基于 QR 码隐写的物流隐私保护的系统

陶 静[1], 罗振豪[1], 王宝生[1], 邢倩倩[1]

[1]国防科技大学计算机学院, 长沙 中国, 410073

摘要 随着物流行业的不断发展, 快递已经在人类生活中扮演了重要角色。企业和个人通过快递可以方便地寄收物件。快递运输中通常使用 QR 码作为快递标签传递物流信息。然而, 未做隐私保护的快递标签可能会导致用户的物流隐私泄露。针对当前物流系统中对企业组织和个人隐私泄露的问题, 本文提出了一种基于 QR 码隐写和多级加密的物流隐私保护系统。在不改变 QR 码扫描结果和功能的条件下, 该系统利用 QR 码自身纠错功能实现了隐写加密数据到物流 QR 码上。本系统采用 AES 算法对隐私信息进行加密并使用 MD5 哈希作为密钥派生算法, 来实现多级加密和限制隐私访问权限; 使用 RSA 算法传递密钥, 确保加密传递的密钥只能被收件方解密。隐写了加密数据的 QR 码与原始 QR 码相比, 差异不大并且解码结果一致。这有效地从用户的角度保护了收件/寄件方的隐私并且隐藏了隐写数据的存在事实。本文首次针对自身隐私保密要求高的企业组织快递物流的隐私保护问题提出了有效的解决方案。实验部分显示了提出的系统在物流隐私保护方面的有效性。容量实验显示 QR 码最高可隐写 9720 比特的隐私数据, 同时鲁棒性实验表明隐写后的 QR 码能够抵抗不同的物理失真(例如, 自然纹理, 角度和距离)。为推动快递物流隐私保护的发展, 本系统代码已开源于 Gitee。

关键词 物流隐私; 隐写术; 二维码

中图法分类号 TP391.1 **DOI 号** 10.19363/J.cnki.cn10-1380/tn.2023.03.01

A Logistics Privacy Protection System Based on Quick-Response Barcodes Steganography

TAO Jing[1], LUO Zhenhao[1], WANG Baosheng[1], XING Qianqian[1]

[1] College of Computer, National University of Defense and Technology, Changsha 410073, China

Abstract With the continuous development of the logistics industry, express delivery has played an important role in society. Businesses and individuals can easily send and receive items through express delivery. In the process of express delivery, QR code is usually used as express labels to transmit express information. However, unprotected express labels may lead to the leakage of user privacy. Aiming at protecting the leakage of companies and personal privacy in the current logistics system, we propose and implement a logistics privacy protection system based on QR code steganography and multi-level encryption. Without changing the QR code scanning results, the system uses the error correction capability of the QR code to realize the steganography of encrypted data into the logistics QR code. The proposed system uses the AES algorithm as the encryption algorithm for private information, and the MD5 hash algorithm as the key derivation algorithm to achieve multi-level encryption and restricted privacy access rights; the RSA algorithm is used to ensure that the encrypted key can only be decrypted by the recipient. The QR code that hides the encrypted data is almost the same as the original QR code and the decoded result is the same, which effectively protects the privacy of the recipient/sender from the user's perspective and hides the existence of steganographic data. As far as we know, this is the first effective solution designed for companies that require high privacy and confidentiality to protect their sensitive logistics information. The experiments demonstrate the effectiveness of the proposed system in logistics privacy information protection. The capacity experiment shows that our system can accommodate up to 9720 bits of secret data, and the robustness experiment shows that the steganographic QR code is robust against different physical noises (e.g., natural texture, angle, and distance). To promote the development of express logistics privacy protection, we release the source code of our program on Gitee.

Key words logistics privacy; steganography; quick-response barcodes

1 引言

随着互联网技术的成熟及其应用的不断推广, 更多的行业与互联网进行深度融合, 形成了“互联网+”的发展模式。其中, 物流行业随着基于网络的电子商务的发展, 也发展迅猛并促使了现代物流的发展。

通讯作者: 陶静, 硕士, 副研究员, Email: ellen5702@aliyun.com。

收稿日期: 2021-10-12; 修改日期: 2022-02-11; 定稿日期: 2023-01-03

2020 年全年, 全国快递业务收入达 8795.4 亿元, 快递量达到了 833.6 亿件, 同比增长 31.2%, 平均物流快递企业每天处理 2.3 亿件物流包裹[1]。

现代物流能够在线追踪发出的货物, 规划投递路线和物流调度。实现这些功能的一项重要支撑技术就是条形码。快速响应码(Quick Response Code, QR 码)是现行主流的条形码。与传统文字传递信息不同, QR 码具有高容量、高纠错能力、便于机器定位和识别等优点。通过扫描物流快递上的 QR 码, 快递员可以快速获取物流快递的详细信息并进行分发运送, 点对点投递。

