社交照片隐私保护机制研究进展

李凤华[1,2], 孙 哲[1,2], 吕梦凡[1,2], 牛 犇[1*]

[1] 中国科学院信息工程研究所 信息安全国家重点实验室 北京 中国 100093

[2] 中国科学院大学 网络空间安全学院 北京 中国 100049

摘要 照片分享服务已经广泛渗透到人们生活的各个角落, 这种新型服务模式及应用给人们带来便利的同时无形中泄露了用户的隐私信息, 而照片分享平台上数量激增的社交照片和多样化的用户需求给隐私保护带来了更大挑战。结合国内外相关最新研究趋势, 对社交网络中的照片隐私保护问题进行研究展望。首先介绍基于全生命周期的社交照片隐私保护模型、照片隐私信息分类、照片隐私保护动机和攻击者模型, 以及所面临的隐私威胁; 其次从社交照片生成与感知、发布与交换、存储与销毁、融合与分析 4 个环节介绍了其隐私保护机制研究现状; 最后展望了该领域的未来研究方向。

关键词 照片隐私保护; 全生命周期; 社交网络; 访问控制; 图像加密

中图法分类号 TP309.2 DOI 号 10.19363/j.cnki.cn10-1380/tn.2018.03.04

Research Progress of Photo Privacy-Preserving Mechanisms in Online Social Network

LI Fenghua[1,2], SUN Zhe[1,2], LV Mengfan[1,2], NIU Ben[1*]

[1] State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China

[2] School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049, China

Abstract The well-developed photo sharing service has infiltrated our daily life, this novel service mode provides us conveniences at the cost of our privacy. This situation becomes more serious with the increasing number of social photos and individuation of user's requirements. In this paper, we study the photo privacy-preserving mechanisms in online social networks based on analyzing the state-of-the-art research trends. Specifically, we first introduce the full life-cycle model of photo privacy, and review the information classification, defense motivation and attack models. We also present the threats and challenges of photo privacy leakage. Then, we make a comprehensive study on this research field in terms of Generation and Perception, Publication and Exchange, Storage and Disposal, Combining Analysis. Finally, we discuss the future research directions on photo privacy preservation.

Key words photo privacy protection; full life-cycle; online social network; access control; image encryption

1 引言

数字拍摄和因特网技术的飞速发展给社交照片的拍摄和分享方式带来了巨大的变革。传统照片受制于胶卷成本和传播媒介, 很难大范围分享。但受惠于高分辨率数码拍摄设备的广泛普及, 以及便利的网络互连服务, 数字照片以其低廉的拍摄和传播成本, 为人们提供了丰富的拍摄乐趣和便利的实时分享体验, 社交照片分享已逐渐成为人们日常生活中的一种习惯。通过社交网络, 用户可以方便地将照片分享给亲人和朋友, 以巩固现有的社交关系; 此外, 以交友为目的的照片展示也成为照片分享的另一个研究热点。

随着数字拍摄和因特网技术的日趋成熟, 照片分享服务提供商(Photo Sharing Service Providers, PSPs)在商业上也取得了巨大成功, 社交网络中分享的照片数量呈现出爆炸式增长。据报道 Facebook 在 2015 年每月有超过 14.9 亿的活跃用户, 平均每天就

通讯作者: 牛犇, 博士, 助理研究员, Email:niuben@iie.ac.cn。

本课题得到国家重点研发计划(No. 2017YFB0802203); 国家自然科学基金-面上基金资助项目(No. 61672515); 国家科技部高技术研究发展计划(“863”计划)基金资助项目(No. 2015AA016007); 国家自然科学基金-广东联合基金资助项目(No. U1401251)资助。

收稿日期: 2016-12-20; 修改日期: 2017-02-17; 定稿日期: 2018-02-05

有 9680 万活跃用户, 每天上传约 3.5 亿张照片①; 与此同时, Instagram 平均每天也有约 6000 万张照片被上传②; 近期一个照片消息服务应用 Snapchat 更是超过了前两者, 其用户每天分享了约 7 亿张照片③。

然而, 在照片分享服务给用户带来便利的同时, 其带来的照片隐私问题也远超用户预料。一项调查表明, 年轻用户在社交网络中分享的照片有 80%并不愿意展现给他们的父母或老师[1]。但他们并没有意识到自己分享的社交照片可能会展示给他们的父母、老师甚至未来的雇主。他们将自己、同学和朋友的社交照片不加限制地公布到社交网络, 一些公司和法律执行机构利用网络上分享的社交照片进行强制调查, 不少公司在招聘过程中已经将检查应聘者的网上照片作为惯例。据微软调查[2], 在美国约有 70%的招聘者会因为在网上找到的信息(包括社交照片)将应聘者婉拒。因此, 有研究建议[3], 用户对社交照片隐私的认知程度仍需提高, 且在多种目标不能同时达到时, 应做出符合自己最大利益的权衡。

虽然隐私敏感的用户可以通过精心设计隐私策略来保护自己的照片隐私, 然而提供隐私保护服务的第三方也并非无懈可击。Facebook、Instagram、iCloud 等主要的照片分享服务商均发生过隐私照片泄露事件。其中最著名的是 fucking 攻击事件④, 攻击者通过逆向工程, 绕过访问控制机制, 直接访问服务器存储的照片。此外, 服务提供商对用户社交照片的利用渐渐偏离了用户上传的初衷, 它们利用日趋成熟的人脸识别技术和积累的大量照片资源, 对用户没有明确表示的社交关系进行推断(如 Facebook 部署的人脸识别功能已经在多个国家引起严重的隐私问题⑤)。在此类技术的基础上, 攻击者结合其他分析技术还可以挖掘出用户的行为习惯、个人偏好等敏感信息, 严重威胁用户隐私。

针对当前社交照片隐私面临的误拍偷拍易于实现、照片受控共享机制不够完善、PSPs 利用照片数据不合理、攻击者融合分析技术愈发成熟等威胁, 近年来研究者们进行了系统化的研究, 提出了大量解决方案, 取得了一定进展。本文在分析已有研究成果基础上, 对照片隐私保护机制进行总结与展望。

本文第 2 节介绍照片隐私保护的背景知识与威胁挑战; 第 3 节从照片隐私全生命周期不同环节的角度介绍其隐私保护机制现状, 并进行分析和比较; 第 4 节讨论照片隐私保护机制未来的发展方向; 最后第 5 节对全文进行总结。

2 照片隐私保护背景知识

2.1 照片隐私信息的全生命周期模型

随着公众对个人隐私的关注不断提高, 隐私保护已成为研究者们关注的热点之一。Li 等提出了隐私计算理论, 归纳了通用环境下的隐私信息全生命周期模型[4]。社交照片隐私作为隐私信息的一种, 也包括生成与感知、发布与交换、存储与销毁、融合与分析等环节, 其全生命周期模型如图 1 所示。

1) 产生与感知环节(Generation and Perception)

该环节中, 照片经由自己(自拍)、朋友(协作)或者其他陌生人(误拍)等不同途径产生并被记录在相应存储空间中, 用户可对存储在本地的社交照片进行分类和管理。

2) 发布与交换环节(Publication and Exchange)

该环节中, 用户将照片上传到社交网络中, 通过向不同的用户展示照片, 获得维系感情、相互认可、评论交流、招募粉丝等社交收益, 并可借此建立有效的线下关系。

3) 存储与销毁环节(Storage and Disposal)

该环节中, 用户的照片被存储在服务提供商的存储设备中, 以实现随时随地都可以通过网络访问到用户照片的目的, 当用户认为照片不应继续发布时, 则选择将照片从网络中删除。

4) 融合与分析环节(Combining Analysis)

该环节作为一个特殊的环节, 主要描述攻击者通过在社交网络中收集用户曾经发布的照片, 或通过攻击手段盗取用户的照片, 并利用机器学习等融合分析技术, 挖掘和还原出用户的隐私信息。

2.2 照片隐私信息分类

在照片隐私信息的全生命周期中, 不同环节产生或向已有照片标记了大量隐私信息。例如, 在产生与识别环节, 照片的图像内容、元数据等隐私信息伴随照片的拍摄而产生, 随后新的隐私信息如标记、评

① B. Insider, “Facebook Users Are Uploading 350 Million New Photos Each Day.”http://www.businessinsider.com/facebook-350-million-photos-each-day-2013-9.

② GigaOM, “Instagram reports 200M users.” http://gigaom.com/2014/03/25/ instagram-reports-200m-users-20b-photos-shared/

③ T. Verge. The new Snapchat brilliantly mixes video and texting. http://www.theverge.com/2014/5/1/5670260/real-talk-the-new-snapchat-makes-texting-fu n-again-video-calls.

④ CNN: Photobucket leaves users exposed. http://www.cnn.com/2012/08/09/ tech/photobucket-privacy-breach

⑤ Facebook Shuts Down Face Recognition APIs After All. http://www.theregister.co.uk/2012/ 07/09/facebook_face_apis_dead.

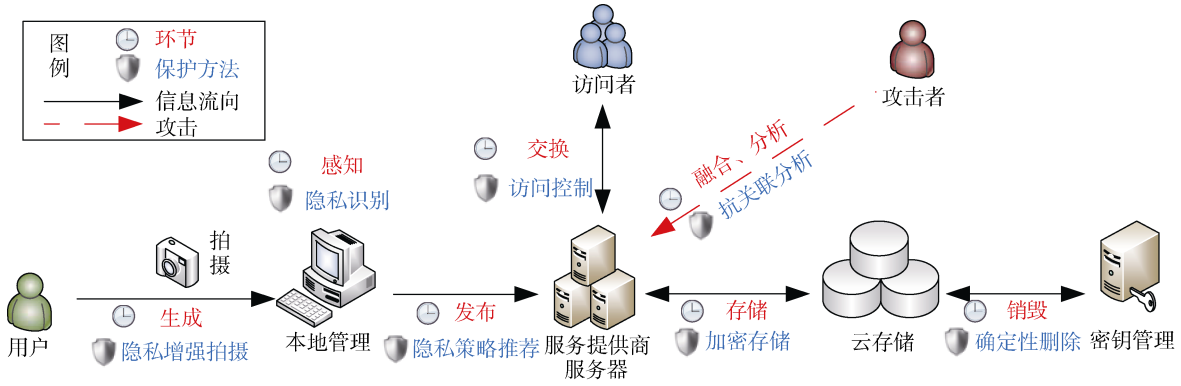


图 1　照片隐私信息全生命周期模型

Figure 1　The Full Lifecycle Model of Photo Privacy

论、归档等, 在发布与交换环节中被不断标记到已有照片中, 这些信息将与原始照片一起历经隐私信息全生命周期的所有环节。目前, 社交网络中的隐私信息大致分为 3 类, 即内容信息、属性信息、关系信息[5]。而照片作为一种特殊的分享媒介, 涵盖了以上 3 类隐私信息。

1) 内容信息(Content Information)

社交网络中照片分享的典型应用场景包括用户在上班路上拍摄了交通拥堵情况, 午休时在某家餐馆拍摄了与美食的合影, 在旅游过程中拍摄了自己与地标的合影, 并欲将这些照片内容上传到社交网络平台上, 甚至给特定用户附上链接, 通过内容信息分享的方式建立线下的人际关系。

然而, 如果照片内容管理不当, 甚至是在不被本人知晓的情况下, 被别人拍摄并作为参与某个特殊活动的证明, 如参与某个私人聚会、示威游行、宗教仪式或者政治集会等, 照片中内容信息的泄露将给用户的隐私带来直接侵害。

2) 属性信息(Profile Information)

目前, 人们在使用数字相机或者带摄像功能的移动设备拍摄照片时, 拍摄设备会自动生成拍摄时的属性信息, 并将其以 EXIF、IPTC 或者 XMP 等通用格式嵌入到照片文件中。常见的属性信息包括拍摄日期、拍摄时间、拍摄位置的经纬度与海拔, 以及焦距、光圈、设备型号等照相机自身相关的参数信息。当照片文件在社交网络中分享时, 用户根据照片的内容给照片添加标签、注释, 并将照片归类到不同的相册中, 从某种角度上都是给照片贴上了不同属性的标签。

照片的各类属性信息通常是以文本形式记录, 并与照片建立映射关系, 相当于给非结构化的照片文件增加了文字索引, 降低了通过图像处理技术挖掘隐私信息的难度, 极大地增加了照片隐私信息泄露的可能。

3) 关系信息(Relationship Information)

照片分享中的关系信息通常是指合影中人与人之间隐含的社会关系, 包括家庭合影、朋友合影、商务合作、集体照等。当用户将包含关系信息的合影向社交网络中的其他用户展示时, 照片成为某种人际关系的物证, 比如自己与某个人认识甚至亲密、自己是某个团体或者圈子的成员等。

相较于内容信息和属性信息, 合影中所包含的关系信息则更为隐蔽, 但承载的信息量反而更大。如果用户对关系信息的保护不够重视, 当用户泄露的照片积累到一定数量后, 通过挖掘照片中所隐含的用户社交关系网络(人与人之间的关系)、所属兴趣团体(人与团体之间的关系)等情报, 攻击者甚至可以将线下用户进行关联。此外, 结合公开数据的用户与隐藏数据的用户之间的关联关系, 攻击者可推测出更多的隐私信息[6]。

2.3　照片隐私保护动机

随着可穿戴拍摄设备的出现与商业化, 自动拍摄的照片从数量上远远超过了原本的拍摄方式, 这加剧了照片在其生命周期中隐私保护的工作量。为了分析出哪些因素使得照片成为一张隐私照片, 研究者们对用户在照片分享过程中的分享行为、决策和需求进行了详细的调研[7]。调研结果表明, 用户限制照片分享范围的主要原因包括隐私信息、印象管理、他人的隐私。

