

基于隐私保护的数字图像取证外包技术框架研究

冯丙文¹, 翁健¹, 卢伟²

(1. 暨南大学网络空间安全学院, 广东 广州 510632;
2. 中山大学数据科学与计算机学院, 广东 广州 510006)

摘要: 将数字取证服务外包给云有利于满足公众和各类机构的鉴别服务需求, 也是数字取证技术发展的必然。然而这也带来了隐私泄露的风险。在已有的相关领域的研究基础上, 给出了具有隐私保护能力的数字图像取证云平台的系统框架, 并对 Copy-move 检测、图像来源鉴别和基于自然图像统计特性的篡改检测 3 种数字图像取证技术给出了隐私保护实现框架, 进而探讨了隐私保护的数字图像取证可能的发展方向。

关键词: 隐私保护; 数字图像取证; 图像处理; 云计算

中图分类号: TP302

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2016.00085

Research on outsourcing of digital image forensics based on privacy preserving

FENG Bing-wen¹, WENG Jian¹, LU Wei²

(1. School of Information Science and Technology, Jinan University, Guangzhou 510632, China;
2. School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510006, China)

Abstract: The rapid growth of digital image forensics calls for outsourcing digital forensics services to the cloud. It meets the requirement of image content authentication from both the public and organizations. However, it also brings the privacy concerns over the abuse of sensitive information. Based on the research in related fields, a digital forensics cloud platform with privacy preservation was presented. Furthermore, privacy-preserving outsourcing of three types of digital forensics methods, including copy-move detection, source identification and natural image statistic-based tamper detection, were designed. Various future research directions were also discussed.

Key words: privacy preservation, digital image forensics, image process, cloud computing

1 引言

在数字图像的广泛使用和传播中, 内容的真实性一直是一个严重的安全问题。蓄意篡改或伪装的数字图像作品一旦流入正式媒体、科学发现、保险和法庭证物等场景, 都可能会对个人、社会, 乃至国家产生严重的影响。数字图像取证技术(digital image forensics)便是一类鉴别数字媒体来源和真实性的技术。其通过对图像数据特征的分析来判断内容的原始性和真实性, 并通过分析

数字信号处理残留的信息追溯图像数据的来源和可能经历的操作^[1,2]。近年来, 随着图像编辑技术和数字取证技术的发展, 数字图像的篡改越来越难以辨认, 其鉴别所需要的专业知识也越来越复杂。单独个人或机构都难以再掌握取证分析所需要的大量技术和样本, 难以承担搭建数字取证系统所需的软硬件资源。因此, 借助于同样迅速发展的云计算, 将数字取证任务外包给第三方云平台无疑是一个经济有效的解决方案。

然而, 将取证任务外包给云也带来了隐私泄

收稿日期: 2016-06-02; 修回日期: 2016-07-30。通信作者: 冯丙文, bingwfeng@gmail.com

露的风险。用户需要取证的图像数据经常包含敏感内容,而且用户和图像内容也经常存在利益关系。这些信息通常用户不希望泄露给数字取证云服务提供方。另一方面,云平台的计算资源共享更加剧了数据外泄的安全隐患。这些被取证的图像数据如果没有适当的保护,很可能给其拥有者带来危害。近年来,层出不穷的信息泄露事件也充分说明网络环境下数据安全的严峻性^[3]。因此,如何在各类数字取证外包服务中保护图像数据的安全和隐私,是决定数字图像取证外包能否实际应用的一个重要决定因素。

近年来,隐私保护技术吸引了大量学者的关注。隐私保护的图像处理技术^[4-10]、图像识别和检索技术^[11-16]、机器学习技术^[17-19]更是其中的热门方向。数字图像取证本质上也是对图像的分析,因此其核心技术可以借鉴已有的隐私保护方案。如隐私保护的图像局部特征提取^[4-6]、频率域变换^[7]、图像去噪^[8,9],以及图像统计特征生成^[10]等,这些图像处理技术常被应用于各类数字取证方案,用来生成图像的描述特征。而特征分析常用的机器学习工具,如支持向量机^[17]、随机森林^[18]、神经网络^[19]等也可以借鉴已有的隐私保护实现。然而,不同于图像识别^[11]或图像检索^[12],数字取证需要分析被掩盖于图像内容之下的微弱信号,再简单的数字取证方案都需要结合多种图像处理技术将能够反应图像篡改的特征提取出来。这增加了隐私保护实现的难度,方案的复杂度也更加难以控制。现在仍然鲜有文献描述如何实现隐私保护的数字取证外包技术。因此本文从常用的几类数字取证技术着手,探讨其外包实现的技术框架。本文首先搭建了一个具有隐私保护能力的数字图像取证云平台的系统框架,给出其应实现的功能和应满足的安全需求。该云平台利用多种数字取证技术对图像的完整性、真实性、来源做出鉴定,并定位可疑的篡改区域。同时,不会泄露待取证图像的内容和取证的结果。随后本文探讨了云平台使用的几种主要的数字取证技术,即 Copy-move 检测、图像来源鉴别和基于自然图像统计特性的篡改检测的隐私保护实现,并分析可选的核心技术。最后对这些技术进行讨论并指出可能的发展方向。

2 数字图像取证云平台系统框架

在数字图像取证服务系统云平台中,核心鉴别技术由云计算服务方提供,并部署到云计算平台上。取证系统一方面利用存储在各个云平台中的海量图像数据作为学习样本,优化分析工具;另一方面为各类用户提供各类图像取证服务。因此,取证系统需要和 2 个外部实体,即用户和其他云存储平台进行交互。系统的数据流如图 1 所示。

