol{K}$ 后的密文 $\boldsymbol{C}_{\mathrm{K}}$ 为

$$\boldsymbol{C}_{\mathrm{K}}=[\Gamma,\bar{\boldsymbol{C}}=\boldsymbol{K}f^{\alpha\delta}(g,g),\ C=g^{\delta\beta},$$
$$C_t=g^{q_t(0)},\ C'_t=h^{q_t(0)}(\mathrm{at}(t))],\ \forall t\in T \quad (9)$$

式中：T 为树 Γ 中所有叶子节点；$\mathrm{at}(t)$ 表示与叶子节点 t 相关联的属性值；$q_t(0)$ 为叶子节点 t 的多项式中的常数项。

3.2.4　索引生成阶段　IndexGen($\boldsymbol{K}_P$, D)→$\boldsymbol{I}'$。DO 根据 3.1 小节中的算法对图像集 D 的多个特征进行提取、融合与降维，得到明文特征集 $\boldsymbol{V}'=[\boldsymbol{V}'_1,\boldsymbol{V}'_2,\cdots,\boldsymbol{V}'_n]^T$，然后利用 ρ stable LSH 函数构建明文索引 $\boldsymbol{I}$，接着通过单项散列函数和安全近邻算法得到安全索引 $\boldsymbol{I}'$。

(1)Index($\boldsymbol{K}_P$, D)→$\boldsymbol{V}'$, $\boldsymbol{I}$，未加密索引生成。对于每幅图像 D_i，DO 利用抗碰撞散列函数集合 $H=\{h_1,h_2,\cdots,h_\gamma\}$ 对其特征向量 $\boldsymbol{V}'_i$ 进行哈希运算，得到 $\boldsymbol{H}(\boldsymbol{V}'_i)=[h_1(\boldsymbol{V}'_i),h_2(\boldsymbol{V}'_i),\cdots,h_\gamma(\boldsymbol{V}'_i)]$。重复 σ 遍运算，得到 σ 个哈希表，以减少错误率。$\boldsymbol{H}_t(\boldsymbol{V}'_i)$ 为第 t 次哈希运算得到的哈希表的桶值集，其中，$t\in[1,\sigma]$，N_t 表示第 t 个哈希表中的桶数，相似图像会被放在同一个哈希桶中。图像 D_i 的特征向量 $\boldsymbol{V}'_i$、标识 P_i 共同存储在哈希表里，即索引 $\boldsymbol{I}$ 中。

(2)IndexEnc($\boldsymbol{K}_P$, $\boldsymbol{V}'$, $\boldsymbol{I}$)→$\boldsymbol{I}'$，安全索引生成。对每幅图片 D_i，将其 d 维特征向量 $\boldsymbol{V}'_i$ 扩展为 $d+\mu+1$ 维向量 $\breve{\boldsymbol{V}}'_i=[\boldsymbol{V}'_i,\|\boldsymbol{V}'_i\|^2,\boldsymbol{\vartheta}]$，$\boldsymbol{\vartheta}$ 为 μ 维的随机二值向量，再根据分裂规则和随机二值向量 $\boldsymbol{U}$ 将 $\breve{\boldsymbol{V}}'_i$ 分裂成两部分 $\breve{\boldsymbol{V}}'^1_i$ 和 $\breve{\boldsymbol{V}}'^2_i$。接着，用随机可逆矩阵 $\boldsymbol{A}_1^T$ 和 $\boldsymbol{A}_2^{-1}$ 分别乘以 $\breve{\boldsymbol{V}}'^{1T}_i$ 和 $\breve{\boldsymbol{V}}'^{2T}_i$，可得加密后的特征向量 $\boldsymbol{F}'_i=[\boldsymbol{A}_1^T\breve{\boldsymbol{V}}'^{1T}_i,\boldsymbol{A}_2^{-1}\breve{\boldsymbol{V}}'^{2T}_i]$，用单项散列函数 $\varphi(\cdot)$ 加密 σ 个哈希表的桶值。DO 将安全索引 $\boldsymbol{I}'$ 发送给 CS。其中，分裂规则的伪代码如下。