但是这种便利却可能导致用户的物流隐私发生泄露。对于个人而言, 在物流快递运输的任何一个环节中快递的标识(包括 QR 码)被他人获取, 都可能直接或间接导致寄件人和收件人的隐私泄露; 在收到快递后, 快递单的不正确保存和处理同样会导致寄件人和收件人的隐私泄露。对于组织、机构而言, 利用大量内部人员的快递物流信息泄露并结合开源情报(Open Source Intelligence), 攻击者能够分析出组织架构, 人员组成, 详细人员信息及偏好, 以及该组织的职能等敏感涉密信息。例如, 某组织长期接收 n 个不同收货人的物流快递, 而这些快递主要是来自电子零件加工厂以及计算机相关销售公司, 另外同一个地址包含 m 个子部门。攻击者利用这些长期泄露的组织人员个人隐私, 结合地图等开源情报能够快速推断出该组织是由 m 个部门组成大约 n 个人的从事计算机研究相关工作的组织。这些可能会对组织的隐私带来严重的威胁。

2018 年顺丰速运[2]提出了“隐址件”服务, 在寄/收件的整个流程中寄件和收件双方的关键个人信息会进行隐藏。这从用户端入手防止隐私泄漏风险。但是, 寄件和收件双方的隐私信息仍存储于物流公司, 物流快递服务端的隐私泄露风险并没有排除, 物理隐私保护依然面临着严峻的挑战。一种直观的方法[3]是将需要保密的包裹快递单上的信息进行加密后生成特制的快递标签。但是, 由于快递标签的明显差异, 这类方法直接让攻击者快速定位到保密的包裹。因此, 物流隐私保护急需一种能够嵌入隐私信息又不被察觉的方法。隐写术, 作为一种能够不让除预期接收者之外的任何人知晓隐藏信息的传递信息方法, 能够用于保护快递隐私信息不被他人察觉[4]。

隐写术被广泛应用于电子文档[5], 音频文件[6]以及图像[7-8]中, 并且主要存在于计算机虚拟世界中, 而在物理世界重放后, 其隐写内容则难以复原。传统的采用隐写墨水的隐写术[9]难以在物理世界和计算机世界间反复传递。在物流快递标签中进行隐写内容, 实际的物流环境和流程要求隐写的载体既能够存在物理世界中用于运输, 又能够存在计算机世界中便于提取隐写信息加快快递分发。而上述隐写术难以满足物流快递中的实际要求。

现有的研究[10]使用 JPEG 图像隐写算法利用 QR 码中黑白方块交界处的纹理信息嵌入隐写信息, 或者人为引入模糊和噪声, 采用中频系数归类统计的方法实现水印嵌入[11]。但是在复杂的物流运输环节中, 磨损、风化、污染都可能使得该纹理信息发生较大变化, 导致嵌入的加密地址信息无法提取而影响正常的物流运输。

为了从用户方的角度解决物流隐私保护的问题, 本文提出了一种基于 QR 码隐写的物流隐私保护方案并实现了终端软件和管理系统, 适用于对自身隐私要求级别高企业组织, 例如军民融合的物流系统。该系统实现了一种轻量级, 抗物理磨损, 分级加密, 但不影响原有 QR 码正常使用的物流隐私保护技术。它首先使用压缩和分级加密技术对敏感信息进行处理, 然后利用 QR 码的本身纠错功能, 在不改变 QR 码内容的情况下将用户隐私信息嵌入快递面单上的 QR 码中。与“隐址件”不同, 敏感信息隐写在 QR 码本身, 存储于寄件/收件企业内部, 而不是物流公司, 这能够保护因物流公司和运输途中导致的隐私泄露。另外, 隐写在 QR 码的敏感信息进行了多级加密, 这可以限制敏感信息访问范围, 防止敏感信息的知情的范围扩大。由于基于 QR 码的纠错功能, 因此在 QR 码可识别的条件下, 可以正常提取隐写内容, 可用于常规的物流运输。

该系统主要由隐写密文生成环节, 订单 QR 码生成环节, 物流快递配送环节以及隐写内容提取和快递分发环节等 4 个环节组成。其中, 隐写密文生成环节和订单 QR 码生成环节有发件组织完成; 物流快递配送环节与正常的物流派送无异; 隐写内容提取和快递分发环节由收件组织完成。

2 背景知识和相关工作

2.1 QR 码

QR 码[12]是由规则排列的深、浅色矩形像素构成的二维条码。其中, 深色像素表示 1, 浅色像素表示 0, 构成一段二进制序列。QR 码目前有 40 个版本, 规格从 21×21 像素至 177×177 像素。QR 码采用 RS 码(Reed-Solomon Codes)实现纠错功能[13]。它的纠错级别(Error Correction Level, ECL)分为 4 种(L, M, Q 和 H), 分别表示约 7%、15%、25%和 30%的字码可