1) 隐私信息(Private Information)

照片中包含的隐私信息是用户设置是否分享照片以及分享范围的一个重要因素。最常见的隐私信息包括照片中有清晰可见的敏感文字、不愿展示的行为以及隐私的拍摄环境等。例如, 用户会禁止含有自己身份证号或银行卡号的照片在网络上公开传播,

限制自己参加游行、集会、宗教仪式或者喝酒、抽烟等行为照片的分享范围。

2) 印象管理(Impression Management)

为了向他人展示自己意欲表达的形象, 印象管理成为很多用户选择是否展示照片的原因之一。印象管理包括分享照片以表现一个正面形象和禁止展示某些负面特征两个方面。例如, 一位用户不愿意展示自己看电影的照片, 因为他不想给别人留下在应当学习的时候“没有学习”的印象; 另外, 一些用户拍摄一些描述电脑工作的照片并分享到朋友圈, 以表现出一种更积极的“工作”印象。

3) 旁观者的隐私(Privacy of Bystanders)

尊重旁观者的隐私已成为大多数有隐私意识用户的共识, 很多用户会因为照片中出现了旁观者而限制照片的传播。例如, 用户拍摄了一张背景中存在路人的照片, 并认为未经许可展示他的信息会侵害该路人的隐私; 此外, 还有很多用户赞同照片中的其他用户也有保护自己隐私的权利。

2.4 攻击者模型

针对社交网络中照片隐私发动的攻击行为, 通常表现为通过用户标识来查询用户的相关照片及其标签、元数据信息的正向搜索; 通过人脸照片特征来查询一个未知的用户, 并识别其身份信息的逆向搜索; 以及未经自己知晓情况下被拍摄发布的非意愿展示[8]。其中, 直接搜索与逆向搜索都是对隐私保护机制的逆过程, 是融合与分析环节的主要工作, 它们与非意愿展示共同成为构成社交照片隐私各环节保护方法防护的主要目标。

1) 正向搜索(Directed Searchability)

攻击者通过用户名字、账户 ID 等可以明确区分用户身份的信息在社交网络和照片分享平台中查询并采集指定用户的照片信息, 从中分析并获取用户的隐私信息。

用户享受到各种照片分享服务带来的便利同时, 也面临着异常严重照片隐私泄露问题, 攻击者很容易从照片及其元数据中获取用户的身份(职业信息)、地标信息(工作单位、家庭住址)和喜好(饮食、购物)等信息; 可以推演得到用户的生活习惯、工作状态、私人时间安排等隐私信息, 可用于雇员调查和法院取证; 甚至还可以推理得到该用户的政治倾向、社会地位和宗教信仰等敏感信息, 尤其敌手将上述信息进行融合分析, 对用户的安全和隐私问题将构成异常严峻的威胁。

2) 逆向搜索(Reverse Searchability)

攻击者利用图像处理技术获得人脸特征, 用来查询未知用户的相关照片数据, 根据与其关联的标签、元数据等信息分辨用户的身份, 并籍此获取更多隐私信息。

目前, 社交网络环境中攻击者对用户照片信息的采集日趋便捷和自动化。与此同时, 人脸识别等图像处理技术也日趋成熟, 从所收集的海量照片中定位某个用户和确定用户之间的人际关系变得更加容易; 此外, 照片分享过程中常常与各类标签和元数据绑定, 各类数据的关联整合可以准确地还原并预测个人的社会关系全貌。

3) 非意愿展示(Unintentional Discovery)

非意愿展示是指含有用户隐私的照片传播到用户无意展示的用户。造成非意愿展示的主要途径包括因用户的错误设置和朋友转发而扩散到社交网络中的照片、因朋友和路人拍摄的照片被不受控制地上传到社交网络等。

最经典的非意愿展示威胁表现在某些特殊场景中, 例如政治集会、宗教仪式等, 拍摄者未经过用户允许甚至知晓的情况下被拍摄到照片里, 这种随时可能发生的威胁将极其严重地威胁用户隐私。长此以往, 社交网络中无处不在的照片隐私泄露风险将严重损害人们的隐私利益。

2.5 照片隐私在全生命周期中所面临的威胁

在生成与识别环节, 用户通常将 EXIF、IPTC 或者 XMP 等通用格式将照片与其标签、元数据信息存储在一起。独立的信息往往不会泄露太多的隐私信息, 而将照片文件与诸如谁、何时、何地、何事等不同的信息一同存储, 则极大增加了隐私泄露的可能。为了更好地保护用户隐私, 需要通过更加精准的隐私发现与预警方案对用户信息进行识别和保护, 确保用户在社交网络场景下能够及时发现并处理照片隐私泄露的情况。

在发布与交换环节, 目前服务提供商的照片分享访问控制机制在具体实施过程中往往会忽略掉照片上传者以外其他用户的隐私, 特别在上传多人合照的时候, 会将其他用户的照片扩散到预设访问控制机制以外的用户群中。这种访问控制机制在非意愿展示模型下, 无疑会给照片上的其他用户带来极大的隐私危害。因此, 照片受控共享问题已成为社交网络中隐私保护技术需要考虑的核心问题之一。

在存储与销毁环节, 随着上传到社交网络的照片数量急剧增长, 大多数社交网络数据服务器上的照片文件都是明文存储以节省存储与计算资源的开销, 很多机构甚至服务提供商本身, 通过公开的 API 或者爬虫技术从数据库中获取大量的照片资源, 并借助大数据挖掘技术对照片信息及其标签、元数据

进行分析, 从中发现用户的行为习惯、爱好等隐私信息, 极大地损害了用户的个人隐私。

在融合与分析环节, 分享照片给人们带来便利同时也给攻击者提供了海量有价值的数据集, 人脸识别、数据挖掘等技术的发展也给用户隐私分析带来便利, 攻击者不仅可以从照片中挖掘出可见的隐私, 还可能通过融合其他信息挖掘一些不可见的隐私。因此, 亟需研究抗融合分析的社交照片隐私保护机制, 力求为用户提供更加全面的照片隐私保护。

综上, 社交网络中照片隐私保护机制旨在从照片隐私的生成与感知、发布与交换、存储与销毁、融合与分析四个环节, 采用多种保护机制相结合的基本思想, 提高社交照片生命周期不同环节中的隐私保护效果, 需要考虑以下四个具有挑战性的问题:

(1) 如何实现用户友好的社交照片安全拍摄和照片隐私信息的精准识别?

(2) 如何保障照片分享过程中非意愿展示照片的受控共享?

(3) 如何确保所有权和管理权分离条件下用户上传照片的安全存储和可信销毁?

(4) 如何抵御大数据环境下基于人脸识别和关联分析的融合分析攻击?

3 社交网络中照片隐私保护研究进展

针对社交网络中照片隐私信息在全生命周期内不同环节所面临的安全威胁, 目前照片隐私保护方案的研究主要从以下几方面展开: ①基于图像处理的安全拍摄与隐私识别技术研究; ②面向社交网络共享的照片隐私策略推荐与访问控制机制研究; ③基于照片加密的隐私增强存储与可信销毁方法研究; ④面向照片隐私的融合分析及应对方案研究。

其中, 基于图像处理的安全拍摄与隐私识别技术研究旨在为用户在拍摄和本地照片管理过程中提供各种照片隐私自动发现和预警功能; 面向社交网络共享的照片隐私策略推荐与访问控制机制主要通过推荐和制定一系列访问规则、标准和详细规范来监管和约束用户的访问行为, 防止照片被非授权的用户访问和使用; 基于照片加密的隐私增强存储与可信销毁方法通过对上传到服务商存储服务器中的照片进行加密, 并在有效期到期时对照片进行确定性删除, 防止攻击者和服务商对照片信息进行不正当使用, 同时有效控制了照片的非意愿扩散, 降低隐私泄露风险; 面向照片隐私的融合分析及应对方案, 提出通过大数据技术从公开的照片中挖掘用户的隐私信息, 并寻找应对的防护方案。

3.1 产生与感知环节的照片隐私保护

随着搭载拍摄功能的移动设备以及可穿戴设备的快速发展和普及, 照片的拍摄方式也由专业人员使用专业设备(如单反相机等)进行拍摄转变为大众可随时随地使用移动设备进行拍摄。同时, 借助 Google Glass 等可穿戴设备的主动式照相和自动拍摄功能, 用户可以体验到更便捷的人机交互, 拍摄下更多传统方式难以捕捉的精彩瞬间。

然而, 拍摄设备便携化给用户带来便利的同时, 也使得拍摄过程日益隐蔽, 不可避免地记录下专业设备难以拍摄的隐私场景。与此同时, 随着对隐私的认识不断提高, 用户对照片隐私的需求也更加多样, 传统的拍摄方法及其保护措施难以满足用户个性化的隐私保护需求。此外, 新型拍摄技术还采用了更大容量的存储介质对拍摄的照片进行存储, 相比于传统拍摄方法, 所拍摄照片的数量不再受到限制, 加剧了用户的照片隐私泄露威胁。

3.1.1 隐私增强拍摄

早期的隐私增强拍摄方法利用文字识别、行人检测等技术对照片中的车牌[9]、行人[10]等隐私区域进行识别并做出相应的隐私化处理。然而这些方案对识别出的全部隐私对象往往实施相同粒度的隐私保护措施, 缺乏对用户个性化隐私需求的考虑。随着人脸识别技术的发展, 识别准确率的提高, 将人脸作为个人标识符应用到现实中成为可能, 以人脸识别为基础的个性化隐私保护方案成为隐私增强拍摄方法的主流。

Toubiana 等[11]提出了一种利用手机地理位置信息辅助保护用户隐私的模型 Photo-TaPE(Photo- Tagging Preference Enforcement), 该模型对新拍摄的照片进行人脸识别, 利用地理位置信息将人脸识别的候选集限定到周围的用户, 并根据用户预先设定的策略对所拍摄照片上的人脸区域采取模糊、标记或通过邮件通知本人等措施保护用户隐私。该方法可以有效提高方案中人脸识别的准确率和效率, 从而提高保护用户隐私的效果。但该方案需要获取用户的实时位置信息来辅助筛选人脸识别候选集, 而实时位置信息可能导致尾随抢劫、空屋盗窃等风险, 直接关系到用户的生命财产安全。

为了减少实时位置信息泄露所带来的安全风险, Henne 等[12]提出了一种利用历史位置信息辅助保护隐私的方案 SnapMe, 该方案允许用户指定隐私区域并设定相应的隐私策略。在实际拍摄过程中, 利用用户在隐私区域的签到信息和照片拍摄时间地点等信息查询相关用户, 最终根据人脸识别技术判断该用户是否被拍摄并进行邮件预警。该方案将实时位置

信息改进为历史签到信息, 一定程度上缓解了实时位置信息泄露所带来的隐私侵害。

该类隐私增强拍摄方法流程如图 2 所示, 其中用户 1 和用户 2 使用隐私增强系统的设备, 将自己的位置等信息上传给可信第三方, 可信第三方通过位置筛选、人脸识别等技术, 对用户 1 进行“所在区域有拍摄”预警; 对用户 2 进行“被拍摄”预警, 并根据用户 2 的策略反馈对照片进行处理; 对于未使用该隐私增强平台的用户 3, 隐私保护系统不能对其进行保护。该类方法可以在一定程度上抵抗反向查询和非意愿展示攻击。

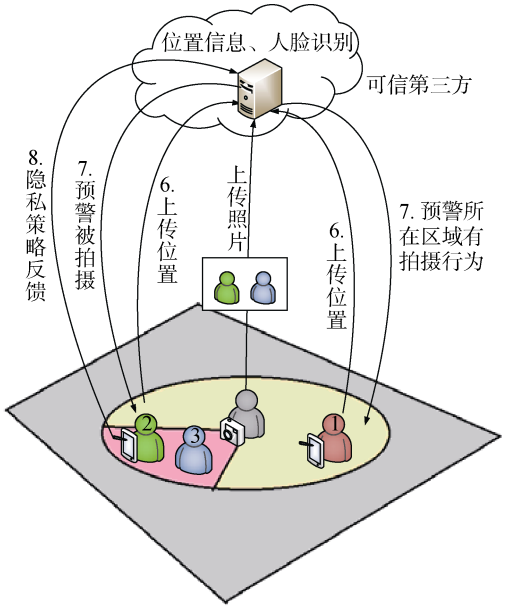


图 2 基于可信第三方的隐私增强拍摄方法示意图
Figure 2 TTP-based Privacy-enhanced Photography

上述方案需要将照片上传至可信第三方, 然而现实中不存在完全可信的第三方。为消除对可信第三方的依赖, 研究者们展开了以下研究。Yus 等[13]提出了一种基于 P2P 通道的隐私保护方案 FaceBlock, 拍摄者通过蓝牙等 P2P 通道收集周围用户的人脸特征和隐私策略。当拍摄者使用 Google 眼镜拍照时, 可以自动根据被拍摄者预先设定的隐私策略, 对人脸进行检测、识别、模糊等操作。该方案中, 用户的隐私策略执行比较严格, 而部分用户希望获得更为灵活的照片隐私控制权限。Li 等[14]提出了一种改进方案 PrivacyCamera, 拍摄者在拍摄时会向附近的用户手机发射信号, 提醒可能的隐私侵害, 可根据反馈进行保护。方案利用被拍摄者与拍摄者的相对位置及被拍摄者的人脸方向等信息, 判断被拍摄者是否为路人, 并通过模糊等方法保护被误拍路人的隐私。