INPUT：$d+\mu+1$ 维的二值向量 $\boldsymbol{U}$ 和向量 $\breve{\boldsymbol{V}}'_i$

OUTPUT：$\breve{\boldsymbol{V}}'^1_i$、$\breve{\boldsymbol{V}}'^2_i$

```
BEGIN
  /* U[θ]等于1时，3个变量赋值相等 */
  if(θ∈[1,d+μ+1]&&U[θ]==1)
    {/* 3个变量相等 */
      V'_i^1[θ]=V'_i[θ];V'_i^2[θ]=V'_i[θ];}
  /* U[θ]不等于1，V'_i^1[θ]+V'_i^2[θ]=V'_i[θ] */
  else if(θ∈[1,d+μ+1]&&U[θ]! =1)
    {/* 直到满足 V'_i^1[θ]+V'_i^2[θ]=V'_i[θ]条件
    才会停止循环 */
      while(! (V'_i^1[θ]+V'_i^2[θ]==V'_i[θ]))
        {/* Random()函数随机生成一个正值 */
          V'_i^1[θ]=Random();
          V'_i^2[θ]=Random();}}
END
```

3.2.5　搜索陷门生成阶段　TrapdoorGen($\boldsymbol{K}_P$, m)→$\boldsymbol{T}_m$。图像查询者 DU 输入公共参数 $\boldsymbol{K}_P$、所查询的图像 m，首先计算图像 m 的融合特征向量 $\boldsymbol{V}'_m=[v'_{m1},v'_{m2},\cdots,v'_{md}]$，然后利用函数集合 $H=\{h_1,h_2,\cdots,h_\gamma\}$ 进行哈希计算得到 $\boldsymbol{H}(\boldsymbol{V}'_m)=[h_1(\boldsymbol{V}'_m),h_2(\boldsymbol{V}'_m),\cdots,h_\gamma(\boldsymbol{V}'_m)]$，重复 σ 遍，得到 σ 个哈希表，$\boldsymbol{H}_t(\boldsymbol{V}'_m)$ 为第 t 次哈希运算得到的哈希表的桶值集，其中 $t\in[1,\sigma]$。然后，通过单项散列函数 $\varphi(\cdot)$ 对其加密得到 $\boldsymbol{\psi}_m=[\varphi(\boldsymbol{H}_1(\boldsymbol{V}'_m)),\cdots,\varphi(\boldsymbol{H}_\sigma(\boldsymbol{V}'_m))]$。设定 $\boldsymbol{\omega}=[\omega_1,\omega_2,\cdots,\omega_\mu]$，将 $\boldsymbol{V}'_m$ 扩展到 $\breve{\boldsymbol{V}}'_m=[-2\boldsymbol{V}'_m,1,\boldsymbol{\omega}]$，根据分裂规则和随机二值向量 $\boldsymbol{U}$ 将 $\breve{\boldsymbol{V}}'_m$ 分裂成 $\breve{\boldsymbol{V}}'^1_m$ 和 $\breve{\boldsymbol{V}}'^2_m$ 两部分。再随机选取正整数 τ，用 $\boldsymbol{A}_1^{-1}$ 和 $\boldsymbol{A}_2^T$ 分别乘以 $\breve{\boldsymbol{V}}'^{1T}_m$ 和 $\breve{\boldsymbol{V}}'^{2T}_m$，得到加密查询向量 $\boldsymbol{F}'_m=[\tau\boldsymbol{A}_1^{-1}\breve{\boldsymbol{V}}'^{1T}_m,\tau\boldsymbol{A}_2^T\breve{\boldsymbol{V}}'^{2T}_m]$。最后，将 $\boldsymbol{F}'_m$ 和 $\boldsymbol{\psi}_m$ 组成的搜索陷门 $\boldsymbol{T}_m$ 发送给云服务器。其中，分裂规则的伪代码如下。

INPUT：$d+\mu+1$ 维的二值向量 $\boldsymbol{U}$ 和向量 $\boldsymbol{V}'_m$

OUTPUT：$\breve{\boldsymbol{V}}'^1_m$、$\breve{\boldsymbol{V}}'^2_m$

```
BEGIN
  /* U[θ]不等于1时，3个变量赋值相等 */
  if(θ∈[1,d+μ+1]&&U[θ]! =1)
      {/* 3个变量相等 */
      V'_m^1[θ]=V'_m[θ];V'_m^2[θ]=V'_m[θ];}
  /* U[θ]等于1时，V'_m^1[θ]+V'_m^2[θ]=V'_m[θ] */
  else if(θ∈[1,d+μ+1]&&U[θ]! =1)
    {/* 直到满足 V'_m^1[θ]+V'_m^2[θ]==V'_m[θ]条
    件才会停止循环 */
      while(! (V'_m^1[θ]+V'_m^2[θ]==V'_m[θ]))
        {/* Random()函数随机生成一个正值 */
          V'_m^1[θ]=Random();
          V'_m^2[θ]=Random();}}
END
```