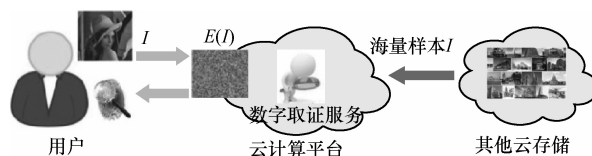


图1 数字图像取证云平台数据流

将数字取证外包到云上可以带来很多好处。首先,外包到云上后,数字取证技术可以充分利用云和其他用户提供的海量样本构造更为复杂、接近实际的分析模型,优化分析工具。其次,服务提供方可以根据不同分析技术的差异进行自动伸缩,通过增、删、改云计算平台支持的分析工具保持系统的先进性,而这些更新不需要每个用户再做相应的调整。最后,将服务外包给云后,用户可以直接通过云来定制数字取证服务,不再需要购置大量的软硬件资源来部署数字取证系统,这极大减少了用户的开支。数字取证云平台的安全和隐私保护要求云计算平台不能获取用户图像数据的相关信息,同时用户也不能获取云计算平台所使用分析工具的相关信息。

取证的图像可能含有大量的敏感信息,因此,在上传前,需要由用户使用隐私保护技术进行处理。隐私保护的实现方式有 2 种:一种是由用户生成特征或完成部分特征运算,再将结果加密后上传给云^[12-14];另一种是用户直接将加密后的图像上传给云计算平台,由云来计算特征并进行鉴别^[11,15]。显然第 2 种方法可以使数字取证系统更灵活地执行特征提取和分析计算,且不会泄露取证系统使用的分析技术,同时用户也不用承担大量的计算或对系统的更新做出改变。这种方式下,图像数据在本地加密后再上传到云计算平台。云

计算平台只能对用户上传的密文图像进行鉴别,并将结果以密文形式返回,在鉴别过程中云计算平台既无法获知原图像内容,也无法得到鉴别中产生的和原图像相关的数据以及鉴别结果。用户最后通过对应的密钥,从云计算平台返回的密文中提取出鉴别结果。另一个交互场景,云计算平台从其他云存储中获取样本并学习、训练,由于不涉及用户,因此可以在明文环境下进行,本文也不予讨论。

一个实用性较好的数字图像鉴别需要从文件头、采集设备、Copy-move 篡改、拼接篡改、图像处理痕迹等角度出发,得出多个分析结果,并利用 Dempster-Shafer 证据理论^[20]等技术融合判决结果。整个检测框架如图 2 所示。下面对其中几类取证方案的隐私保护实现进行设计。由于 Copy-move 检测,图像来源鉴别和基于自然图像统计特性的篡改检测常被结合应用以提高取证的精度^[21,22],而且这 3 类方案使用的图像处理、机器学习等技术涵盖了大部分数字图像取证方案使用的技术。因此这里以这 3 类取证方案为例,其他取证方案可以由类似方式设计。文中使用 $\langle x \rangle$ 表示 x 的同态加密的密文。用 $\oplus$ 和 $\otimes$ 表示同态的 $+$ 和 $\times$ 操作。

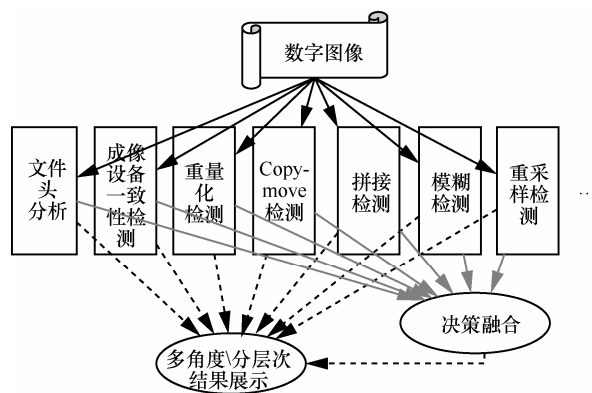


图2 数字图像取证云平台数据流

3 隐私保护的 Copy-move 检测

为了掩盖图像中某些物体,或强调某些内容,篡改者可能将图像的部分区域从一个位置复制到其他位置,这种篡改手段被称为 Copy-move^[23]。图 3 给出了一个 Copy-move 篡改图像实例。由于图像中的复制区域和相对应的粘贴区域基本一

致,可以通过寻找图像中存在的相似区域来检测图像伪造痕迹。一个较好的实现手段是使用特征点匹配来寻找图像内的近似区域^[23~25]。这类方法既可以保证执行效率,又可以对复制区域可能经历的几何变形、图像处理具有较好的顽健性。利用特征点检测图 3(a)的检测结果如图 3(c)所示。



(a) Copy-move 篡改图像



(b) 篡改区域



(c) 利用特征点检测结果

图3 Copy-move 检测实例

为了便于隐私保护的实现,将 Copy-move 检测技术简化为 3 个步骤,如图 4 所示。第 1 步是从用户传来的加密图像中提取特征点;第 2 步对这些特征点进行匹配;第 3 步将能够配对的且满足一定密集度的特征点所覆盖的区域作为疑似篡改区域。