该类隐私增强拍摄方法效果如图 3 所示, 拍摄时, 无论是通过拍摄方发起策略询问, 还是普通用户进行隐私策略广播, 用户间都是通过 P2P 通道传递隐私策略信息和人脸特征, 避免了对可信第三方的依赖。该类方案同样只能保护系统内部用户的照片隐私, 不能保护系统外用户的隐私。

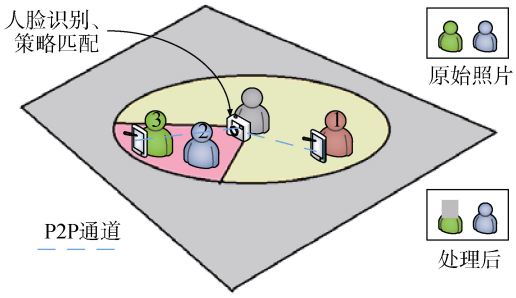


图 3 基于 P2P 通道的隐私增强拍摄方法示意图
Figure 3 P2P-based Privacy-enhanced Photography

由于通过 P2P 通道传递用户的可视特征和隐私选择易被附近其他用户获取, 为防止恶意用户将可视特征和隐私决策用于其他用途。Aditya 等[15]提出了一种集成在相机中的可信软件平台 I-Pic。该平台利用多方安全计算技术对拍摄的内容与用户的隐私策略进行处理, 在保护拍摄者拍摄照片的可视特征和被拍摄者隐私选择的前提下, 实现对拍摄和共享的照片的隐私保护。

因为各安全拍摄方案要求特定的拍摄设备平台, 具有较强的局限性。如果被拍摄者是非平台用户, 则无法实现有效的隐私保护。Pallas 等[8]另辟蹊径, 设计了一种通用离线标记 Offlinetags, 包括“禁止拍摄”“模糊人脸”“允许上传”“允许标记人脸”四种不同的隐私策略。当拍摄者拍摄配戴标记的用户时, 只需拍摄相机能识别出标记的内容, 就可根据标记对拍摄照片进行处理, 以保护用户的隐私。如图 4 所示, 当佩戴蓝色标记时, 对该用户进行拍摄时需要将其脸部进行模糊处理。

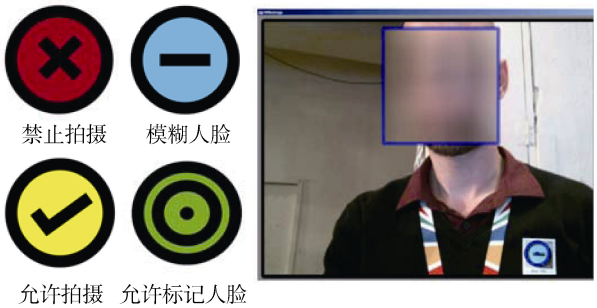


图 4 Offlinetags 标记及效果示意图[8]
Figure 4 Offlinetags Method[8]

以上几种隐私增强方法比较如表 1 所示。

表 1 隐私增强拍摄方法比较

Table 1 Comparison of Privacy-enhanced Photography

方案	基于信息系统	是否使用人脸作为标识符	其他信息辅助
Toubiana[11]	可信第三方	是	实时位置信息
Henne[12]	可信第三方	是	历史签到信息
Yus[13]	P2P 通道	是	—
Li[14]	P2P 通道	是	人脸拍摄角度
Pallas[8]	离线标记	否	图像识别

另外，为提高用户在拍摄过程中的体验，减少人机交互频率，研究者们还提出了一些细粒度、自适应的隐私增强拍摄机制。Templeman 等[16]提出了针对敏感环境的细粒度访问控制框架，利用图像处理技术对不同室内区域进行识别，并使用照片流分析技术提高了识别的准确率。在此基础上，Roesner 等[17]提出了一种通用可扩展的访问控制框架，可以对设备中不同应用进行分别授权。该模型支持用户对现实中的目标设定不同的访问控制策略，例如进入浴室自动停止拍摄或在拍摄过程中自动移除路人的照片等，以减少对用户的提醒次数，并在保证隐私的同时提高可穿戴设备的可用性。Krombholz 等[18]总结了上述隐私增强拍摄方法的特征，提出了一种系统化和量化的隐私增强效果评估方法，希望通过该方法促进隐私增强技术向高可用性、高满意度、低限制的方向发展。

3.1.2 隐私识别与预警

照片隐私作为一种主观性较强的信息，为了更准确地将隐私信息从普通照片中识别出来，研究者们对用户照片隐私内涵开展了多项调研。Ahern 等[19]通过调研 Flickr 用户管理个人隐私照片的习惯和行为，将用户对照片隐私的考量归纳为以下四方面：安全、社会展示、身份识别和便利性。此外，通过进一步调研，研究者们发现用户的隐私策略并不是固定的，用户会根据照片内容和环境对隐私策略的侧重点进行调整。Hoyle 等[7]针对用户隐私照片的分类开展了量化的研究，邀请 36 位志愿者们佩戴可穿戴设备一周，并记录下 14 477 张所拍摄的照片，通过概率分布情况对照片是否含有隐私信息以及是否可以分享的选择进行了汇总。研究结果表明，用户主要会出于隐私信息、印象管理和他人的隐私三方面原因，考虑是否分享照片。通过对用户照片隐私决策的充分调研，可以确定隐私照片与公开照片在样本分布上可区分，这使得根据图像特征和机器学习方法来预测照片隐私成为可能。

从图像处理的角度出发，Zerr 等[1]在邀请网络用户对照片是否属于隐私照片进行标记的基础上，选取可视特征与元数据相结合，利用图像处理和模式识别技术，实现了大量照片数据中隐私照片的自动发现。在此基础上，该研究团队研发了一个实用的隐私预警系统 PicAlert[20]，在用户上传隐私照片时进行辅助预警，并实现了对隐私照片的自动检索。

自 Zerr 等将图像识别技术引入照片隐私领域，其后的研究者们分别在个性化、多分类和准确率 3 个角度对其工作进行深化。如图 5 所示，该类方案基本遵循图像识别的流程。

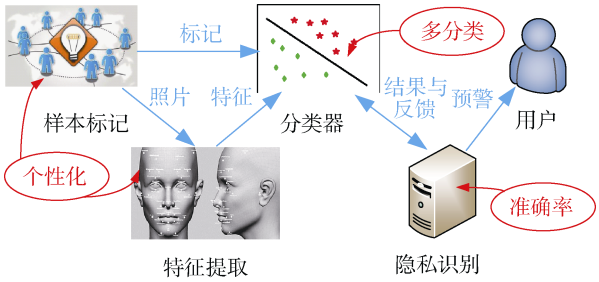


图 5 隐私识别技术流程图

Figure 5 Process of Privacy Recognition

由于方案[1]中所使用的训练集需均要用户共同标记，属于通用型的隐私识别技术。但在实际使用中，不同用户的隐私观具有很大的差异性。为了适应该差异性，Shimada 等[21]提出了“隐私社区”(Privacy Community)的概念，将不同隐私倾向的用户以“社区”为单位进行划分。因为同一社区用户的隐私标准具有一定的群体倾向性，并且较为固定，通过提取社区用户所发布的文本、照片和属性三类信息中的隐私特征，采用基于隐私社区的方案对新发布信息进行的分类，并根据该社区的隐私倾向对可能发生的隐私侵害事件进行预测和提醒。该方案一定程度上在个性化与通用型隐私识别中做出均衡，但仍未对个体用户提供隐私识别服务。为了进一步提高个性化隐私识别，Spyromitros 等[22]提出了一种针对单个用户的隐私照片分类模型。该模型利用语义特征来描述照片内容，并加入用户的少量回馈，构建了个性化的隐私预警和分类模型。

此外，Buschek 等[23]从另一个角度完善了 Zerr 等的工作。该方案通过将可视特征与标签相结合，利用随机森林分类器构建了一种用户隐私照片的细粒度多分类模型。同时，该方案将 Zerr 等工作中的“隐私”与“公开”两类分类扩展到用户自定义的三类以上，并提高了分类的准确率。

在照片隐私识别准确度方面，Tonge 等[24]通过对比可视特征、图像内容描述标签和用户标签等影响

照片隐私预测的主要特征及其组合, 发现利用图像内容描述标签和用户标签进行识别所得到的结果准确率最高。同时, 通过引入深度学习算法, Tran 等[25]提出了一种基于分层特征的隐私检测框架。利用对象和卷积特征建立深度学习模型来检测照片中的隐私。相比于 Zerr 等人所提出的方案[1]以及标准卷积神经网络算法, 该方案的识别结果更为准确。

由于隐私照片数据集往往携带被拍摄者的个人隐私, 研究者们往往出于保护数据提供者隐私的考虑, 无法直接提供原始数据以供对比参考。Zerr 等人在方案[1]中提供的数据集, 是目前网络上少数几个可用的数据集之一。以上几种隐私识别方法在 Zerr 数据集上的实验结果比较如表 2 所示。

表 2 隐私识别方法实验结果比较

Table 2 Experiment Comparison of Privacy Recognition

方案	精准率 (Precision)	召回率 (Recall)	综合评价指标 (F1-Measure)	准确率 (Accuracy)
Zerr[1]	0.65	0.40	0.50	0.65
Tonge[24]	0.83	0.83	0.83	0.83
Tran[25]	0.94	0.85	0.89	0.95

(注: Zerr 方案[1]中数据集包含从 Flickr 网站上收集的多类对象, 并对照片是否为隐私照片进行了标记。)

由于隐私识别与预警技术可以帮助用户有效的区分隐私用公开的照片信息, 所以成为目前隐私照片管理常用的方法, 在其基础上发展的隐私策略推荐与访问控制机制研究是目前照片隐私研究的重要内容之一。

3.2 发布与交换环节的照片隐私保护

随着照片分享方式的不断创新, 在线照片分享服务变得越来越流行。很多平台将不同种类的标签功能集成到照片分享服务中, 例如, 照片在上传到 Facebook 平台时, 可增加上传时间、上传位置、文字注释描述、人脸标记、特殊区域标记等信息。这些标签信息在给用户展示自己和建立线下人际关系提供便利的同时, 也带来了很大的隐私泄露隐患。

此外, 在社交网络照片分享过程中, 已发布信息的收听范围往往不可预测, 而接收者往往又是影响用户是否分享照片的重要因素之一。Bernstein 等[26]通过对 22 万 Facebook 用户的日志数据核查, 并配合调查问卷, 量化分析了发布信息间接听众的范围。调研结果显示用户对听众规模的预测只有真实听众数量级的 27%。发布在互联网上的照片其接收者往往超出用户的预料, 如不加强保护, 会对用户隐私造成严重的侵害。

3.2.1 隐私照片分类与策略推荐

为防止未整理的照片数据中包含有易忽略的隐私信息, 将不同的照片根据内容进行标记分类是照片发布阶段隐私保护的重要工作之一。由于人工标识隐私照片的工作过于乏味且耗费大量的时间和精力, Fesnin 等[27]提出了一种基于语义的隐私照片分类技术, 利用人脸识别、图像分类等图像处理技术对相册中的照片进行语义标记, 并将标记自动扩散至整个相册, 结合拍摄时间、拍摄设备标识等标签, 对不同场景不同时间段的照片进行分类。

然而, 网络上的照片标记不准确且不完全, 同时纠正照片标记工作也是耗费时间和劳动密集的工作。Liu 等[28]提出了标签排序方案, 从照片的所有标签中选出最能表达照片内容的标签。随后他们在工作[29]中提出图片标签排序和图片重新标记方案, 可以通过视觉和语义的相似性对照片进行重新标签, 使标记能够更好的描述照片本身。在他们近期的工作中[30]提出了一个半自动的照片标签标记方案, 仅需要用户给相册中的一些代表性例子进行人工标记, 剩余照片则用邻近传播算法进行标记, 帮助用户减少标记工作, 并提高标记的准确率。当新照片需要加入的时候, 也可以利用邻近传播算法进行标记。

基于标签技术的照片分类方法流程如图 6 所示。该类方法在实现对用户照片分类后, 仍需用户分别对各类照片设置隐私策略, 并且照片内容与分享行为之间并没有一一映射关系。

为提供更自动化的策略推荐服务, 研究者们开展了相应研究。Squicciarini 等[31]设计了一种辅助用户设置照片隐私策略的系统。该系统首先将照片根据图像内容、元数据进行分组, 再根据个人用户的历史隐私策略推测新照片的隐私策略; 若用户的历史策略很少, 或者近期隐私策略突变导致不足以作为预测照片的标准, 则向社交网络中其他类似用户进行请求, 借用多位其他用户的历史信息, 作为预测的训练集, 对新照片进行预测。该系统作为一种典型的照片隐私策略推荐系统, 其架构如图 7 所示。

社交网络中的用户必须平衡考虑“分享什么”和“与谁分享”, 分享的过程中虽然存在隐私泄露风险但也伴随着潜在的社交利益。文献[31]仅从如何保护隐私信息的角度设定策略, 而未考虑用户的分享需求。Kairam 等[32]首次将美学因素纳入策略推荐的考虑, 他们调研了 96 位用户, 并采集了 1040 万张照