3.2.6　搜索阶段　Search($\boldsymbol{I}'$, $\boldsymbol{T}_m$)→$\boldsymbol{C}_K$, $\boldsymbol{C}_m$。云服务器 CS 接收到密文图像集安全索引 $\boldsymbol{I}'$、搜索令牌 $\boldsymbol{T}_m$，首先查找与 $\boldsymbol{\psi}_m$ 相同的哈希桶，然后对该桶中所有图片的特征向量与查询图像的特征向量进行内积计算，并根据值的大小衡量两张图像的相似性，公式如下

$$\begin{aligned}&\boldsymbol{F}'^T_m\boldsymbol{F}'_i=\\&(\tau\boldsymbol{A}_1^{-1}\breve{\boldsymbol{V}}'^{T1}_m)^T\boldsymbol{A}_1^T\breve{\boldsymbol{V}}'^T_m+(\tau\boldsymbol{A}_2^T\breve{\boldsymbol{V}}'^{T2}_m)^T\boldsymbol{A}_2^{-1}\breve{\boldsymbol{V}}'^{T2}_i=\\&\tau\breve{\boldsymbol{V}}'_m\breve{\boldsymbol{V}}'^T_i=\tau(\|\boldsymbol{V}'_i\|^2-2\boldsymbol{V}'^T_i\boldsymbol{V}'_m+\boldsymbol{\omega}\boldsymbol{\vartheta}^T)=\\&\tau(\|\boldsymbol{V}'_m-\boldsymbol{V}'_i\|^2-\|\boldsymbol{V}'_m\|^2+\boldsymbol{\omega}\boldsymbol{\vartheta}^T)\end{aligned}\quad(10)$$

根据内积值大小输出相似度排在前 j 张的密文图像集 $C_m=\{C_1,\cdots,C_j\}$ 以及包含了对应图像访问策略的密钥密文 $\boldsymbol{C}_K$。

3.2.7　解密阶段　Dec($\boldsymbol{K}_U$, $\boldsymbol{C}_K$, C_m)→D_m/ϕ。图像

数据请求者 DU 需输入属性私钥 $\boldsymbol{K}_{\mathrm{U}}$、云服务器 CS 返回的检索结果 C_m 和包含了对应图像访问策略的密钥密文 $\boldsymbol{C}_{\mathrm{K}}$，本地验证是否具有解密权限。

(1) $\mathrm{Test}(\boldsymbol{K}_{\mathrm{U}},\boldsymbol{C}_{\mathrm{K}})\rightarrow\boldsymbol{K}/\phi$，验证密钥。DU 收到 $\boldsymbol{C}_{\mathrm{K}}$ 后，检查属性私钥 $\boldsymbol{K}_{\mathrm{U}}$ 和访问策略 Γ 是否匹配，如果匹配不上则返回 ϕ，否则采用递归算法，得到 $B=f^{r\delta}(g,g)$。DU 恢复对称密钥的公式为

$$\boldsymbol{K}=\frac{\bar{\boldsymbol{C}}}{B^{-1}f(C,L)}=\frac{\boldsymbol{K}f^{\alpha\delta}(g,g)f(g^{\delta\beta},g^{\frac{\alpha+r}{\beta}})}{f^{r\delta}(g,g)}=\frac{\boldsymbol{K}f^{\alpha\delta}(g,g)f^{r\delta}(g,g)}{f^{\delta(\alpha+r)}(g,g)} \tag{11}$$

(2) $\mathrm{Dec}(C_m,\boldsymbol{K})\rightarrow D_m$，用户解密。利用解密而来的对称密钥 $\boldsymbol{K}$ 解密检索结果 $C_m=\{C_1,C_2,\cdots,C_j\}$，从而得到明文图像集 $D_m=\{D_1,D_2,\cdots,D_j\}$。

4 安全性分析与证明

4.1 安全性分析

本文提出的方案中，加密图像集的方法为传统对称加密算法，采用高安全性的图像加密算法能够强力地保护图像的隐私。其中，由于安全近邻算法中的随机矩阵相乘与添加冗余项等操作，安全索引表中存储的密文特征向量受到了更高强度的隐私保护。若云服务器 CS 想要得到明文特征向量，就首先需要获得向量 $\boldsymbol{U}$ 和矩阵 $\boldsymbol{A}_1$、$\boldsymbol{A}_2$。但是，由于向量 $\boldsymbol{U}$ 和矩阵 $\boldsymbol{A}_1$、$\boldsymbol{A}_2$ 都是随机生成的，云服务器 CS 几乎没有可能得到这些信息的确切数据，即无法解密安全索引表中的密文特征向量。搜索陷门的隐私安全性同样依赖于随机向量 $\boldsymbol{U}$ 和随机矩阵 $\boldsymbol{A}_1$、$\boldsymbol{A}_2$ 的数据是否会被窃取。除此之外，随机数 τ 使搜索陷门的生成存在随机性，故云服务器 CS 无法对搜索陷门与查询图像进行关联推理。