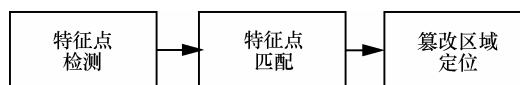


图4 隐私保护的 Copy-move 检测技术框架

Copy-move 检测中最常使用的特征点是 SIFT (scale-invariant feature transform) 特征点^[23]。SIFT 也几乎是隐私保护实现研究最多的特征点。Hsu 等^[26]首次研究隐私保护的 SIFT 实现, 利用加同态加密, 即 Paillier 密码^[27]来实现 SIFT 计算。加密域下的比较是 SIFT 中一个关键的技术, 是 Hsu 等通过将 Difference-of-Gaussian (DoG) 系数的取值范围量划分为几个子阈值, 并判断这些子阈值和系数的密文是否相同来确定每个系数的近似取值。然而该方法需要多传递大量的信息, 同时图像内容也可能被服务器通过比较算法大体猜到。由于 SIFT 的位置本身也会透露出图像的内容, Qin 等^[4,6]进一步对 SIFT 计算中 SIFT 位置的信息进行隐藏, 其计算外包给了 3 个独立的云平台: 2 个计算方和一个比较方, 比较方在返回特征点位置时会增加一些干扰点, 以隐藏真实特征点的位置。Hu 等^[5]进一步改进方案, 仅使用 2 个独立云平台, 通过两方交互协议实现密文下的乘法运算与比较运算, 并借由这 2 种协议实现了 SIFT 的计算。和之前文献相比, Hu 等实现的 SIFT 流程较为完整, 因此提取的特征点的性质更接近于明文下提取的 SIFT 点。

在 Hu 等^[5]的方案中, 客户端首先将图像明文数据 I 加密为 2 个密文 I_1 和 I_2 , 其满足关系 $I(x, y) = I_1(x, y) - I_2(x, y)$ 。随后将 2 个密文分发给 2 个云。每个云分别利用自己的密文计算加密域下的 DoG 系数。由于 DoG 的计算本身也具有加法同态性, 因此, 明文域下的 DoG 系数可以由 2 个加密域下的 DoG 系数相减得到。在下一步计算 DoG 局部极值点步骤中, 需要 2 个云均得到每个 DoG 系数和其空间或尺度位置相邻的 DoG 系数的大小关系。因此需要一个比较协议来比较 2 个密文的大小。这里利用了环上误差学习 (ring-LWE) 全同态加密^[28]实现 2 个整数的密文比较。其将 2 个整数依二进制比特加密。假设 2 个云分别拥有 $x = (x_1, \dots, x_N) \in Z_p^N$ 和 $y = (y_1, \dots, y_N) \in Z_p^N$, 欲比较 2 个数大小, 可以进行如下计算。

如果 $i = k$, 则

$$\begin{aligned} a_i &= \langle x_i \rangle \otimes (\langle 1 \rangle \oplus \langle -1 \rangle \otimes \langle y_i \rangle) \\ b_i &= \langle y_i \rangle \otimes (\langle 1 \rangle \oplus \langle -1 \rangle \otimes \langle x_i \rangle) \end{aligned} \quad (1)$$

如果 $i < k$, 则

$$\begin{aligned} a_i &= (\langle 1 \rangle \oplus \langle -1 \rangle \otimes b_{i+1}) \otimes \\ &\quad [(a_{i+1} \otimes (\langle 1 \rangle \oplus \langle -1 \rangle \otimes a_{i+1}) \otimes \langle x_i \rangle \otimes (\langle 1 \rangle \oplus \langle -1 \rangle \otimes y_i))], \\ b_i &= (\langle 1 \rangle \oplus \langle -1 \rangle \otimes a_{i+1}) \otimes \\ &\quad (b_{i+1} \otimes (\langle 1 \rangle \oplus \langle -1 \rangle \otimes b_{i+1}) \otimes \langle y_i \rangle \otimes (\langle 1 \rangle \oplus \langle -1 \rangle \otimes x_i)) \end{aligned} \quad (2)$$

可以看到, 如果 a_i 解密能得到 1, 那么 $x > y$ 。接下来, 经过协作完成主方向分配, 局部直方图生成等步骤后, 2 个云分别得到加密域下的 SIFT 特征点 V_1 和 V_2 。客户端可以由这 2 个密文相减得到 SIFT 特征点的明文: $V = V_1 - V_2$ 。

拥有特征点后, 下一步是对每个特征点进行匹配, Copy-move 检测中常用的匹配方法是 Best-bin-first, 即和特征点 X 最接近的特征点为

$$\arg \min_{\tilde{X}} \|X - \tilde{X}\| \quad (3)$$

其中, $\| \cdot \|$ 为欧式距离。密文下寻找最近邻点有很多种方法, Lu 等^[29]对图像特征首先使用比特平面随机化, 然后随机投影到低维空间。该加密方式下, 只要密钥相同, 特征间距离便可以近似保持。Wang 等^[13]考虑用户、数据库、云服务器 3 个实体间的隐私保护相似度量, 利用相似矩阵的秩相同的特点, 提出了可以抗云服务器和用户合谋的保序加密。Shantanu 等^[30]给出了可计算隐私安全 (computational privacy) 和信息论隐私安全 (information-theoretic privacy) 下欧式距离的计算和最小欧式距离的搜索方案。并介绍了几种常用于加密域比较的随机化编码方法: Johnson-Lindenstrauss 编码^[31]、局部敏感散列^[32]和基于均匀一标量量化的安全编码^[33]。这里扩展 Hu 的方案^[5]以进一步实现 2 个 SIFT 点的比较。

假设要比较的特征点为 $V(x, y)$ 和 $V(x', y')$ 。2 个云分别计算

$$\Delta V_1 = V_1(x, y) - V_1(x', y') \quad (4)$$

$$\Delta V_2 = V_2(x, y) - V_2(x', y') \quad (5)$$

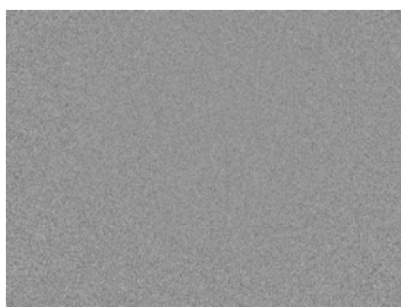
并比较两者大小。可以看到, 如果 $\Delta V_1 \geq \Delta V_2$, 那么 $[V_1(x, y) - V_2(x, y)] - [V_1(x', y') - V_2(x', y')] = V(x, y) - V(x', y') \geq 0$ 。这里需要注意的是, 由于比较协议仅能对整数进行, 因此 SIFT 特征向量生成时需要使用整数近似。