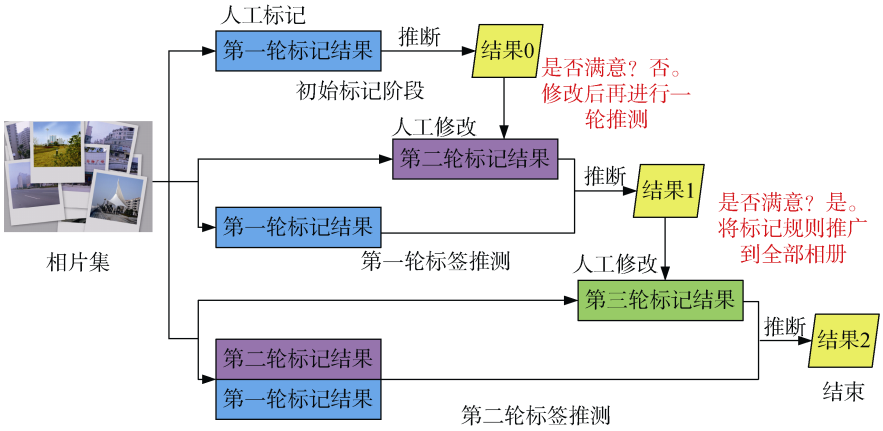


图 6　照片标记方法流程图

Figure 6　Process of Photo Tagging

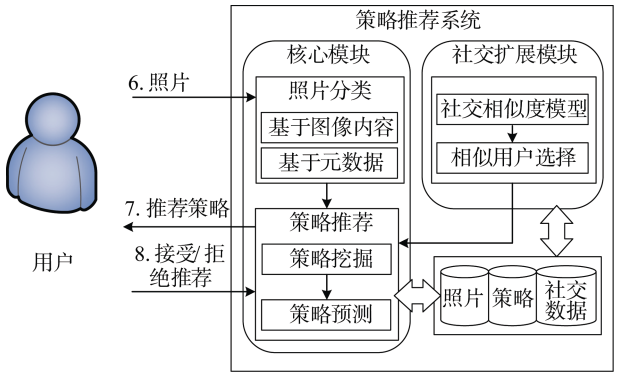


图 7　照片隐私策略推荐系统架构图

Figure 7　Architecture of Policy Recommendation

片，从用户照片分享行为、照片内容特征和美学三个角度对用户的照片分享决策进行分析，利用决策树分类器对测试样本进行分类，希望由此为个人用户提供简单的、可理解的分类规则。实验发现，当照片美学评分高于0.1785且非人物肖像，用户就93.7%的可能将其设置为公开。

由于用户的行为往往与特殊时间和突发事件息息相关, Ni 等[33]对 15 万 Twitter 和 28 万 Instagram 用户的访问控制策略进行调研，发现女性用户、年轻用户和亚洲用户更在意隐私策略，并通过动态分析的方法发现全球性事件和重要节日会影响用户使其更改策略。在此基础上他们结合用户在线行为特征和用户统计资料对隐私策略进行预测，并希望藉此为个人用户自动分配隐私策略，同时为政府机构向公众发布隐私风险预警提供建议。

以上几种隐私增强方法比较如表 3 所示。

此外，在确定访问者范围方面，虽然 Facebook 等社交网络服务商可以允许用户自主选择信息的访问者，但是确定每个信息的每个访问者是非常消耗时间和精力的工作。因此，用户经常将朋友分组，但

表 3　图像隐私策略推荐方法比较

Table 3　Comparison of Policy Recommendation

方案	时间维度	评判要素	适用用户类型
Squicciarini[31]	动态预测	照片内容+策略挖掘	个人用户
Kairam[32]	静态预测	照片内容+策略挖掘+美学	个人用户
Ni[33]	动态预测	策略挖掘	个人用户+政府机构

是再细致的分组都难以满足用户实时变化的分享需求。Yang 等[34]提出了一种基于效用的隐私信息分享框架。通过平衡隐私风险与社交收益两个要素，可以从朋友分组中选择一个子集，成为最适合该消息传播的范围。

3.2.2　面向社交网络交换的访问控制

照片的标记对分类设定访问控制权限也具有重要的意义。Besmer 等[35]研究了用户在社交网络照片分享过程中的隐私风险和保护需求，研究结果表示用户对标记和控制他们的隐私照片具有强烈的需求。他们在后续工作[36]中分析了用户在照片分享领域身份与印象的管理、社会价值、监视隐私、了解攻击威胁、所有权和亲密度等需求，并将标签技术与隐私照片的管理联系起来，以标签为单位对其他用户的照片访问请求进行授权。在此基础上, Klemperer 等[37]通过调研发现，出于组织照片目的设定的标签可以用于对照片的细粒度访问控制，并提出一种基于标签的访问控制模型。实验表明建立合适的标签规则和减少标签的重叠可以有效提高访问控制策略的准确性。

然而标签的标记工作量和准确率问题并不能由自动技术完全解决，除了利用标签进行访问控制的技术以外，还有研究者从其他角度对照片隐私受控

共享开展了研究。Pang 等[38]提出了一种根据不同关系类型进行授权的访问控制方案，方案通过提取 Facebook 中公开信息和关系拓扑中用户群体的属性、共同兴趣和活动等信息，将社交网络关系分成可以被语义理解的类别，针对社交关系设计更接近现实场景的访问控制策略。

由于照片本身除了包含照片所有者的隐私，还包含被拍摄主体、路人甚至拍摄者等利益相关者的隐私信息。上述文献虽然向单个用户提供访问控制机制帮助他们管理自己发布的照片，但尚无提供多用户共同协作控制隐私传播的机制。Hu 等[39]提出了一种多用户隐私信息的合作管理方法，利用量化隐私分享和共享损失，并通过权衡它们来解决隐私冲突问题。他们在现有工作的基础上完成了一个体系化的多用户隐私信息管理模型[5]，该模型利用阈值调整不同用户在最终决策中的影响力以适应不同的场景，并通过投票机制来决定最终的访问控制策略。Palomar 等[40]从属性角度出发提出了一种细粒度访问控制模型和共有数据分享管理系统，该系统收集各个利益相关者的隐私倾向，在保证用户相互之间隐私倾向保密性的前提下，支持对多媒体隐私属性的细粒度访问控制。

在实际系统中，为了识别照片中用户的身份，人脸识别技术被引入照片管理系统。Xu 等[41]提出了一种基于人脸识别技术的多用户照片隐私保护系统。该系统对上传的照片进行分析识别，找出照片中的用户，通知并提醒这些用户与照片上传者协商决定照片发送、评论和标签的控制权限。在此基础上，Ilia 等[42]提出了一种基于人脸的细粒度访问控制系统 Face/off，从而改变了照片分享过程中需要多用户在不同的隐私选择中做出权衡的现状。该系统将 Facebook 中的人脸标记功能与人脸识别结合起来，通过识别出的人脸自动关联到 Facebook 中用户的访问控制策略，并根据策略对用户的人脸区域实施模糊等隐私保护措施。

一个通用的多主体访问控制模型如图 8 所示，在照片发布时，照片的上传者与所有者、利益相关者之间可能存在隐私冲突，需要对冲突的策略进行消解；在传播过程中，原始的隐私策略也可能与传播者的隐私策略产生冲突，也需要进行冲突消解。

以上几种访问控制机制比较如表 4 所示。

现在的社交网站如 Flicker 虽然加入了限制转发的功能，并提供相应的访问控制策略选择，但是通过其他拍摄设备拍摄后再转发的行为仍可规避这种防护方式。Zhang 等[43]提出了一种轻量级硬件无关的

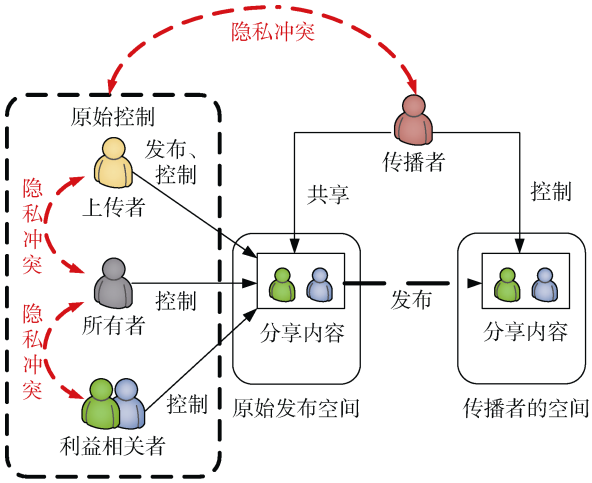


图 8　多方决策隐私冲突模型图

Figure 8　Model of Multiparty Policy Conflicts

表 4　图像隐私访问控制机制比较

Table 4　Comparison of Photo Access Control Methods

方案	访问控制粒度	参与决策者	是否使用人脸作为标识符
Klemperer[37]	同类标签的照片集	照片上传者	否
Pang[38]	照片	照片上传者	否
Hu[5]	照片	利益相关者	否
Palomar[40]	照片的元数据属性	利益相关者	否
Xu[41]	一张照片	利益相关者	是
Ilia[42]	合影上的人脸区域	利益相关者	是

屏幕拍摄干扰系统 Kaleido，通过将照片重新编码成多帧，并提高显示频率、减少延迟时间，实现了在保证肉眼正常阅览的情况下，阻止照相机对屏幕的二次拍摄，其流程如图 9 所示。

Tan 等[44]提出了一种基于人脸识别的实时手机隐私照片保护方案 CHIPS，该方案在安卓 4.2 系统上利用人脸识别技术，对存储卡上的隐私照片进行分类，并根据隐私策略对手机应用进行细粒度的访问授权，防止恶意应用非法盗走照片。

3.3　存储与销毁环节的照片隐私保护

随着移动设备广泛应用和社交网络分享服务的快速发展，存储在社交网络服务商的用户照片也达到了一个惊人的量级。当隐私照片上传到服务提供商时，用户便失去对照片的物理控制，导致用户数据遇到泄露和非授权访问等安全问题。近年来，Facebook、Twitter 等主要社交网站接连发生过隐私照片泄露事件，意味着上传到服务提供商的照片也并非安全无忧；此外，照片分享服务商的存储系统还成为了兴趣发现系统的关注重点[45]，大量照片在用户不知晓的情况下被用来研究用户的兴趣倾向。

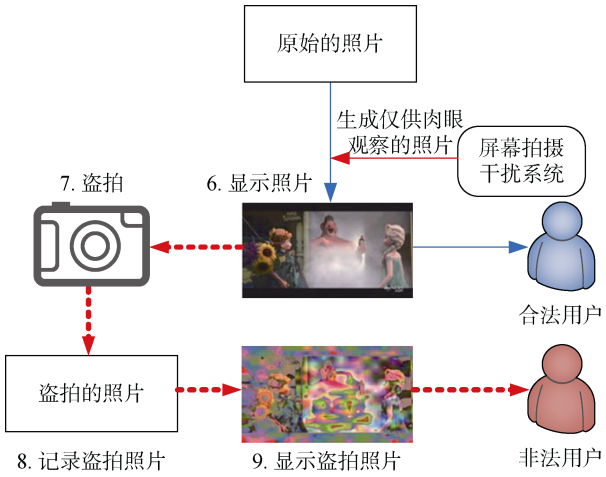


图 9 屏幕拍摄干扰系统示意图①

Figure 9 Screen-shooting Countermeasure System

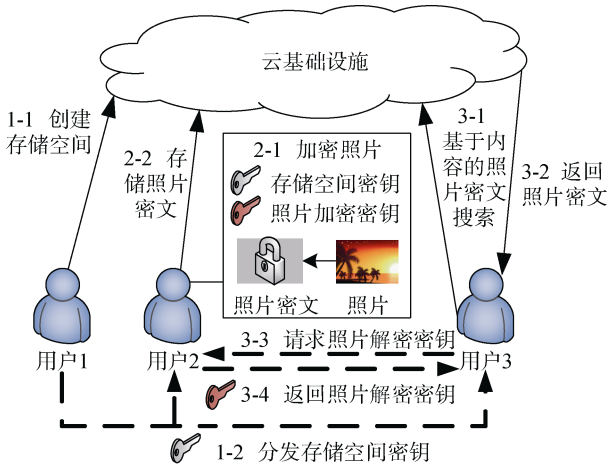


图 10 基于内容的照片密文存储示意图

Figure 10 Content-based Encryption Storage

3.3.1 基于图像加密的隐私增强存储

随着个人用户的照片存储需求越来越大, 将数据外包给云服务提供商或者上传到社交网络服务商的数据中心的做法变得越来越普遍, 将加密技术引入云存储是用户最常用的选择之一。Sosa 等[46]设计了一种隐私增强的照片分享文件系统 PicFS, 该系统只允许所有者修改上传的加密照片并分享给指定的用户, 用户的上传和下载操作匿名于照片分享平台。随后, Rane 等[47]设计了一个基于属性的图像加密与检索系统, 该系统只有当检索的属性向量符合条件时, 才可以解密照片密文, 从而降低了计算复杂度。

为了提高云服务器中加密存储的可用性, 研究者们开展了基于照片内容的密文检索研究。Zhang 等[48]设计了一种云上的隐私增强图像检索系统 PIC, 支持授权用户在加密状态下检索其他用户的图像内容, 实现基于内容的细粒度访问控制。Ferreira 等[49]提出了一种针对外包数据隐私保护存储和检索的安全框架, 该框架支持基于内容图像检索的加密技术 IES-CBIR(Image Encryption Scheme that displays Content-Based Image Retrieval properties), 实验证明 IES-CBIR 适用于多种现实应用场景。利用图像内容的密文检索方面的特性, Yuan 等[50]提出了一种基于图像特征的安全社会关系发现系统, 根据拥有类似照片内容的两个用户更容易成为朋友的同质性理论, 在保护照片内容信息的前提下, 实现了以照片内容为中心的社交关系隐私匹配。

基于内容的照片存储方案如图 10 所示, 该类方案可以有效实现对照片的安全存储和检索, 但存在保护粒度粗、计算开销大等问题。

现有方案主要关注照片内容隐私的保护与安全存储, 然而随着照片拍摄设备越来越自动化, 照片文件中的元数据信息也随之增长。Henne 等[51]提出了一种基于 CaaS(Confidentiality as a Service)加密的元数据受控共享方案, 将访问控制所保护的对象从照片内容本身, 扩展到了其相关的元数据信息。通过将元数据中的属性信息与图像本身的内容信息分开存储并设置不同的密钥, 有效抵抗了基于属性信息的攻击对照片隐私的威胁。

照片中的隐私信息往往集中在照片中的部分区域, 而整张照片加密从某种程度上减少了公开信息的传播范围。Ra 等[52]将选择性照片加密引入照片分享服务中, 将照片划分为隐私区域和公开区域, 并分开存储, 明确了用户对自己照片数据的所有权和管理权, 有效降低了照片自动处理技术对隐私的侵害。随后, Yuan 等[53]提出了一种安全 JEPG 干扰算法, 可对照片上多个不规则区域进行干扰, 经过加密上传到分享服务平台上, 不同的用户可以对图像上的不同区域分别进行加密, 实现了相关用户细粒度个性化的图像安全共享。He 等[54]提出了一种灵活的照片加密方法, 该方法在给不同的隐私区域分配不同密钥的基础上, 还支持对加密后图像进行裁剪、缩放和压缩等常规操作。

面向照片区域的加密算法大多基于图像文件的像素数据按矩阵分布的属性, 对不同的像素区域进行加密, 其流程如图 11 所示。该类算法由于使用图像处理算法对矩阵数据进行变化, 具有支持裁剪、缩放等操作的优点, 但同样存在着加密效果与计算开销相矛盾的问题。

图像加密与图像的密文操作往往伴随着巨大的计算量, 为减轻隐私增强存储中的计算开销问题,

① 图 9 中效果图片引自文献[44], 图中流程根据文献内容作了适当调整.