不管是系统用户还是外部攻击者，想要获取明文图像都得先对对称加密密钥的密文 $\boldsymbol{C}_{\mathrm{K}}$ 进行解密。由式(9)可得，$\bar{\boldsymbol{C}}=\boldsymbol{K}f^{\alpha\delta}(g,g)$，故想要得到对称加密密钥 $\boldsymbol{K}$，就必须先算出 $f^{\alpha\delta}(g,g)$，只有当拥有满足访问结构树属性的用户才能得到 $f^{\alpha\delta}(g,g)$。为了能够在系统正常访问到检索结果，不满足访问策略的用户可能会进行串谋，将各自不同的属性集合到一起，以期满足访问结构树的规则。恶意用户可以根据自己的密钥密文中 $\boldsymbol{C}_{\mathrm{K}}$ 的 $C=g^{\beta\delta}$ 和自己的属性私钥 $\boldsymbol{K}_{\mathrm{U}}$ 中的 $L=g^{(\alpha+r)/\beta}$ 计算出 $f^{\delta(\alpha+r)}(g,g)$。恶意用户要得到 $f^{\alpha\delta}(g,g)$ 需先得到 $f^{r\delta}(g,g)$。系统中每个用户的 r 是由 TA 随机生成且对用户隐藏，用户不知道自己或其他人的 r。除此之外，用户属性私钥的其他组成部分 $L_s=g^r h^{r_s}(s)$ 和 $L'_s=g^{r_s}$ 也受到随机数 r 的影响。因此，哪怕多个恶意用户想要通过分享各自的属性密钥或对称加密密钥密文，但各自的组成部分并不相同，仍然无法解析出对称加密密钥 $\boldsymbol{K}$。本文所提方案 RAESSG 可以抵抗系统中用户之间的串谋攻击。

4.2 安全性证明

定理 1 云服务器 CS 和外部敌手无法学习到明文图像的任何信息，除了密文图像数据。

证明 在本文提出的方案中，DO 通过随机生成的对称密钥 $\boldsymbol{K}$ 对明文图像 D_i 进行加密，基于 CP-ABE 将对称加密密钥 $\boldsymbol{K}$ 加密为密文 $\boldsymbol{C}_{\mathrm{K}}$，然后将密文图像和 $\boldsymbol{C}_{\mathrm{K}}$ 上传至 CS 中。因此，当 CS 和外部敌手截获到信道中的密文图像，想要获取明文图像必须先对 $\boldsymbol{C}_{\mathrm{K}}$ 进行解密。由式(9)可得，$\bar{\boldsymbol{C}}=\boldsymbol{K}f^{\alpha\delta}(g,g)$，故想要得到 $\boldsymbol{K}$，就必须先算出 $f^{\alpha\delta}(g,g)$。基于离散对数问题，CS 和外部敌手由 $f^{\alpha}(g,g)$ 和 $C=g^{\beta\delta}$ 计算得到 $f^{\alpha\delta}(g,g)$ 是困难的。同样，基于迪菲-赫尔曼问题，云服务器 CS 和外部敌手无法得到与访问策略相匹配的属性私钥，故无法计算得知访问树 Γ 根节点的值 $f^{\alpha\delta}(g,g)$，即 $\boldsymbol{K}$ 无法被算出。只有当属性私钥与访问树结构 Γ 相匹配时，$\boldsymbol{K}$ 才能被算出以对密文图像进行解密。但是，由于每个属性私钥都是由用户本身自己保留，故除密文图像外，CS 和外部敌手学习不到明文图像的任何信息。