寻找到最佳特征点后, 需要根据这些特征点的位置关系聚类得到疑似篡改区域。由于在 Hu 的方案中 2 个云均能知道最匹配特征点的位置,

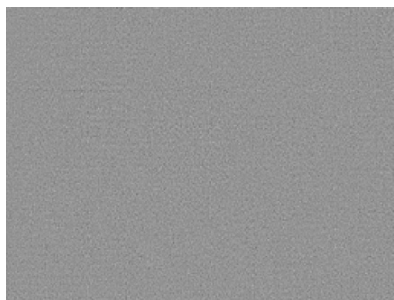
因此,这一步可以很容易在明文环境下实现。至此简单地实现了 Copy-move 检测的隐私保护外包。然而需要注意的是,该实现方法会向云服务器泄露疑似篡改区域的位置。另一方面,性能较好的 Copy-move 检测技术通常利用图像自适应分割、几何变换估计、迭代优化等技术^[34,35]。其他特征点,如 SURF^[36]等,也可能达到优于 SIFT 的检测结果。因此,真正实用的隐私保护的 Copy-move 检测还需要进一步的研究。

4 隐私保护的图像来源鉴别

受工作原理和物理特性的影响,数码相机的镜头、成像传感器和数字信号后处理,如去马赛克、Gamma 矫正、色彩矫正、白平衡等,会在成像过程中留下特有的设备痕迹和噪音,称其为设备指纹(fingerprint)^[37]。同一设备所获取的所有图像均带有相似的指纹,这些指纹只与成像设备使用的硬件元器件有关,与图像内容无关。因此可以用设备指纹来辨别图像来源,或由其是否一致来判断单幅图像是否用多个图像拼接生成。常用的设备指纹是光敏材料的光子响应非均匀性引起的光子响应非均匀性噪音(PRNU, photo-response non-uniformity noise)^[37]。图5给出了2种不同相机的PRNU模式噪音^[38]。



(a) 一部 Apple iPhone 3G 手机的 PRNU



(b) 一部 Samsung i760 手机的 PRNU

图5 不同相机的 PRNU 示例

依据文献[37],基于 PRNU 模式噪音的图像来源鉴别流程如图6所示。一副图像的 PRNU 可以由该幅图像减去其去噪图像,即

$$p = I - \text{Denoise}(I) \quad (6)$$

而为了得到稳健的结果,通常使用来于同一相机的多幅图像的平均来得到该相机的模式噪音。图像来源鉴别过程如下。第1步:从图像库中得到所有候选成像设备的 PRNU。这一步可以由数字取证云服务器从其他公共云存储获取图像并计算出来。由于不涉及用户参与,这一步可以在明文环境下进行。第2步:用户将待检测图像加密并传给云服务器,由云服务器提取出 PRNU 模式噪音。在第3步:利用相关性检测寻找具有最相似 PRNU 的设备,并将此作为匹配设备返回给用户。

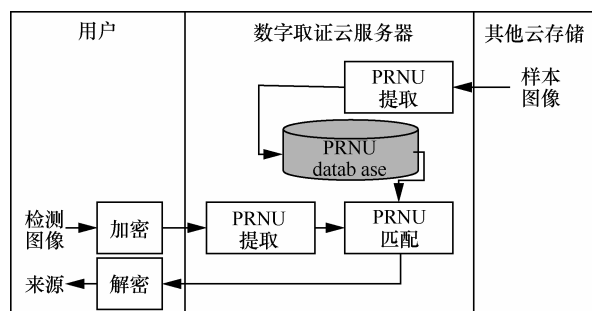


图6 数字图像取证云平台数据流

PRNU 提取的一个重要步骤是加密域下的图像去噪。Saghiaannejadesfahani 等^[8]利用多方计算实现了信息论安全的基于小波的分布式去噪。其基于密码分享(secret share)的同态性质设计了多方线性卷积协议和比较协议。Hu 等^[9]设计了加密域下的非局部均值(non-local means)去噪技术。其使用2个密文:一个是由保序加密生成,以支持加密域下权重的计算;另一个使用 Paillier 加密生成,以完成去噪计算。这里借鉴 Hu 等^[9]的方案。

该方案中,去噪图像 $\tilde{I}$ 的每个像素的值由原图 I 利用式(7)确定。

$$\tilde{I}(i) = \sum_{j \in \Omega} w(i, j) I(i, j) \quad (7)$$

其中,权重 $w(i, j)$ 描述的是图像 I 中以 i 为中心的像素块 $N(i)$ 和以 j 为中心的像素块 $N(j)$ 的相似程度,由式(8)计算

$$w(i, j) = \frac{\exp\left(-\frac{\|N(i) - N(j)\|^2}{h^2}\right)}{\sum_{j' \in \Omega} \exp\left(-\frac{\|N(i) - N(j')\|^2}{h^2}\right)} \quad (8)$$

可以看到, 加密域下图像去噪算法需要计算任意 2 个图像块密文的欧式距离。利用 Johnson-Lindenstrauss 编码^[31], 可以计算每个图像块的保序密文

$$E_j(i) = PN_i + \Delta_i, \Delta_i \sim N(0, \zeta) \quad (9)$$

并用其求得 $w(i, j)$ 。最后, Paillier 加密域下的去噪图像为

$$\langle \tilde{I}_c(i) \rangle = \bigoplus_{j \in \Omega} \langle I(i) \rangle^{w(i, j)} \quad (10)$$