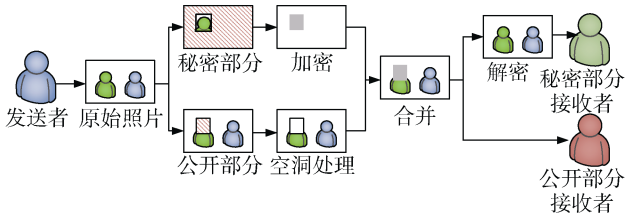


图 11　面向照片区域的加密算法流程图
Figure 11　Process of Partial Encryption

一些研究者开始寻找非加密的隐私增强存储方案。Nourian 等[55]提出了一种基于混沌映射的照片数据混淆编码方案, 该方案支持照片在混淆状态下的像素级操作, 并实现了多种像素级的过滤算法。为防止攻击者通过观察找到混淆编码后的照片与原始照片的关系, 在后续工作中[56], 他们增加了外界照片数据对原有照片集的干扰, 从而增大了密文空间, 提高了照片数据的安全性。

混沌映射算法的基本流程如图 12 所示, 该类算法通过将照片划分 N*N 的方格, 再通过位置置乱的方式将不同方格、不同像素、不同照片的位置进行混淆。为抵抗统计特征分析, 研究者们还对混淆后的像素增加随机浓度, 对像素特征进行隐藏。此类算法可以有效地降低照片加密的算法复杂度。

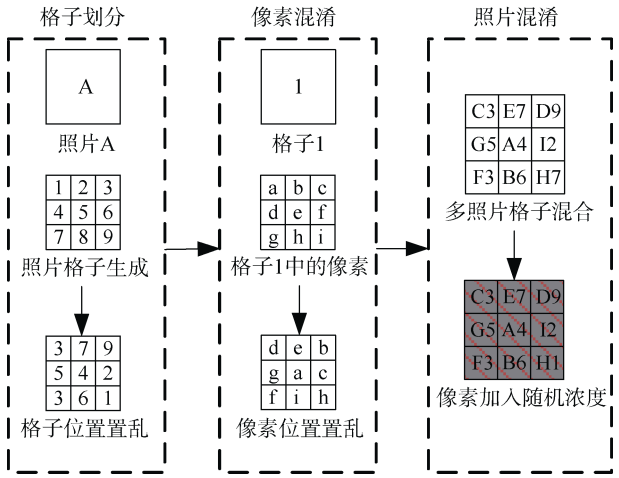


图 12　基于混乱映射的加密算法示意图
Figure 12　Chaotic-mapping-based Encryption

以上几种隐私增强存储方案比较如表 5 所示。

3.3.2　社交网络云存储中的确定性删除

当代的互联网用户, 特别是年轻人, 经常会发布一些隐私照片到互联网上, 并且他们往往没有意识到这些信息会影响到他们未来的生活和工作。此外, 全球有超过 10 亿的用户使用云存储来保管私人文件。很多人并没有充分意识到将文件存储到云存储中就是存储到了互联网上。Clark 等[57]做了一个调查, 并发现很多用户没有意识到自己将隐私照片上传到云上, 当得知后, 大多数人选择将这些上传的照片永久删除。

表 5　图像隐私增加存储方案比较
Table 5　Comparison of Photo Privacy-enhanced Storage

方案	加密算法	算法复杂度	加密对象
Rane[47]	基于属性的加密	$O(m)$	照片文件
Zhang[48]	基于内容的密文检索	$O(m^5+2^m)$	照片内容
Ferreira[49]	基于 IES-CBIR 图像加密	$O(n^2)$	照片内容
Henne[51]	基于 CaaS 加密	调用服务	图像元数据
Ra[52]	P3 算法	$O(m*n)$	照片部分区域
Yuan[53]	安全 JEPG 干扰算法	$O(m*n)$	照片部分区域
Nourian[55]	基于混沌映射算法	$O(m)$	照片内容
Nourian[56]	基于混沌映射算法	$O(m)$	照片内容

(注: n 指加密所用大素数, m 指特征向量长度)

将数据从服务器上永久删除其实并不容易, 因为隐私照片上传到服务商服务器后, 用户便失去对照片的物理控制, 而以数据自毁为基础的隐私信息确定性删除技术从某种程度上使用户重新获得了数据控制权。基于密钥管理的确定性删除技术主要包括密钥集中管理和分散管理两种类型。在密钥集中管理方面, Tang 等[58]设计并实现了一个安全增强存储系统 FADE, 该系统在云存储系统外增加一层细粒度、基于策略的访问控制和文件确定性删除的组件, 并维护了一个独立于第三方云存储的密钥管理系统, 该系统可以无缝接入到现有的云存储服务中。为实现对隐私数据更细粒度的控制, Xiong 等[59]提出了带时间约束的基于属性访问控制方案, 该方案能对每个属性设定不同的授权期限, 达到不同粒度数据块的单独删除。

密钥集中式管理的确定性删除流程如图 13 所示, 该类方法过于依赖可信服务器, 将成为攻击者重点攻击目标而导致单点失效问题。因此, 研究者们开展了对密钥分散式管理下确定性删除的研究。

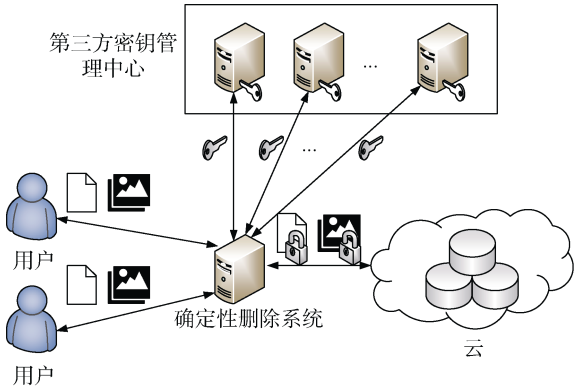


图 13　密钥集中管理的确定性删除方案架构图
Figure 13　Assured Deletion Schemes based on Root-controlling Key Management

Geambasu 等[60]设计了一种集成加密技术的数据自毁系统, 如图 14 所示, 该方案利用一种 P2P 网络——离散哈希表(Distributed Hash Tables, DHTs)的网络特性, 将密钥分量分散到 DHTs 节点上, 其每个网络节点只保存密钥分量 8 小时。随着节点不断更新, 密钥将无法重构, 实现了密钥分散管理下的确定性删除。

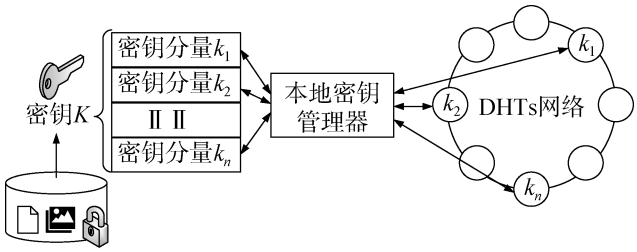


图 14 基于 DHT 网络的密钥分散管理方法示意图
Figure 14 DHT-Network-based Distributing Key Management

由于上述方法的密钥有效期受 DHTs 节点更新周期的限制, Backes 等[61]设计了一种可自定义有效期的隐私照片自动销毁系统——X-pire, 该系统无须向网页中嵌入其他工具, 就可以在设定的期限到期时销毁照片, 用户可以动态的延长或者缩短照片的有效期。该系统需要在 JPEG 加密文件中嵌入有效期和销毁操作等信息, 并且该信息在照片被压缩后仍然有效。随后, 他们在现有基础上提出了升级版本的 X-pire 2.0[62], 该系统在前代的基础上利用可信计算技术, 可以有效防止攻击者生成照片副本, 并可以无缝的接入到现有互联网中。

目前, 确定性删除技术还处于起步阶段, 绝大多数研究者的方案都是针对通用文件设计的, 虽然可以运用在照片隐私领域, 但仍可以根据图像文件的特点进行适配, 文献[62]为面向照片隐私的确定性删除指出了未来研究方向。

3.4 融合与分析环节的照片隐私保护

随着社交照片分享平台中照片数量的快速增长, 海量照片数据已成为一种蕴含巨大价值的资源。然而, 用户所上传的照片很难抵抗数据采集工具的采集和挖掘, 攻击者可以通过社交网站爬虫、窃取手机存储、朋友授权等手段轻易得到用户隐私照片[63]。为评估隐私保护算法的效果并防止针对用户照片隐私的攻击, 研究者们围绕照片隐私融合攻击与抗关联分析“攻防”两个方面开展了多项研究。

3.4.1 隐私融合与分析

在社交网络中共享的照片往往包含着一些零碎的隐私信息, 例如照片合影中就包含一种可视的人际关系。如何将照片中的人际关系映射到现实生活中, 成为部分研究者的兴趣点。Yang 等[64]提出了一种判断人与人之间的姿势与相对位置的图像处理模型, 利用合影中人与人之间的接触编码、连通域模型和姿势评估等技术, 实现对合影中人物的人际关系分类。在随后的工作中, Chakraborty 等[65]将拍摄时人物的空间距离计算出来, 并与人际距离学(Proxemics)的理论结合起来, 实现了对照片人物更细粒度的亲密度分类。

因为图像可以传达一些事实信息, 包括参与某个聚会、遇见某个人等, 因此照片分享在个人隐私中扮演了一个特别的角色。再加上自然社交网络与在线社交网络的重合, 增加了照片内容访问控制的复杂性, Shoshitaishvili 等[63]证明了一种可以根据社交网络中分享照片融合分析恋爱人际关系的攻击方式, 该方法流程如图 15 所示, 通过 Flickr 下载、手机存储窃取、朋友授权 3 种途径获得大量具有元数据的照片集, 将照片中识别的用户按私人亲密场景、公开亲密场景、私人社交场景、公开社交场景 4 个场景进行亲密程度分类, 再根据从照片中获取的环境特征和一些随时间变化的特征对关系进行细分, 来感知用户亲密度随时间的变化。最终判断用户在与谁约会, 甚至可以判断恋爱关系, 并得到了可接受的准确度。

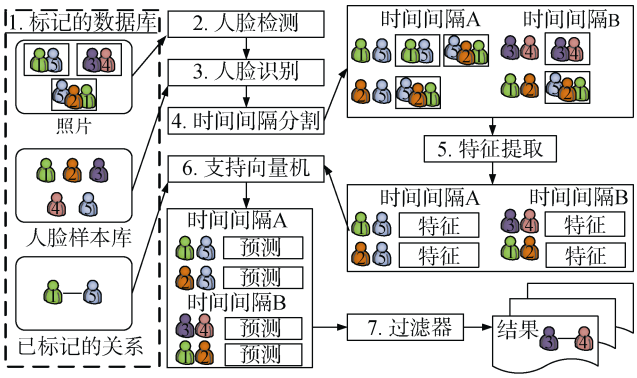


图 15 推断人际关系方法示意图
Figure 15 Process of Relationship Prediction

另一方面, 由于社交照片通常包含人脸等高敏感级别的信息, 为避免照片隐私信息泄露, 人们会利用诸如模糊、用图形遮挡眼睛或脸部等方式来保护用户隐私。Nakashima 等[66]调研了 108 位参与者, 邀请他们对照片比例修改、脸部区域遮蔽、眼部区域遮蔽、模糊等不同防护技术处理后的政治家照片进行识别, 以确认其隐私保护效果。实验表明, 这些技术的效果由访问者与照片主体之间的熟悉程度以及照片主体的显著性决定, 并不能完全保证隐私保护的效果。

除了在线上社交网络的照片交换共享活动, 在现实生活中也常常发生用手机向其他用户展示照片

的情况。在线下社交网络交换照片过程中, 为了防止其他用户看到自己智能手机相册中部分不愿意展示的照片, Zezschwitz 等[67]提出了一种方法对手机中的照片缩略图利用马赛克、结晶化、油画等多种过滤器进行抽象化, 在手机用户本人可以识别的同时防止其他用户猜测到照片内容。