定理 2 在已知背景攻击模型下，内积计算可在多项式时间内确保密文特征向量和搜索陷门安全。

证明 敌手能够探测到密文图像集、安全索引表和搜索陷门集，甚至于其他背景知识。假设敌手 A 可以获取到一组特征向量的明密文对$[\boldsymbol{V}',\boldsymbol{F}']$、一组陷门中特征向量的明密文对$[\boldsymbol{V}'_{\mathrm{td}},\boldsymbol{F}'_{\mathrm{td}}]$，加密后的特征向量集 $\boldsymbol{E}=[\boldsymbol{F}'_1,\boldsymbol{F}'_2,\cdots,\boldsymbol{F}'_j]$、未加密的特征向量集 $\boldsymbol{E}=[\boldsymbol{V}'_1,\boldsymbol{V}'_2,\cdots,\boldsymbol{V}'_j]$、陷门中的查询向量集 $\boldsymbol{T}'_{\mathrm{td}}=[\boldsymbol{F}'_{m1},\boldsymbol{F}'_{m2},\cdots,\boldsymbol{F}'_{ms}]$。在已知密文模型中，图像内容是安全的。然而，对于图像来说，图像的特征向量能够泄露图像内容，搜索陷门与密文特征向量的生成过程类似，因此主要讨论图像特征向量的隐私性。

图像向量间的相似度由式(10)度量。敌手 A 将$[\boldsymbol{V}',\boldsymbol{F}']$和$[\boldsymbol{V}'_{\mathrm{td}},\boldsymbol{F}'_{\mathrm{td}}]$代入式(10)，可得

$$(\boldsymbol{F}'^{\mathrm{T}}_{\mathrm{td}}\boldsymbol{F}'=\tau(\|\boldsymbol{V}'\|^2-2\boldsymbol{V}'^{\mathrm{T}}\boldsymbol{V}'_{\mathrm{td}}+\boldsymbol{\omega}\boldsymbol{\vartheta}^{\mathrm{T}}) \tag{12}$$

式中的 τ 和 $\boldsymbol{\omega}\boldsymbol{\vartheta}^{\mathrm{T}}$ 均未知，密文特征向量无法泄露明文特征向量的相关信息。即便明文特征向量集 $\boldsymbol{E}$ 已知，在多项式时间内，敌手也无法暴力破解密文特征向量。故在已知背景攻击模型下，多项式时间敌

手无法得到加密数据集和搜索陷门的明文。

5　性能分析与比较

5.1　理论分析对比

表 1 给出了文献[7]、文献[11]和本文 RAESSG 共 3 个对比方案的理论计算开销分析，对方案中代价较高的密码运算操作进行比较。表中：Q_G 为循环群指数运算量；Q_{Z_p} 为有限域指数运算量；Q_E 为双线性对运算量；Q_H 为哈希运算量；Q_M 为矩阵运算量；w 为图像拥有者数；n 为图像数；d 为明文特征向量维数；d' 为密文特征向量维数；γ 为哈希函数数量；σ 为哈希表数量；$|T|$ 为访问树结构的叶子节点数。

由表 1 可看出，在密钥生成阶段，RAESSG 方案的密钥生成时间随着访问树结构的叶子节点数的增加而增加，为了加强内积计算的安全性，在安全近邻算法中在密文特征向量中添加了冗余项，特征向量维数的变多使得密钥生成时间也跟着增加。在索引生成阶段，3 个方案的索引生成时间都与图像总数呈正相关。在搜索陷门生成阶段，3 个方案的陷门生成时间主要受特征向量维数、哈希计算和矩阵相乘的影响，将其考虑为图片数量为 1 的索引建立，故省略。在图片检索阶段，图像数和特征向量维数对计算开销的影响很大。RAESSG 方案的密钥生成的计算开销较大，主要是因为属性私钥的生成，但以此为代价可以换来对检索结果的细粒度访问控制；文献[7]的索引建立很快，但只是利用图像加密密钥对特征向量进行加密，并未建立真正意义上的隐私，故在图像检索阶段，需要根据搜索陷门对所有密文图像的密文特征向量进行内积计算，然后两两比对得到最后的结果，因此当图像数量较多时，计算开销很大；群指数运算对文献[11]各阶段的计算开销都有影响，而 RAESSG 只有两个阶段的开销与其相关。

表 1　计算开销的理论分析对比

Table 1　Theoretical comparison of computational expenses

方案来源	计算开销		
	密钥生成	索引生成	图像检索
文献[7]	$2dwQ_{Z_p}$	$2nQ_M$	$2d^2n^2$
文献[11]	$(2+3d)wQ_{Z_p}$	$ndpQ_{Z_p}$	$2nQ_{Z_p}$
本文	$(2\|T\|)wQ_G+wQ_{Z_p}+2d'^2$	$n(\gamma\sigma Q_H+2Q_M)$	$2nd'^2+Q_E$