待检测图像的 PRNU (即 $p_c = [p_{c,1}, \dots, p_{c,N}]^T$) 和最匹配设备的 PRNU (即 $p_j = [p_{j,1}, \dots, p_{j,N}]^T$) 可以由式(11)求得

$$\begin{aligned} j &= \arg \max_{j'} \frac{(p_c - \bar{p}_c)^T (p_{j'} - \bar{p}_{j'})}{\|p_c - \bar{p}_c\| \|p_{j'} - \bar{p}_{j'}\|} \\ &= \arg \max_{j'} \frac{(p_c - \bar{p}_c)^T (p_{j'} - \bar{p}_{j'})}{\|p_{j'} - \bar{p}_{j'}\|} \end{aligned} \quad (11)$$

通过去噪图像的密文 $\langle \tilde{I}_c \rangle$, 云服务器可以得到待检测图像的 PRNU 密文 $\langle p_c \rangle = \langle I \rangle \oplus \langle \tilde{I}_c \rangle^{-1}$, 而均值密文 $\langle \bar{p}_c \rangle = \left\langle \frac{1}{N} \sum_i p_{c,i} \right\rangle$ 也容易求得。云服务器上已存有所有候选设备 PRNU 的明文, 因此 2 个 PRNU 在加密域下的相关性为

$$\langle D(p_c, p_i) \rangle = \bigoplus_{1 \leq j \leq N} \left(\langle p_{c,j} \rangle \oplus \langle \bar{p}_c \rangle^{-1} \right)^{\frac{(p_{i,j} - \bar{p}_i)}{\|p_i - \bar{p}_i\|}} \quad (12)$$

而比较任意 2 个相关性的大小可以借助于第 3 节介绍的加密域比较方法。由于使用了 Paillier 加密, 可以使用百万富翁协议 (millionaire protocol)^[11,39] 实现比较。

通过以上步骤, 可以获得最匹配设备序号的密文。该密文可以由用户使用对应私钥解密。需要注意的是, 这种隐私保护的图像来源鉴别实现方法计算量较大, 而且在比较密文大小的时候需要额外的用户和云服务器通信。在图像去噪方法

的选用上, 还与 PRNU 提取实际使用的去噪方法有一定差距。

5 隐私保护的基于自然图像统计特性的篡改检测

本文最后给出一个一般性拼接篡改检测技术的隐私保护外包实现。由于图像篡改会改变原自然图像的统计特征, 因此可以利用一个二分类器, 通过分析大量的已认证自然图像和篡改图像来确定这些特征的差异。称这类方法为基于自然图像统计特性的检测技术。

已有学者给出了多种可应用于篡改检测的自然图像统计模型^[40-42]。由于 He 等^[40]的方法检测效果较好, 且计算量较少, 本文选用它来实现隐私保护的篡改检测方案, 其考虑原图像的像素相关性和小波系数的相邻位置的系数相关性。流程如图 7 所示。在特征提取阶段, 对给定的数字图像使用三级小波分解, 然后计算在空域和小波不同尺度下, 或不同子带下的差分矩阵, 并从中提取二阶共生矩阵作为特征向量。该特征向量用于分类器的训练和检测。由于分类器的训练由云

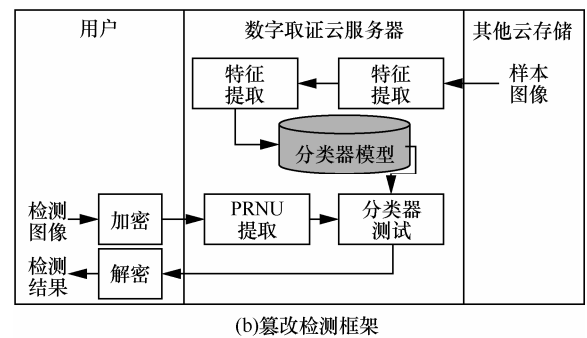
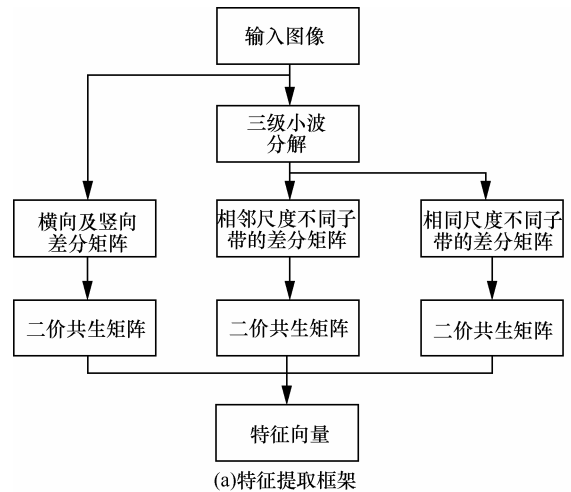


图 7 隐私保护的基于自然图像统计特性的篡改检测技术框架

服务器使用其他云存储的图像实现，因此可以在明文下进行。而利用用户提供的待检测图像计算特征，并用训练好的分类器检测图像则必须在加密域下进行以保护隐私。

特征提取阶段2个重要的技术是小波变换和共生矩阵生成。学者已经给了加密域下的各类图像频率域变换，如DFT^[43]、DCT^[44]、DWT^[7]。由于频域变换本质上等同于图像的线性卷积操作，而标准图像变换的卷积核并不需要隐私保护，因此这类变换使用仅具有加法的同态加密便可以实现。在加密域的小波变换中，近似（或细节）系数可以由式(13)计算^[7]

$$\begin{aligned} \langle A_j(k) \rangle &= \left\langle \sum_{l \in \mathbb{Z}} G_d(2k-l) A_{j-1}(l) \right\rangle \\ &= \bigoplus_{1 \leq j \leq N} \langle A_{j-1}(l) \rangle^{G_d(2k-l)} \end{aligned} \quad (13)$$