在最近的研究中, McPherson 等[68]提出了一种针对马赛克、高斯滤波器模糊和 P3 加密等泛化区域算法的复原攻击方法, 该方法假设攻击者获得了包含被攻击对象的公开照片集, 并知晓泛化算法所用的参数, 利用参数对公开信息进行泛化处理并提取图像特征, 通过图像识别和神经网络技术, 识别出照片中被遮蔽、模糊或者泛化加密的隐私区域。实验结果表示, 目前通用的泛化参数下, 泛化处理很难抵挡这种攻击。

常见的人脸区域和敏感照片的图像处理算法保护效果如图 16 所示。人脸区域局部遮蔽方法的缺陷在于, 攻击者可通过其他未遮蔽区域的特征推测遮蔽区域信息, 在文献[42]的实验中, 仍有 12.6%遮蔽脸部区域被用户的朋友识别出来。而通过过滤器和泛化的保护方法, 很难抵挡文献[68]中所描述的攻击方法, 其适用范围比较有限。

图 16 图像处理算法保护效果示意图①

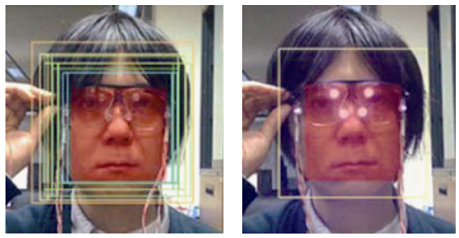
Figure 16 Protective Effects of Picture Processing

随着互联网上发布的各类信息不断积累, 以及机器学习和数据挖掘等技术的飞速发展, 攻击者可以挖掘到的用户隐私的数量和种类都与日俱增, 如何应对融合分析攻击成为研究者们的工作之一。

3.4.2 抗关联分析的隐私保护

早在 2005 年, 研究者们就开始对人脸识别在图像领域的过分应用所产生的隐私威胁展开研究。Newton 等[69]提出了一种抗人脸识别的隐私保护方法, 在保留大多数人脸特征的前提下, 防止人脸被机器学习方法识别。

随后, Yamada 等[70]提出了一种佩戴不可见噪声信号发射器抵抗人脸检测的方案。该方案在脸部佩戴近红外线信号发射系统, 这种信号不可被肉眼捕获, 不会影响到面对面的交流和肉眼观察已拍摄的照片, 却可以被数码相机的感光设备接受从而干扰机器对照片中人脸的自动检测。其效果对比如图 17 所示, 图中方框表示不同特征的人脸检测算法, 为保证误判率不影响到人脸检测的可用性, 超过 3 个方框同时标记则表示人脸被检测出来, 图 17(a)和图 17(b)均可被肉眼观察, 但图 17(b)无法通过机器学习方法检测出用户的脸部区域。



(a) 未使用红外线干扰 (b) 使用红外线干扰

图 17 基于硬件的抗人脸识别方法效果图②

Figure 17 Hardware-based Face Recognition Countermeasure Methods

研究发现将照片打码或者模糊的方式难以完全应对融合分析等攻击, 并且会影响照片在分享阅览过程中的美观。Nakashima 等[71]提出了一种基于人脸融合的抗人脸识别方案, 通过将两张脸的可视特征融合成一张虚拟人脸, 并用这张新的脸部图像替代照片上的脸部区域, 从而达到保护隐私和维持照片美观的目的。该方案效果及流程如图 18 所示, 当选定目标照片后, 将源照片中人脸根据目标人脸轮廓图形进行范围修改, 再将修改后源照片与目标照片的视觉特征进行融合, 形成一张虚拟人脸照片。

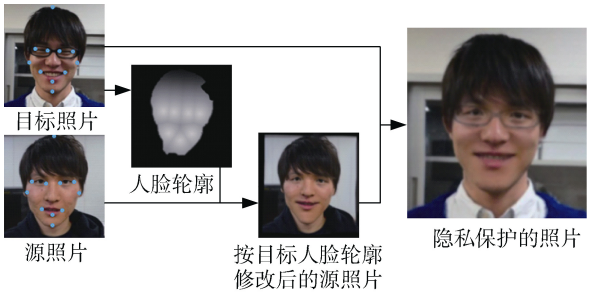


图 18 可视特征融合流程及效果图③

Figure 18 Process of Visual Feature Fusion

① 图 16 中人脸系列照片引自文献[66], 风景系列照片引自文献[67].
② 图 17 引自文献[70].
③ 图 18 中人物照片引自文献[71], 图中流程根据文献内容作了适当调整.

3.5 研究进展小结

针对本文在第 2 章中指出的照片隐私在产生与感知、发布与交换、存储与销毁、融合与分析四个环节中面临的四个挑战, 研究者们从不同角度进行了较为系统的研究, 取得了丰富的成果。本文归纳了以上四个环节中隐私保护方法, 其综合分析与比较如表 6 所示。在生成与感知环节, 隐私增强拍摄方法和隐私识别方法可以有效防止照片隐私信息离开用户的物理控制, 但在高可用性、高满意度、低限制方面仍有改进的空间; 在发布与交换环节, 隐私策略推荐与访问控制可以帮助用户将照片分享给合适的接受者, 但仍需要在用户友好方面对隐私策略的推荐和设置上加以改进; 在存储与销毁环节, 加密存储和确定性删除技术使得用户重新获得照片隐私信息的控制权, 但现有工作大多集中在通用文件的可信销毁上, 仅 X-pire 2.0[62]对照片确定性删除进行了探讨; 在融合与分析环节, 抗关联分析的隐私保护方法可以阻断用户与照片隐私内容的关联, 但抗关联分析方案多聚集在抗人脸识别的工作上, 缺少对其他关联信息的关注。

除了针对照片内容信息保护的研究以外, 研究者们还开展了针对照片属性和关系信息的研究工作。例如, 在照片属性信息保护的研究中, 研究者们在存储与销毁环节提出了将属性信息与照片文件分开存储的方式[51]来提高照片属性信息的安全, 并通过设置不同密钥阻断了攻击者获取照片内容与属性之间的关联关系; 在关系属性保护的研究中, 研究者们在融合与分析环节中提出了抗人脸识别的保护方案[69-71], 有效防止了利用人脸作为个人标识符的关系信息自动发现与推测攻击。

综合上述分析, 现有工作主要实现了单一环节、特定场景的照片隐私保护, 均未全面考虑照片隐私信息在不同环节衔接、多场景切换等情形下的安全问题, 未来将围绕照片隐私全生命周期探索统一的隐私保护框架机制。

4 发展趋势与未来研究方向

目前, 社交网络中照片隐私保护还是一个相对年轻的研究领域, 其研究脉络如图 19 所示, 在理论和应用上都还存在一些难点以及新的方向需要进一步研究探讨, 主要包括:

1) 用户友好的图像隐私增强拍摄方法

随着搭载相机的移动设备和可穿戴设备的飞速发展, 涉及隐私的照片无论是数量还是种类都远远超过了胶卷时代。目前还没有产品化的隐私增强拍摄方法, 学术界的解决方案往往针对某一特定场景或者限制用户必须使用某种平台, 并且需要用户频繁参与隐私决策。此外, 一些方案需要收集用户的位置或生理特征等隐私信息, 使得用户必须得在多种隐私中做出权衡。因此, 隐私增强拍摄方法还需要在适应多种拍摄场景和提高可用性等方面进行更深入的研究。

2) 照片的二次转发控制方法

社交网络中照片的传播越来越便利, 互联网上的照片接收者往往超出照片发布者的预料[26], 现有服务提供商的访问控制机制基本只考虑照片首次传播, 虽然有一些限制二次转发的方法 Kaleido[43], 但该方法过于严格, 仅可作为防止通过拍屏恶意制作照片副本的手段。从目前的研究来看[72], 面向照片二次传播及其传播链路的细粒度访问控制研究具有良好的应用前景, 能够为现实中的隐私保护问题提供有效的解决方法。

3) 针对照片隐私的确定性删除方法

目前的确定性删除方法大多是针对通用格式的文件, 虽然可以应用在照片文件上, 但照片文件仍然存在一些其独有的特征, 可能导致其他相关问题。X-pire 2.0[62]虽然通过在 JPEG 加密文件嵌入有效期和销毁操作等信息实现了 JPEG 格式照片的确定性删除。但照片文件的格式多样且易于产生副本, 如何实现多格式照片数据确定性删除和构建相似图像内容副本的关联关系都具有技术上的难点。尽管如此, 作者认为在现有通用文件确定性删除技术基础上的完善和延伸, 对于扩展照片确定性删除的应用领域具有重要的意义。

4) 抗融合分析的隐私保护方法

互联网中每天发布的照片以及历史积累的照片已经达到一个惊人的量级, 从某种程度上成为了大数据挖掘的最佳资源。另一方面, 人脸识别技术为有效关联照片隐私内容与用户身份搭建了桥梁。因此, 攻击者可以通过融合分析得出用户的隐私信息。目前应对照片隐私融合分析攻击的保护方法主要集中在干扰人脸识别技术上, 其目的在于使得人脸无法作为个人身份标识符, 从而阻断用户与照片的关联。随着攻击技术的发展, 针对非人脸信息关联的攻击方法将不断涌现, 其应对防护技术也将不断适应新攻击技术的特征, 螺旋交替式发展。

5) 通用的照片隐私度量标准

在社交照片的全生命周期过程中, 隐私泄露的途径多种多样, 通过不同途径泄露的隐私信息种类也不尽相同。现有隐私保护方案大多针对单一种类

表 6　社交照片隐私保护机制比较

Table 6　Comparison of Photo Privacy-Preserving Mechanisms

环节	技术	研究角度	方案	保护粒度	标识符种类	是否依赖可信第三方	是否使用元数据	是否抵抗直接查询	是否抵抗逆向查询	是否抵抗非意愿展示	优点	缺点
生成与感知	隐私增强拍摄	基于位置辅助	Toubiana[11]	人脸	人脸	是	是	否	是	是	在生成环节就阻断了隐私信息的传播	非同平台用户难以获得隐私策略制定权利, 且人机交互设置较多
			Henne[12]	人脸	人脸	是	是	否	是	是		
		基于 P2P 通道	Yus[13]	人脸	人脸	否	否	否	是	是		
			Li[14]	人脸	人脸	否	否	否	是	是		
		基于离线标记	Pallas[8]	人脸	离线标记	否	否	否	是	是		
		优化人机交互	Templeman[16]	环境	内容	否	是	否	—	否		
			Roesner[17]	环境+人脸	内容	否	是	否	是	是		
	隐私识别与预警	基础	Zerr[1]	照片	内容	否	是	是	是	是	在发布前将隐私信息过滤出来, 防止其传播	存在准确率问题, 无法提供严格的保护机制
		个性化优化	Shimada[21]	照片	内容	否	是	是	是	是		
			Spyromitros[22]	照片	内容	否	是	是	是	是		
		多分类	Buschek[23]	照片	内容	否	是	是	是	是		
		准确率提高	Tonge[24]	照片	内容	否	是	是	是	是		
			Tran[25]	照片	内容	否	是	是	是	是		
发布与交换	照片分类与策略推荐	照片分类	Fesnin[27]	相册	语义	否	是	—	—	—	在发布时推荐合适的隐私策略, 减少工作量	存在准确率问题, 无法提供严格的保护机制
			Liu[30]	相册	语义	否	是	—	—	—		
		策略推荐	Squicciarini[31]	照片	内容	否	是	是	是	否		
			Kairam[32]	照片	内容	否	是	是	是	否		
			Ni[33]	照片	行为	否	是	是	是	否		
	面向社交网络交换的访问控制	基于标签	Klemperer[37]	相册	标签	否	是	是	是	否	可提供严格的访问控制机制	策略变更时, 需要较多设置
		基于关系	Pang[38]	照片	关系	否	否	是	是	否		
		共有数据的访问控制	Hu[5]	照片	—	否	否	是	是	是		
			Palomar[40]	属性	属性	否	是	是	是	是		
			Xu[41]	照片	人脸	是	否	是	是	是		
			Ilia[42]	人脸	人脸	是	否	是	是	是		
存储与销毁	基于图像加密的隐私增强存储	图像加密与检索	Rane[47]	照片	属性	是	是	是	是	是	可防止攻击者获得隐私数据	需要较大的计算、存储开销
			Zhang[48]	照片	内容	是	否	是	是	是		
			Ferreira[49]	照片	内容	是	否	是	是	是		
		部分图像加密	Ra[52]	区域	区域	否	否	是	是	是		
			Yuan[53]	区域	区域	否	否	是	是	是		
		轻量级加密	Nourian[55]	照片	—	是	否	是	是	是		
			Nourian[56]	照片	—	是	否	是	是	是		
	确定性删除	集中密钥管理	Tang[58]	文件	—	是	否	是	是	是	帮助用户重新获得数据的控制权	缺少针对照片特征的研究
			Xiong[59]	文件	—	是	否	是	是	是		
		分散密钥管理	Geambasu[60]	文件	—	否	否	是	是	是		
			Backes[62]	照片	—	否	否	是	是	是		
融合与分析	隐私融合分析	人际关系识别	Yang[64]	—	—	—	—	—	—	—	评估现有方案有效性, 为抗攻击方法提供目标	该研究点尚处于起步阶段
			Chakraborty[65]	—	—	—	—	—	—	—		
			Shoshitaishvili[63]	—	人脸	否	是	—	—	—		
		泛化区域识别	McPherson[68]	—	人脸	否	否	—	—	—		
	抗关联分析	软件方法	Newton[69]	人脸	人脸	否	否	否	是	是	可以有效阻断机器识别用户人脸标识	缺少对其他关联标识的研究
			Nakashima[71]	人脸	人脸	否	否	否	是	是		
		硬件方法	Yamada[70]	人脸	人脸	否	否	否	是	是		