5.2　实验性能分析对比

实验环境为 RedniBook14、CPU Intel(R) Core (TM) i5-10210U @1.60 GHz、Win10 系统的笔记本电脑，实验工具为 Java1.8 和 python3.7。实验从 Corel 10K 数据集中随机选取 20 类图像，并添加了电力巡检工人类型图像1 000张、电网安全事故现场类型图像 600 张、绝缘子类型图像1 000张、电网合同类型图像 400 张，共计5 000张图像。设定有 4 个哈希表，每个哈希表需利用 4 个哈希函数，降维后的融合特征维度为 120，访问树 Γ 的叶子节点数为 10。

密计算开销的仿真结果如表 2 所示。在图像检索阶段，对比的是 3 个方案在5 000张图像中进行检索的平均时间。文献[7]的密文生成时间和索引生成时间较短，分别约为 60 和 70 ms，但采用遍历内积计算并比较图像密文特征的方式，导致图像检索平均时间太长，约为 358 s；文献[11]采取了聚类方法，文图像检索时间在 3 个对比方案中最短，约为 56 ms，但索引生成时间较长，约为 235 s；由于 RAESSG 需要对用户生成属性私钥，故密钥生成时间为 615 ms，相较于文献[7]和文献[11]高出许多，但其为一次性过程，是可以接受的。虽然 RAESSG 的索引生成时间和图像检索时间并非是最短的，但与表 2 中的最短时间相差不大，远远小于表 2 中的最长时间。在表 2 的实验中，3 个方案采用的特征均为 4 个特征融合后的特征。需要说明的是，由于特征提取和融合时间较长，为方便比较，表 2 中的索引生成时间未包括这些时间。此时，PCA 降维后的融合特征维数为 120 维，花费时间为 496.5 s。

表 2　计算开销的仿真对比

Table 2　Simulation comparison of computational cost

方案来源	时间/ms		
	密钥生成	索引生成	图像检索
文献[7]	60	70	358 000
文献[11]	46	235 000	56
本文	615	76	246

密文图像检索时间如图 4 所示。可以看出，本文利用基于键值对的方法构建安全索引，即便待检索图像的数量不断增加，检索效率依旧维持在同一水平，平均测试检索时间为 246 ms，局部敏感哈希

函数有效缩短了检索时间。

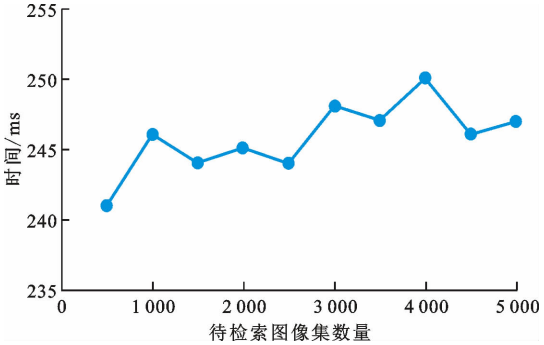

图 4 密文图像检索时间

Fig. 4 Encrypted image retrieval time comparison

设定访问树中叶子节点数为 10，在用户属性数为 2、4、6、8、10 的情况下，分别对待检索图像集建立安全索引，得到平均生成时间为 76 ms。用户属性数对安全索引构建与密钥生成的影响如图 5 所示。可以看出：用户的属性数并不会影响安全索引的生成速度；由于搜索陷门的生成相当于图片为 1 的安全索引构建，所以用户属性数也不会影响搜索陷门的构建。用户属性私钥生成时间随属性数的增加而增加，但增强了安全性。而且，对于用户个体而言，该过程无需重复进行。

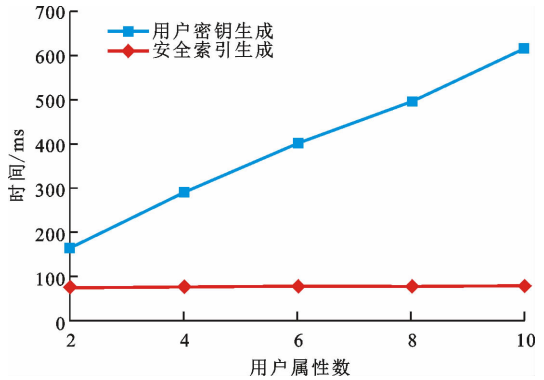

图 5 用户属性数对安全索引构建与密钥生成的影响

Fig. 5 Influence of number of user attributes on generation of security index and attribute private key