其中， $G_d(2k-l)$ 为低通（或高通）滤波器系数。需要注意的是计算中所有的浮点数值都要放大取整。接下来需要计算各类差分图像。以空域图像的水平差分为例，差分图像的密文为

$$\langle D(x, y) \rangle = \langle I(x, y) \rangle \oplus \langle I(x, y+1) \rangle^{-1} \quad (14)$$

密文域下计算图像的统计特征比较困难，可以参考Hu等^[5]的方案，利用多个云服务器共同完成统计特征提取。而Qin等^[10]利用Ring-LWE全同态加密给出了另一种实现方法，其将每一个像素按比特加密，并根据其最高几个比特的值将像素分配到一个桶（Bucket）中，例如，如果要计算具有4个区间0~63、64~127、128~191、192~255的直方图，落入其中的像素的最高2个比特应分别为00,01,10,11。因此，可以计算

$$\begin{aligned} B(x, y, 00) &= (\langle 1 \rangle \oplus \langle -1 \rangle \otimes \langle D(x, y)_7 \rangle) \otimes \\ &(\langle 1 \rangle \oplus \langle -1 \rangle \otimes \langle D(x, y)_6 \rangle) \end{aligned} \quad (15)$$

$$\begin{aligned} B(x, y, 01) &= (\langle 1 \rangle \oplus \langle -1 \rangle \otimes \langle D(x, y)_7 \rangle) \otimes \\ &\langle D(x, y)_6 \rangle \end{aligned} \quad (16)$$

$$\begin{aligned} B(x, y, 00) &= \langle D(x, y)_7 \rangle \otimes \\ &(\langle 1 \rangle \oplus \langle -1 \rangle \otimes \langle D(x, y)_6 \rangle) \end{aligned} \quad (17)$$

$$B(x, y, 00) = \langle D(x, y)_7 \rangle \otimes \langle D(x, y)_6 \rangle \quad (18)$$

其中， $D(x, y)_i$ 是像素 $D(x, y)$ 的第 i 个比特。当且

仅当像素的最高两位为 j 时， $B(x, y, j) = 1$ ，否则， $B(x, y, j) = 0$ 。二阶共生矩阵可以借此计算。以水平向右方向的共生矩阵为例，其加密域计算公式为

$$\begin{aligned} M^{\rightarrow}(a, b) &= \sum_{x, y} \#(D(x, y) = a, D(x, y+1) = b) \\ &= \bigoplus_{x, y} (B(x, y, a) \otimes B(x, y+1, b)) \end{aligned} \quad (19)$$

通过计算8个方向的共生矩阵，即 $M^{\rightarrow}$ 、 $M^{\leftarrow}$ 、 $M^{\downarrow}$ 、 $M^{\uparrow}$ 、 $M^{\searrow}$ 、 $M^{\swarrow}$ 和 $M^{\nearrow}$ ，可以得到图像特征的密文形式。

最后一步是利用训练好的分类器对特征进行分类。常用的分类器为支持向量机（SVM, support vector machine）。具备隐私保护能力的SVM已有一些相关成果。Mangasarian等^[45,46]将完全随机核引入到1-范数隐私保护SVM，所有参与者通过计算自己的随机矩阵来构造各自的局部核矩阵，然后将所有这些参与方的局部核矩阵求和，即可构建全局约简核矩阵，进而得到隐私保护支持向量机模型。Vaidya等^[47]采用同态加密的安全多方计算协议，通过半可信的第三方完成保护敏感信息的内积矩阵计算，并由这个第三方采用SVM算法进行分类和训练。Rahulamathavan等^[17]仅利用加同态加密和一个云服务器实现了非线性SVM的分类过程。由于本方案仅需要对分类过程进行隐私保护，可以借助于Rahulamathavan等^[17]的方案。

对于归一化的待分类特征 t ，基于多项式核函数的非线性分类器分类的判决函数为

$$d(t) = \sum_{s \in S} \alpha_s y_s (x_s^T t + 1)^p + b \quad (20)$$

其中， x_s 为第 s 个支持向量， y_s 是它的标签， α_s 为Lagrangian变量。记多项式核函数为 K_s^p ，即

$$K_s^p = (x_s^T t + 1)^p \quad (21)$$

由于隐私保护，云服务器只能获得特征的密文 $\langle t \rangle$ ，而其他值均能得到其明文。可以看到，除了 $\langle K_s^p \rangle$ ，其他运算均为加密域下的加法和数乘。对于 $\langle K_s^p \rangle$ 的幂运算的计算，采用服务器和客户端的交互实现。服务器先计算 $\langle K_s \rangle$ ，然后利用一个随机数盲化 $\langle K_s \rangle^{\zeta}$ ，并将其交由客户端计算幂运算。收到客户端返回结果后，服务器除去盲化部分即可获得密文 $\langle K_s^p \rangle$ 。最后分类器给出的类标签

为 $d(t)$ 的符号

$$f(t) = \text{sign}[d(t)] \in \{-1, 1\} \quad (22)$$

这里同样需要借助于百万富翁协议实现密文 $\langle d(t) \rangle$ 符号的判断。

和第 4 节的问题相同, 这种隐私保护的篡改检测实现方法同样存在计算量较大, 需要额外的用户和云服务器通信的问题。另外, 里面存在大量的整数近似, 会影响方案的性能, 现在已有学者提出有理数和浮点数的同态加密运算方案^[48,49], 这都推动了寻找更好的实现方案的进程。