(注: “—”表示该方案无法参与该评分项。)

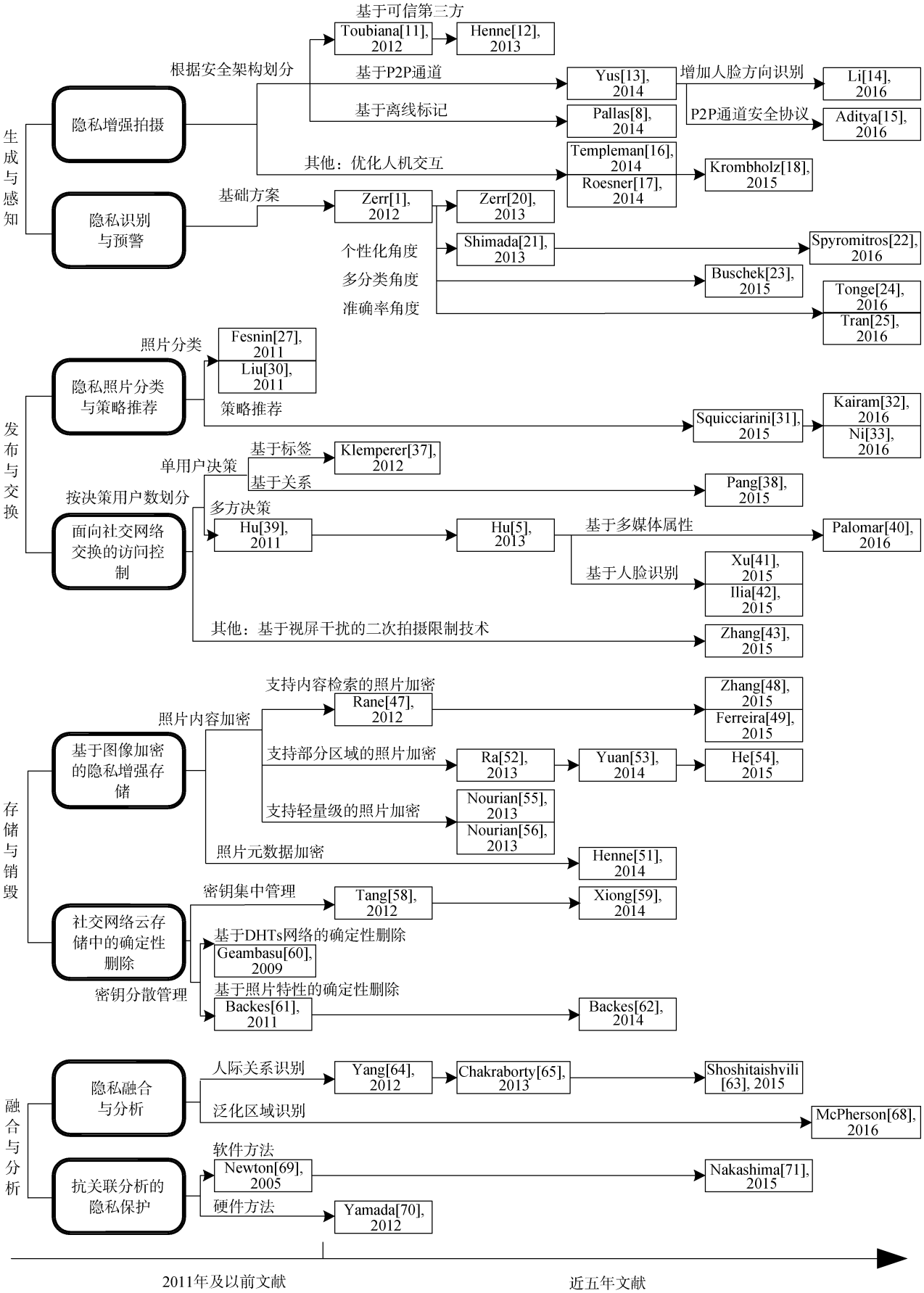


图 19　照片隐私保护研究发展脉络图

Figure 19　Pedigree Chart of Photo Privacy-Preserving Mechanisms

的隐私信息。由于系统资源有限, 难以同时满足多种隐私信息的保护需求, 需要对各类隐私信息进行量化并作出权衡, 从而实现不同种类隐私保护方法的取舍。然而, 在目前的研究成果中还未见成熟的跨环节跨种类的照片隐私度量标准。因此, 制定通用的照片隐私风险度量体系是建立全生命周期保护方案的基础, 值得研究者们进一步研究和探索。

6) 照片隐私全生命周期的保护框架系统

照片隐私保护是目前信息保护领域的研究热点之一, 取得了丰富的研究成果。但从实际应用角度而言, 现有研究成果大多聚焦于照片隐私的某一个环节, 其应对的场景也相对单一, 缺乏跨环节、多场景的隐私方案, 亟需从全生命周期的角度设计更为通用的隐私保护框架。此类框架的研究难点在于如何实现不同环节隐私保护方法的衔接与优化、根据场景变化自适应地选择隐私保护算法与参数, 并从整体上考虑照片隐私保护需求与代价的均衡, 避免出现“木桶效应”。

5 结束语

伴随着便携式拍摄技术的快速发展和社交网络中用户分享照片的飞速增长, 用户对自己的照片隐私愈发重视。照片信息在其生命周期过程中面临着多样的隐私泄露风险, 若不妥善处理, 将严重侵害用户的隐私利益。因此, 社交网络中照片隐私成为了信息安全领域中一个新兴的研究热点。总体而言, 其研究还处于起步阶段, 尚未建立一套完整的隐私保护体系, 其系统距实际应用还有不小的差距。

本文首先列举了社交网络环境中照片在不同环节面临的主要挑战, 包括照片隐私信息的产生与识别、发布与交换、存储与销毁、分析与融合四个环节, 指出了照片的隐私信息分类除应包括内容信息、属性信息外, 合影中的人际关系信息也是照片隐私中需要的关注的问题; 本文给出了直接查询、反向查询和非意愿展示三个类型的攻击者模型; 然后, 回顾了近年来学术界在社交网络中各个环节隐私保护研究领域的主要成果, 从隐私增强拍摄、隐私识别与预警、隐私照片分类与策略推荐、访问控制、加密存储、确定性删除、融合分析及其应对方法等方面对相关研究工作的基本思想、工作原理等进行了深入分析、归纳与总结, 分别指出了各种技术方法的优缺点及存在的共性问题; 最后预测了该领域的未来研究方向。

致 谢 在此向对本文成文过程中给予指导的老师、提供帮助的同学和给本文提出建议的评审专家表示感谢。

参考文献

[1] S. Zerr, S. Siersdorfer, J. Hare, and E. Demidova, “Privacy-Aware Image Classification and Search,” in *Proc. ACM Int’l Conf. Research and Development in Information Retrieval (SIGIR*’12), pp. 35–44, 2012.

[2] “Online Reputation in a Connected World,” Microsoft, http://download.microsoft.com/download/c/d/2/cd233e13-a600-482f-9c97-545bb4ae93b1/dpd_online reputation research_overview.doc, 2009.

[3] B. Henne and M. Smith, “Awareness about Photos on the Web and How Privacy-Privacy-Tradeoffs Could Help,” in *Proc. Springer Int’l Conf. Financial Cryptography and Data Security (FC’13)*, pp. 131–148, 2013.

[4] Fenhua Li, Hui Li, Yan Jia, Nanghai Yu, and Jian Weng, “Privacy Computing: Concept, Connotation and Its Research Trend,” *Journal on Communications*, vol. 37, no. 4, p. 1–11 (in Chinese), 2016.
(李凤华, 李晖, 贾焰, 俞能海, 翁健, “隐私计算研究范畴及发展趋势”, *通信学报*, 2016, 37(4): 1–11.)

[5] H. Hu, G.-J. Ahn, and J. Jorgensen, “Multiparty Access Control for Online Social Networks: Model and Mechanisms,” *IEEE Trans. on Knowledge and Data Engineering*, vol. 25, no. 7, pp. 1614–1627, 2013.

[6] E. Zheleva and L. Getoor, “To Join or Not to Join: The Illusion of Privacy in Social Networks with Mixed Public and Private User Profiles,” in *Proc. ACM Int’l Conf. World Wide Web (WWW’09)*, pp. 531–540, 2009.

[7] R. Hoyle, R. Templeman, D. Anthony, D. Crandall, and A. Kapadia, “Sensitive Lifelogs: A Privacy Analysis of Photos from Wearable Cameras,” in *Proc. ACM Conf. Human Factors in Computing Systems (CHI’15)*, pp. 1645–1648, 2015.

[8] F. Pallas, M.-R. Ulbricht, L. Jaume-Palasí, and U. Höppner, “Offlinetags: A Novel Privacy Approach to Online Photo Sharing,” in *Proc. ACM Conf. Human Factors in Computing Systems (CHI’14)*, pp. 2179–2184, 2014.

[9] I. Azogu and H. Liu, “Privacy-Preserving License Plate Image Processing,” in *Proc. IEEE Global Communications Conf. Workshops (Globecom’11)*, pp. 34–39, 2011.

[10] A. Nodari, M. Vanetti, and I. Gallo, “Digital Privacy: Replacing Pedestrians from Google Street View Images,” in *Proc. IEEE Int’l Conf. Pattern Recognition (ICPR’12)*, pp. 2889–2893, 2012.

[11] V. Toubiana, V. Verdot, B. Christophe, and M. Boussard, “Photo-TaPE: User Privacy Preferences in Photo Tagging,” in *Proc. ACM Int’l Conf. World Wide Web (WWW’12)*, pp. 617–618, 2012.

[12] B. Henne, C. Szongott, and M. Smith, “SnapMe if You Can: Privacy Threats of Other Peoples’ Geo-tagged Media and What We Can Do About It,” in *Proc. ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec’13)*, pp. 95–106, 2013.

[13] R. Yus, P. Pappachan, P. K. Das, E. Mena, A. Joshi, and T. Finin, “FaceBlock: Privacy-Aware Pictures for Google Glass,” in *Proc. ACM Int’l Conf. Mobile Systems, Applications, and Services*

(MobiSys'14), p. 336, 2014.

[14] A. Li, Q. Li, and W. Gao, "PrivacyCamera: Cooperative Privacy-Aware Photographing with Mobile Phones," in *Proc. IEEE Int'l Conf. Sensing, Communication, and Networking (SECON'16)*, pp. 1–9, 2016.

[15] P. Aditya, R. Sen, P. Druschel, S. J. Oh, R. Benenson, M. Fritz, B. Schiele, B. Bhattacharjee, and T. T. Wu, "I-pic: A platform for privacy-compliant image capture," in *Proc. ACM Int'l Conf. Mobile Systems, Applications, and Services (MobiSys'16)*, pp. 235–248, 2016.

[16] R. Templeman, M. Korayem, D. J. Crandall, and A. Kapadia, "PlaceAvoider: Steering First-Person Cameras away from Sensitive Spaces," in *Proc. ISOC Network and Distributed System Security Symp. (NDSS'14)*, pp. 1–15, 2014.

[17] F. Roesner, D. Molnar, A. Moshchuk, T. Kohno, and H. J. Wang, "World-Driven Access Control for Continuous Sensing," in *Proc. ACM Conf. Computer and Communications Security (CCS'14)*, pp. 1169–1181, 2014.

[18] K. Krombholz, A. Dabrowski, M. Smith, and E. Weippl, "Ok Glass, Leave me Alone: Towards a Systematization of Privacy Enhancing Technologies for Wearable Computing," in *Proc. Springer Int'l Conf. Financial Cryptography and Data Security (FC'15)*, pp. 274–280, 2015.

[19] S. Ahern, D. Eckles, N. S. Good, S. King, M. Naaman, and R. Nair, "Over-Exposed? Privacy Patterns and Considerations in Online and Mobile Photo Sharing," in *Proc. ACM Conf. Human Factors in Computing Systems (CHI'15)*, pp. 357–366, 2007.

[20] S. Zerr, S. Siersdorfer, and J. Hare, "PicAlert!: A System for Privacy-aware Image Classification and Retrieval," in *Proc. ACM Int'l Conf. Information and Knowledge Management (CIKM'12)*, pp. 2710–2712, 2012.

[21] S. Shimada and I. Echizen, "Predictive Notification Service of Privacy Invasion on Posting Pictures to SNS," in *Proc. IEEE Int'l Computer Software and Applications Conf. Workshops (COMPSAC'13)*, pp. 193–199, 2013.

[22] E. Spyromitros-Xioufis, S. Papadopoulos, A. Popescu, and Y. Kompatsiaris, "Personalized Privacy-Aware Image Classification," in *Proc.ACM Int'l Conf. Multimedia Retrieval (ICMR'16)*, pp. 71–78, 2016.

[23] D. Buschek, M. Bader, E. von Zezschwitz, and A. De Luca, "Automatic Privacy Classification of Personal Photos," in *Proc. IFIP Conf. Human-Computer Interaction (INTERACT'15)*, pp. 428–435, 2015.

[24] A. Tonge and C. Caragea, "Image Privacy Prediction Using Deep Features," in *Proc. AAAI Conf. Artificial Intelligence (AAAI'16)*, pp. 1–2, 2016.

[25] L. Tran, D. Kong, H. Jin, and J. Liu, "Privacy-CNH: A Framework to Detect Photo Privacy with Convolutional Neural Network using Hierarchical Features," in *Proc. AAAI Conf. Artificial Intelligence (AAAI'16)*, pp. 1–7, 2016.