利用公式 $P_r = s'/s$ 计算各个方案的检索精确率，其中，s 表示检索返回的图像数，s' 表示与查询图像相同种类的检索返回图像数。云服务器 CS 在整个数据集中进行检索，返回检索结果数为 35，3 个方案检索精确率的对比结果如表 3 所示。可以看出：文献[7]以时间开销为代价换来了较高的检索精确率；文献[11]简化了欧氏距离计算方法，通过放弃检索精度以得到较高的检索效率，两个方案的实用性都受到严重影响；当对图片集进行特征提取时，只对 3 种传统特征进行处理时，即为方案 RAESSG—CNN 特征，此时精确率为 82.9%，与文献[7]的准确率相当；将 CNN 特征纳入考虑时，文献[7]和文献[11]的准确率也都有一定的提高，分别提高 8.5%和 11.4%；本文方案 RAESSG 的检索精确率弱于加强后的文献[7]，但 CNN 特征使检索精确率提高了 5.7%，这说明了 CNN 深度特征与传统特征结合的有效性。

表 3 密文图像检索精确率对比

Table 3 Comparison of ciphertext image retrieval accuracy

方案来源	检索精确率/%
文献[7]	82.9
文献[7]+CNN 特征	82.9+8.5=91.4
文献[11]	60.0
文献[11]+CNN 特征	60.0+11.4=71.4
本文-CNN 特征	88.6−5.7=82.9
本文	88.6

实验中对数据集分别提取了 1 024 维 CNN 特征、3 780 维 HOG 特征、120 维 BOW 特征及 120 维 ColorSpace 特征，根据 RAESSG 方案分别对经 PCA 降维后的 80、100、120、140 维的融合特征进行实验，实验结果如表 4 所示。可以看出：当降维后融合特征维度处在一个合适的范围内，特征维度越高，检索时间越长，但是由于索引的结构，检索时间十分相近；当特征维度控制在 120 维时，检索精确率最高；融合特征的特征维度对检索精确率的影响较大，维度太低或太高，会造成图像特征信息的缺失或冗余，进而降低密文图像的检索精确率。本文方案的示例检索结果如图 6 所示。

表 4 不同融合特征降维维度的检索性能对比

Table 4 Comparison of retrieval performances for different reduced dimensionality of fusion features

融合特征维度	检索时间/ms	检索精确率/%
80	240	74.3
100	244	80.0
120	246	88.6
140	253	85.7

6 结 论

本文针对智能电网环境中敏感图像数据面临的隐私安全、密文图像检索效率及检索精度较低的问题，提出一种基于多特征融合的智能电网敏感图像

(a)检索原图　　(b)检索的相似图像

图 6　密文图像检索结果

Fig. 6 Ciphertext image retrieval results

可检索属性基加密方案，将对称可搜索加密与密文策略属性基加密机制相结合，在确保图像数据的隐私的前提下实现对敏感图像的高效安全地检索与精确地访问。本文得出的主要结论如下。

(1)安全性分析表明，本文方案可抵抗已知密文攻击、已知背景攻击和抗串谋攻击。理论性能分析和在电网数据集的实际性能测试表明，本文方案能对敏感图像进行高效地安全检索，其密钥生成时间、索引生成时间及图像检索时间皆为毫秒级，检索精度为 88.6%。在本文的 3 个对比方案中，所提方案在检索效率和检索精度中取得了最好的平衡，在智能电网场景中进行实际应用是可行有效的。

(2)利用局部敏感哈希函数建立基于键值对的查找表方式，采用添加了冗余项的安全最近邻算法对特征向量进行加密，可以在增强安全性的同时有效地缩短检索时间，且即便待检索图像库的数量不断增加，检索效率依旧维持在同一水平，其检索时间大约为 246 ms。

(3)深度图像特征与传统图像特征的结合，增强了最终特征向量的表达能力，进而提高密文图像的检索精度。深度图像特征的融合将文中 3 种对比方案的检索精度提高了 5.7%～11.4%。

(4)本文的密文图像搜索模式较为单一，只允许以图搜图，会给用户体验造成一定的局限性。下一步研究工作将会针对方案的关键字查询图像功能或多模态查询图像功能进行展开。

参考文献：

[1] 冯国平，古明生，吉小恒. 电网非结构化数据管理平台研究与实现 [J]. 南方能源建设，2015，2(S1)：222-225.
FENG Guoping，GU Mingsheng，JI Xiaoheng. Research and implementation of unstructured data management platform for power grid enterprises [J]. Southern Energy Construction，2015，2(S1)：222-225.