6 结束语

云计算和大数据的发展给数字图像取证领域带来了新的进步空间。而云平台上数字取证技术的隐私保护也必将成为该领域面临的挑战。本文给出了基于云平台的数字图像取证系统的一般架构, 并探讨了 3 类取证技术, 即 Copy-move 检测, 图像来源鉴别和基于自然图像统计特性的篡改检测如何在满足隐私保护的前提下外包给云。为了满足安全性和性能的要求, 本文参考了现有各类隐私保护的图像处理、图像识别和机器学习实现方案, 利用同态加密、安全多方计算等多种技术, 给出了取证方案的大体框架。这些框架的功能与性能还有待进一步实验的验证。另外, 由于数字图像取证领域的特殊性, 现有的方案还不足以实现具有实用价值的隐私保护取证方案。还有很多取证技术, 如基于图像内容的物理一致性, 需要大量的人工干预, 难以实现隐私保护。虽然隐私保护的安全外包在其他领域已经取得了一定的进展, 但在数字图像取证这个特殊应用上, 还有许多地方值得研究。

参考文献:

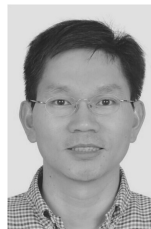
- [1] WANG W, DONG J, TAN T. A survey of passive image tampering detection[C]//International Workshop on Digital Watermarking (IWDW 2009). c2009:308-322.
- [2] VERDOLIVA L. Handbook of digital forensics of multimedia data and devices[J]. IEEE Signal Processing Magazine, 2016, 33(1): 164-165.
- [3] 杨帅峰, 肖俊芳, 江浩, 等. 世界网络安全发展特征与趋势[M]//世界网络安全发展报告(2014~2015). 北京: 社会科学文献出版社, 2015: 21-42.
- [4] QIN Z, YAN J, REN K, et al. Towards efficient privacy-preserving image feature extraction in cloud computing[C]//The 22nd ACM International Conference on Multimedia. c2014:497-506.
- [5] HU S, WANG Q, WANG J, et al. Securing SIFT: privacy-preserving outsourcing computation of feature extractions over encrypted image data[J]. IEEE Transactions Image Process, 2016.
- [6] QIN Z, YAN J, REN K. Private image computation: the case of cloud based privacy-preserving SIFT[C]//IEEE Conference on Computer Communications Workshops (Infocom WKSHPS). c2014: 179-180.
- [7] ZHENG P, HUANG J. Discrete wavelet transform and data expansion reduction in homomorphic encrypted domain[J]. IEEE Transactions on Image Processing, 2013, 22(6): 2455-2468.
- [8] SAGHAIAN-NEJAD-ESFAHANI S M, LUO Y, SEN-CHING S C. Privacy protected image denoising with secret shares[C]//The 19th IEEE International Conference on Image Processing. c2012: 253-256.
- [9] HU X, ZHANG W, HU H, et al. Non-local denoising in encrypted images[C]//The International Conference on Internet of Vehicles. c2014. 386-395.
- [10] QIN Z, YAN J, REN K, et al. Privacy-preserving outsourcing of image global feature detection[C]//2014 IEEE Global Communications Conference. c2014: 710-715.
- [11] SADEGHI A R, SCHNEIDER T, WEHRENBURG I. Efficient privacy-preserving face recognition[C]//The International Conference on Information Security and Cryptology. c2009:229-244.
- [12] XIA Z, ZHU Y, SUN X, et al. Towards privacy-preserving content-based image retrieval in cloud computing[J]. IEEE: Transactions on Cloud Computing, 2015.
- [13] WANG Q, HU S, REN K, et al. Cloudbi: practical privacy-preserving outsourcing of biometric identification in the cloud[C]//European Symposium on Research in Computer Security. c2015: 186-205.
- [14] BARNI M, BIANCHI T, CATALANO D, et al. Privacy-preserving fingerprint authentication[C]//The 12th ACM Workshop on Multimedia and Security. c2010:231-240.
- [15] OSADCHY M, PINKAS B, JARROUS A, et al. Scifi-a system for secure face identification[C]//2010 IEEE Symposium on Security and Privacy. c2010: 239-254.
- [16] PETER A, HARTMANN T, MÜLLER S, et al. Privacy-preserving architecture for forensic image recognition[C]//IEEE International Workshop on Information Forensics and Security (WIFS). c2012: 79-84.
- [17] RAHULAMATHAVAN Y, PHAN R C W, VELURU S, et al. Privacy-preserving multi-class support vector machine for outsourcing the data classification in cloud[J]. IEEE Transactions on Dependable and Secure Computing, 2014, 11(5): 467-479.
- [18] VAIDYA J, SHAFIQ B, FAN W, et al. A random decision tree framework for privacy-preserving data mining[J]. IEEE Transactions on Dependable and Secure Computing, 2014, 11(5): 399-411.
- [19] YUAN J, YU S. Privacy preserving back-propagation neural network learning made practical with cloud computing[J]. IEEE Transactions on Parallel and Distributed Systems, 2014, 25(1): 212-221.
- [20] FONTANI M, BIANCHI T, ROSA D A, et al. A framework for decision fusion in image forensics based on Dempster-Shafer theory of evidence[J]. IEEE Transactions on Information Forensics & Security, 2013, 8(4):593-607.
- [21] COZZOLINO D, GRAGNANIello D, VERDOLIVA L. A novel framework for image forgery localization[J]. Computer Science,