[26] M. S. Bernstein, E. Bakshy, M. Burke, and B. Karrer, "Quantifying the Invisible Audience in Social Networks," in *Proc. ACM Conf. Human Factors in Computing Systems (CHI'13)*, pp. 21–30, 2013.

[27] A. Fesnin, V. Gouet-Brunet, S. Kominen, V. Oria, and J. Sun, "Towards a Privacy Preserving Personal Photo Album Manager with Semantic Classification, Indexing and Querying Capabilities," in *Proc. ACM Int'l Conf. Multimedia (MM'11)*, pp. 835–836, 2011.

[28] D. Liu, X.-S. Hua, L. Yang, M. Wang, and H.-J. Zhang, "Tag Ranking," in *Proc.ACM Int'l Conf. World Wide Web (WWW'09)*, pp. 351–360, 2009.

[29] D. Liu, X.-S. Hua, M. Wang, and H.-J. Zhang, "Retagging Social Images Based on Visual and Semantic Consistency," in *Proc.ACM Int'l Conf. World Wide Web (WWW'10)*, pp. 1149–1150, 2010.

[30] D. Liu, M. Wang, X.-S. Hua, and H.-J. Zhang, "Semi-Automatic Tagging of Photo Albums via Exemplar Selection and Tag Inference," *IEEE Trans. on Multimedia*, vol. 13, no. 1, pp. 82–91, 2011.

[31] A. C. Squicciarini, D. Lin, S. Sundareswaran, and J. Wede, "Privacy Policy Inference of User-Uploaded Images on Content Sharing Sites," *IEEE Trans. Knowledge and Data Engineering*, vol. 27, no. 1, pp. 193–206, 2015.

[32] S. Kairam, J. Kaye, J. A. G. Gómez, and D. A. Shamma, "Snap Decisions? How Users, Content, and Aesthetics Interact to Shape Photo Sharing Behaviors," in *Proc. ACM Conf. Human Factors in Computing Systems (CHI'16)*, pp. 113–124, 2016.

[33] M. Ni, Y. Zhang, W. Han, and J. Pang, "An Empirical Study on User Access Control in Online Social Networks," in *Proc. ACM Symp. Access Control Models and Technologies (SACMAT'16)*, pp. 13–23, 2016.

[34] M. Yang, Y. Yu, A. K. Bandara, and B. Nuseibeh, "Adaptive Sharing for Online Social Networks: A Trade-off between Privacy Risk and Social Benefit," in *Proc. IEEE IInt'l Conf. Trust, Security and Privacy in Computing and Communications (TrustCom'14)*, pp. 45–52, 2014.

[35] A. Besmer and H. R. Lipford, "Privacy Perceptions of Photo Sharing in Facebook," in *Proc. ACM Symp. Usable Privacy and Security (SOUPS'08)*, pp. 1–2, 2008.

[36] A. Besmer and H. Richter Lipford, "Moving Beyond Untagging: Photo Privacy in a Tagged World," in *Proc. ACM Conf. Human Factors in Computing Systems (CHI'10)*, pp. 1563–1572, 2010.

[37]P. Klemperer, Y. Liang, M. Mazurek, M. Sleeper, B. Ur, L. Bauer, L. F. Cranor, N. Gupta, and M. Reiter, "Tag, You Can See It! Using Tags for Access Control in Photo Sharing," in *Proc. ACM Conf. Human Factors in Computing Systems (CHI'12)*, pp. 377–386, 2012.

[38] J. Pang and Y. Zhang, "A New Access Control Scheme for Facebook-Style Social Networks," *Computers & Security*, vol. 54, no. C, pp. 44–59, 2015.

[39] H. Hu, G.-J. Ahn, and J. Jorgensen, "Detecting and Resolving Privacy Conflicts for Collaborative Data Sharing in Online Social Networks," in *Proc. IEEE Annual Computer Security Applications Conference (ACSAC'11)*, pp. 103–112, 2011.

[40] E. Palomar, L. González-Manzano, A. Alcaide, and Á. Galán, "Implementing a Privacy-Enhanced Attribute-Based Credential System for Online Social Networks with Co-Ownership Management," *IET Information Security*, vol. 10, no. 2, pp. 60–68, 2016.

[41] K. Xu, Y. Guo, L. Guo, Y. Fang, and X. Li, "My Privacy My

Decision: Control of Photo Sharing on Online Social Networks," *IEEE Trans. Dependable and Secure Computing*, vol. PP, no. 99, pp. 1–13, 2015.

[42] P. Ilia, I. Polakis, E. Athanasopoulos, F. Maggi, and S. Ioannidis, "Face/Off: Preventing Privacy Leakage from Photos in Social Networks," in *Proc. ACM Conf. Computer and Communications Security (CCS'15)*, pp. 781–792, 2015.

[43] L. Zhang, C. Bo, J. Hou, X.-Y. Li, Y. Wang, K. Liu, and Y. Liu, "Kaleido: You Can Watch It But Cannot Record It," in *Proce. ACM IInt'l Conf. Mobile Computing and Networking (MOBICOM'15)*, pp. 372–385, 2015.

[44] J. Tan, U. Drolia, R. Martins, R. Gandhi, and P. Narasimhan, "Short paper: CHIPS: Content-based Heuristics for Improving Photo Privacy for Smartphones," in *Proc. ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec'14)*, pp. 213–218, 2014.

[45] D. Beaver, S. Kumar, H. C. Li, J. Sobel, P. Vajgel, and I. Facebook, "Finding a Needle in Haystack: Facebook's Photo Storage," in *Proc. USENIX Symp. Operating Systems Design and Implementations (OSDI'10)*, pp. 1–14, 2010.

[46] C. Sosa, B. C. Sutton, and H. H. Huang, "PicFS: The Privacy-Enhancing Image-Based Collaborative File System," in *Proc. IEEE Int'l Conf. Parallel and Distributed Systems (ICPADS'10)*, pp. 99–106, 2010.

[47] S. Rane and W. Sun, "An Attribute-Based Framework for Privacy Preserving Image Querying," in *Proc. IEEE Int'l Conf. Image Processing (ICIP'12)*, pp. 2649–2652, 2012.

[48] L. Zhang, T. Jung, P. Feng, K. Liu, X.-Y. Li, and Y. Liu, "PIC: Enable Large-Scale Privacy Preserving Content-Based Image Search on Cloud," in *Proc. IEEE Int'l Conf. Parallel Processing (ICPP'15)*, pp. 949–958, 2015.

[49] B. Ferreira, J. Rodrigues, J. Leitao, and H. Domingos, "Privacy-Preserving Content-Based Image Retrieval in the Cloud," in *Proc. IEEE Int'l Symp. Reliable Distributed Systems (SRDS'15)*, pp. 11–20, 2015.

[50] X. Yuan, X. Wang, C. Wang, A. Squicciarini, and K. Ren, "Enabling Privacy-Preserving Image-Centric Social Discovery," in *Proc. IEEE Int'l Conf. Distributed Computing Systems (ICDCS'14)*, pp. 198–207, 2014.

[51] B. Henne, M. Koch, and M. Smith, "On the Awareness, Control and Privacy of Shared Photo Metadata," in *Proc. Springer Int'l Conf. Financial Cryptography and Data Security (FC'14)*, Springer, 2014, pp. 77–88.

[52] M.-R. Ra, R. Govindan, and A. Ortega, "P3: Toward Privacy-Preserving Photo Sharing," in *Proc USENIX Symp. Networked Systems Design and Implementation (NSDI'13)*, pp. 515–528, 2013.

[53] L. Yuan, P. Korshunov, and T. Ebrahimi, "Secure JPEG Scrambling Enabling Privacy in Photo Sharing," in *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition (FG'15)*, pp. 1–6, 2015.

[54] J. He, B. Liu, X. Bao, H. Jin, and G. Kesidis, "On Privacy Preserving Partial Image Sharing," in *Proc. IEEE Int'l Conf. Distributed Computing Systems (ICDCS'15)*, pp. 758–759, 2015.

[55] A. Nourian and M. Maheswaran, "An Approach for Privacy Enhanced Pixel-Level Image Processing in Hybrid Clouds," in *Proc. IEEE Int'l Conf. Computer Communications and Networks (ICCCN'13)*, pp. 1–8, 2013.

[56] A. Nourian and M. Maheswaran, "Privacy Aware Image Template Matching in Clouds Using Ambient Data," *The Journal of Supercomputing*, vol. 66, no. 2, pp. 1049–1070, 2013.

[57] J. W. Clark, P. Snyder, D. McCoy, and C. Kanich, "'I Saw Images I Didn'T Even Know I Had' Understanding User Perceptions of Cloud Storage Privacy," in *Proc. ACM Conf. Human Factors in Computing Systems (CHI'15)*, pp. 1641–1644, 2015.

[58] Y. Tang, P. P. C. Lee, J. C. S. Lui, and R. Perlman, "Secure Overlay Cloud Storage with Access Control and Assured Deletion," *IEEE Trans. Dependable and Secure Computing*, vol. 9, no. 6, pp. 903–916, 2012.

[59] J. Xiong, X. Liu, Z. Yao, J. Ma, Q. Li, K. Geng, and P. S. Chen, "A Secure Data Self-Destructing Scheme in Cloud Computing," *IEEE Trans. Cloud Computing*, vol. 2, no. 4, pp. 448–458, 2014.

[60] R. Geambasu, T. Kohno, A. A. Levy, and H. M. Levy, "Vanish: Increasing Data Privacy with Self-Destructing Data," in *Proc. USENIX Security Symp. (USENIX Security'09)*, pp. 299–316, 2009.

[61] J. Backes, M. Backes, M. Dürmuth, S. Gerling, and S. Lorenz, "X-pire!- A Digital Expiration Date for Images in Social Networks," *arXiv preprint arXiv: 1112.2649*, pp. 1–22, 2011.

[62] M. Backes, S. Gerling, S. Lorenz, and S. Lukas, "X-pire 2.0 - A User-Controlled Expiration Date and Copy Protection Mechanism," in *Proc. ACM Symp. Applied Computing (SAC'14)*, pp. 1633–1640, 2014.

[63] Y. Shoshitaishvili, C. Kruegel, and G. Vigna, "Portrait of a Privacy Invasion," *Proc. Springer Privacy Enhancing Technologies Symp. (PETS'15)*, pp. 41–60, 2015.

[64] Y. Yang, S. Baker, A. Kannan, and D. Ramanan, "Recognizing Proxemics in Personal Photos," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR'12)*, pp. 3522–3529, 2012.

[65] I. Chakraborty, H. Cheng, and O. Javed, "3D Visual Proxemics: Recognizing Human Interactions in 3D from a Single Image," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR'13)*, pp. 3406–3413, 2013.

[66] Y. Nakashima, T. Ikeno, and N. Babaguchi, "Evaluating Protection Capability for Visual Privacy Information," *IEEE Security & Privacy*, vol. 14, no. 1, pp. 55–61, 2016.

[67] E. von Zezschwitz, S. Ebbinghaus, H. Hussmann, and A. De Luca, "You Can't Watch This! Privacy-Respectful Photo Browsing on Smartphones," in *Proc. ACM Conf. Human Factors in Computing Systems (CHI'16)*, pp. 4320–4324, 2016.

[68] R. McPherson, R. Shokri, and V. Shmatikov, "Defeating Image Obfuscation with Deep Learning," *arXiv preprint arXiv: 1609.00408*, pp. 1–12, 2016.

[69] E. M. Newton, L. Sweeney, and B. Malin, "Preserving privacy by de-identifying face images," *IEEE Trans. Knowledge and Data Engineering*, vol. 17, no. 2, pp. 232–243, 2005.

[70] T. Yamada, S. Gohshi, and I. Echizen, "Use of Invisible Noise Signals to Prevent Privacy Invasion Through Face Recognition from Camera Images," in *Proc. ACM Int'l Conf. Multimedia (MM'12)*, pp. 1315–1316, 2012.

[71] Y. Nakashima, T. Koyama, N. Yokoya, and N. Babaguchi, "Facial

Expression Preserving Privacy Protection using Image Melding," in *Proc. IEEE Int'l Conf. Multimedia and Expo (ICME'15)*, pp. 1–6, 2015.

[72] F. Li, Y. Wang, L. Yin, R. Xie, and J. Xiong, "Novel Cyberspace-Oriented Access Control Model," *Journal on Communications*, vol. 37, no. 5, p. 9–20 (in Chinese), 2016.
(李凤华, 王彦超, 殷丽华, 谢绒娜, 熊金波, "面向网络空间的访问控制模型", *通信学报*, 2016, 37(5): 9–20。)

李凤华 于 2009 年在西安电子科技大学计算机系统结构专业获得博士学位。现任中国科学院信息工程研究所副总工程师、研究员、博士生导师。研究领域为网络与系统安全、隐私计算、可信计算。Email: lfh@iie.ac.cn

吕梦凡 于 2015 年在哈尔滨工业大学信息安全专业获得学士学位。现在中国科学院大学网络空间安全学院计算机系统结构专业攻读硕士学位。研究领域为网络与系统安全。研究兴趣为移动系统中的数据安全与隐私保护。Email: lvmengfan@iie.ac.cn

孙哲 于 2012 年在中国科学技术大学软件工程专业获得硕士学位。现在中国科学院大学网络空间安全学院信息安全专业攻读博士学位。研究领域为信息安全、隐私保护。研究兴趣为图像隐私保护、隐私策略推荐。Email: sunzhe@iie.ac.cn

牛犇 于 2014 年在西安电子科技大学密码学专业获得博士学位。现任中国科学院信息工程研究所助理研究员。研究领域方向为网络安全、隐私计算。E-mail: niuben@iie.ac.cn