[2] 施进发，焦合军，赵群力，等. 公共云环境下的多租户数据隐私研究 [J]. 计算机工程与应用，2016，52(20)：138-144.
SHI Jinfa，JIAO Hejun，ZHAO Qunli，et al. Research on multi-tenant privacy preservation in public cloud environment [J]. Computer Engineering and Applications，2016，52(20)：138-144.

[3] FERREIRA B，RODRIGUES J，LEITÃO J，et al. Practical privacy-preserving content-based retrieval in cloud image repositories [J]. IEEE Transactions on Cloud Computing，2019，7(3)：784-798.

[4] 温蜜，李婧，殷脂. 智能电网中数据的可搜索加密机制 [J]. 上海电力学院学报，2013，29(6)：513-517，526.

[5] LI Jiangnan，NIU Xiangyu，SUN J S. A practical searchable symmetric encryption scheme for smart grid data [C]// Proceedings of the 2019 IEEE International Conference on Communications (ICC). Piscataway，NJ，USA：IEEE，2019：18852472.

[6] ELTAYIEB N，ELHABOB R，HASSAN A，et al. An efficient attribute-based online/offline searchable encryption and its application in cloud-based reliable smart grid [J]. Journal of Systems Architecture，2019，98：165-172.

[7] 朱旭东，李晖，郭祯. 云计算环境下加密图像检索 [J]. 西安电子科技大学学报，2014，41(2)：151-158.
ZHU Xudong，LI Hui，GUO Zhen. Privacy-preserving query over the encrypted image in cloud computing [J]. Journal of Xidian University，2014，41(2)：151-158.

[8] YUAN Jiawei，YU Shucheng，GUO Linke. SEISA：Secure and efficient encrypted image search with access control [C]// Proceedings of the 2015 IEEE Conference on Computer Communications (INFOCOM). Piscataway，NJ，USA：IEEE，2015：2083-2091.

[9] 李颖莹，马建峰，苗银宾. 基于边缘计算的支持多密钥的加密图像检索 [J]. 通信学报，2020，41(4)：14-26.
LI Yingying，MA Jianfeng，MIAO Yinbin. Encrypted image retrieval in multi-key settings based on edge computing [J]. Journal on Communications，2020，41(4)：14-26.

[10] XIA Zhihua，XIONG N N，VASILAKOS A V，et al. EPCBIR：an efficient and privacy-preserving content-based image retrieval scheme in cloud computing [J]. Information Sciences，2017，387：195-204.

[11] SHEN Meng，CHENG Guohua，ZHU Liehuang，et

al. Content-based multi-source encrypted image retrieval in clouds with privacy preservation [J]. Future Generation Computer Systems, 2020, 109: 621-632.

[12] NAZIR A, ASHRAF R, HAMDANI T, et al. Content based image retrieval system by using HSV color histogram, discrete wavelet transform and edge histogram descriptor [C]// Proceedings of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET). Piscataway, NJ, USA: IEEE, 2018: 17732657.

[13] YANG Ting, SHI Guohua, YANG Na, et al. Combining multiple feature for texture image classification [C]// Proceedings of the 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). Piscataway, NJ, USA: IEEE, 2017: 17615084.

[14] AHMED K T, UMMESAFI S, IQBAL A. Content based image retrieval using image features information fusion [J]. Information Fusion, 2019, 51: 76-99.

[15] HUSAIN S S, BOBER M. REMAP: multi-layer entropy-guided pooling of dense CNN features for image retrieval [J]. IEEE Transactions on Image Processing, 2019, 28(10): 5201-5213.

[16] LIU Peizhong, GUO Jingming, WU Chiyi, et al. Fusion of deep learning and compressed domain features for content-based image retrieval [J]. IEEE Transactions on Image Processing, 2017, 26(12): 5706-5717.

[17] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[18] 陈享, 何亨, 李鹏, 等. 云环境中基于目标检测的密文图像检索方案 [J]. 计算机工程与应用, 2020, 56(11): 75-82.
CHEN Xiang, HE Heng, LI Peng, et al. Ciphertext image retrieval scheme based on target detection in cloud environment [J]. Computer Engineering and Applications, 2020, 56(11): 75-82.

[19] MIAO Yinbin, MA Jianfeng, LIU Ximeng, et al. Attribute-based keyword search over hierarchical data in cloud computing [J]. IEEE Transactions on Services Computing, 2020, 13(6): 985-998.

[20] HUANG Gao, LIU Zhuang, VAN DER MAATEN L, et al. Densely connected convolutional networks [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 2261-2269.

(编辑 陶晴)