- 2013.
- [22] GABORINI L, BESTAGINI P, MILANI S, et al. Multi-clue image tampering localization[C]//IEEE International Workshop on Information Forensics & Security, IEEE. 2014:125-130.
- [23] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(60): 91-110.
- [24] PAN X, LYU S. Region duplication detection using image feature matching[J]. IEEE Transactions on Information Forensics and Security, 2010, 5(4): 857-867.
- [25] CHRISTLEIN V, RIESS C, JORDAN J, et al. An evaluation of popular copy-move forgery detection approaches[J]. IEEE Transactions on Information Forensics & Security, 2012, 7(6):1841-1854.
- [26] HSU C Y, LU C S, PEI S C. Image feature extraction in encrypted domain with privacy-preserving SIFT[J]. IEEE Transactions on Image Processing, 2012, 21(11): 4593-4607.
- [27] PAILLIER P, POINTCHEVAL D. Efficient public-key cryptosystems provably secure against active adversaries[C]//International Conference on the Theory and Application of Cryptology and Information Security. c1999:165-179.
- [28] BRAKERSKI Z, VAIKUNTANATHAN V. Fully homomorphic encryption from ring-LWE and security for key dependent messages[C]//Annual Cryptology Conference. c2011: 505-524.
- [29] LU W, VARNA A L, SWAMINATHAN A, et al. Secure image retrieval through feature protection[C]//IEEE International Conference on Acoustics. c2009:1533-1536.
- [30] RANE, SHANTANU. Privacy-preserving nearest neighbor methods[J]. IEEE Signal Processing Magazine, 2013, 30(2):18-28.
- [31] JOHNSON W B, LINDENSTRAUSS J. Extensions of Lipschitz mappings into a Hilbert space[J]. Contemporary Mathematics, 1984, 26(1):189-206.
- [32] ANDONI A, INDYK P. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions[J]. Foundations of Computer Science Annual Symposium on, 2006, 51(1):117-122.
- [33] BOUFOUNOS P, RANE S. Secure binary embeddings for privacy preserving nearest neighbors[C]// IEEE International Workshop on Information Forensics and Security (WIFS). c2011:1-6.
- [34] LI J, LI X, YANG B, et al. Segmentation-based image copy-move forgery detection scheme[J]. IEEE Transactions on Information Forensics & Security, 2015, 10(3):507-518.
- [35] PUN C M, YUAN X C, BI X L. Image forgery detection using adaptive over-segmentation and feature points matching[J]. IEEE Transactions on Information Forensics & Security, 2015, 10:1.
- [36] YU B, LI Z, BO C, et al. Surf feature extraction in encrypted domain[C]//IEEE International Conference on Multimedia and Expo. c2014:1-6.
- [37] LUKAS J, FRIDRICH J, GOLJAN M. Digital camera identification from sensor pattern noise[J]. IEEE Transactions on Information Forensics and Security, 2006, 1(2): 205-214.
- [38] STAMM M C, WU M, LIU K J R. Information forensics: an overview of the first decade[J]. Access IEEE, 2013, 1:167-200.
- [39] ATALLAH M J, KERSCHBAUM F, DU W. Secure and private sequence comparisons[C]//The ACM Workshop on Privacy in the Electronic Society(WPES'03). c2003:39-44.
- [40] HE Z, LU W, SUN W, et al. Digital image splicing detection based on Markov features in DCT and DWT domain[J]. Pattern Recognition, 2012, 45(12): 4292-4299.
- [41] NG T T, CHANG S F, SUN Q. Blind detection of photomontage using higher order statistics[C]//IEEE International Symposium on Circuits and Systems. c2004.
- [42] QIU X, LI H, LUO W, et al. A universal image forensic strategy based on steganalytic model[C]//The 2nd ACM Workshop on Information Hiding and Multimedia Security. c2014:165-170.
- [43] BIANCHI T, PIVA A, BARNI M. On the implementation of the discrete Fourier transform in the encrypted domain[J]. IEEE Transactions on Information Forensics & Security, 2009, 4(1):86-97.
- [44] BIANCHI T, PIVA A, BARNI M. Encrypted domain DCT based on homomorphic cryptosystems[J]. Eurasip Journal on Information Security, 2009, 2009(2):1-12.
- [45] YU H, JIANG X, VAIDYA J. Privacy-preserving SVM using nonlinear kernels on horizontally partitioned data[C]//ACM Symposium on Applied Computing. c2006:603-610.
- [46] MANGASARIAN O L, WILD E W. Privacy-preserving classification of horizontally partitioned data via random kernels[C]//International Conference on Data Mining, Las Vegas. c2008.
- [47] VAIDYA J, YU H, JIANG X. Privacy-preserving SVM classification[J]. Knowledge & Information Systems, 2008, 14(2):161-178.
- [48] LIU X, CHOO R, DENG R, et al. Efficient and privacy-preserving outsourced calculation of rational numbers[J]. IEEE Transactions on Dependable & Secure Computing, 2016:1-1.
- [49] LIU X, DENG R, DING W, et al. Privacy-preserving outsourced calculation on floating point numbers[J]. IEEE Transactions on Information Forensics & Security, 2016:1-1.

作者简介：



冯丙文 (1985-), 男, 山东东营人, 博士, 暨南大学讲师, 主要研究方向为多媒体安全与数字取证。



翁健 (1976-), 男, 广东茂名, 博士, 暨南大学教授、博士生导师, 主要研究方向为密码学与信息安全。



卢伟 (1979-), 男, 河南南阳人, 中山大学副教授、硕士生导师, 主要研究方向为多媒体信息安全和数字取证。