被破坏(例如, 褶皱, 破损以及涂鸦), 仍可以正常提取 QR 码中的数据。由于 QR 码是由大量无语义的黑白像素组成, 因此修改少量的 QR 码像素, 但不影响 QR 码正常解码的情况下(QR 码解码结果不变), 其他人难以察觉。本文提出的基于 QR 码隐写的物流隐私保护系统是利用 QR 码的纠错功能, 通过修改少量像素并且不影响 QR 码正常解码的情况下实现敏感数据的信息隐藏。

2.2 隐写术

隐写术是一门关于信息隐藏的实践与技术。它是利用特定的手段将目标信息隐藏, 不让除预期的接收者之外的任何人知晓目标信息的传递的事实和目标信息的内容。

传统的隐写术采用隐形墨水[9],[14]等方法实现信息隐藏和情报传递。现有的隐写术被广泛应用于计算机数字载体中, 例如: Liu 等人[5]提出了一种利用变化跟踪技术在 Microsoft Word 电子文档中隐藏数据的隐写方法。而 Binny 等人[6]提出一种使用基于 LSB (Least Significant Bit)的算法在音频中嵌入文本信息的隐写技术。在图像方面, Lu 等人[15]提出了一种基于 LSB 匹配方法的双图像嵌入技术实现图像隐写。然而该类方法难以用于打印图像上。Tkachenko 等人[16]基于 QR 码提出了 2LQR 的隐写术, 该方法通过用特定的纹理图案替换 QR 码中的黑色模块来嵌入隐写信息。尽管 2LQR 能够实现在打印 QR 码图像上隐写内容, 但是由于其替换了 QR 码中的黑色模块, 使得使用 2LQR 隐写后的 QR 码与正常 QR 码差异较大, 易被发现, 难以实现隐藏含有隐写信息的事实。Lin 等人[17]提出了一种利用 QR 码的纠错能力的隐写术, 该方法不会修改条码的可读性并且可应用于打印图像上。此外, Tancik 等人[18]提出了利用神经网络实现的一种在打印的图像中隐藏数据的隐写技术 Stegastamp。其优势在于可以在任意图像中隐写数据, 并且能够通过扫描设备识别提取, Tancik 也指出该方案对 7 字节的数据隐写是具有鲁棒性的。但是在快递物流中, 7 字节的隐写容量过小, 难以满足物流中寄/收件方复杂的隐写内容需求。

2.3 物流模型及隐私保护问题

当前国内的物流模型如图 1 所示。当用户需要邮寄快递包裹时, 通常需要提供寄件方和收件方详细地址、个人身份信息、联系方式以及快递内容信息给快递员。在经过运输和层层分拣后, 快递包裹交付给收件方。为了便于快递包裹的投递, 收件方地址可能会尽可能的详细, 甚至具体到单位部门、门牌号、收件人具体职务等。然而这些敏感信息的泄露可能对企业和组织的隐私造成严重的威胁。

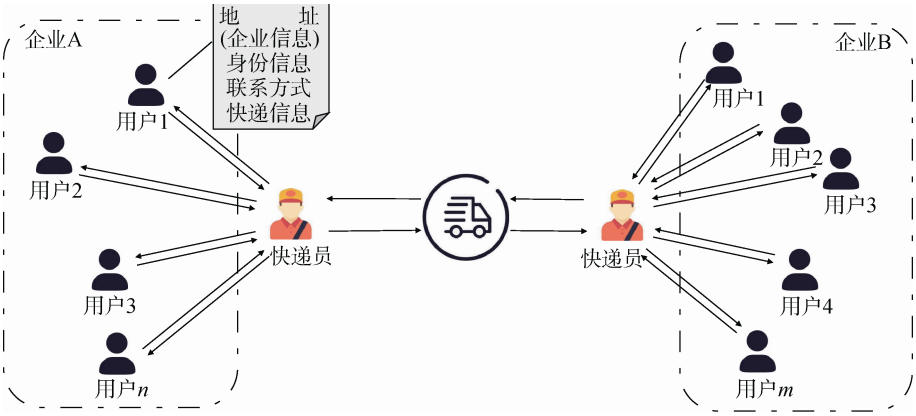

图 1 物流简化模型图

Figure 1 An example of Logistics simplified model

在图 1 中, 快递包裹邮寄的过程中主要参与方有企业中的寄件用户, 收件用户, 快递员以及快递公司。以下情况都可能造成企业组织的隐私信息被泄露:

(1) 寄件用户在填写快递订单时误填敏感信息;

(2) 收件用户在收到快递包裹后未妥善处理快递订单, 快递订单上信息可能会被恶意窃取;

(3) 快递员在运输快递时由于疏忽, 快递包裹信息被他人恶意窃取;

(4) 存在恶意快递人员利用职务之便故意泄露快递信息;

(5) 快递服务公司因自身缺陷或被黑客渗透造成快递信息的泄露, 或者快递服务公司本身存在该恶意的泄露行为。

面对快递包裹隐私泄露的问题, 快递服务公司对快递隐私信息的提出附加了权限要求。只有具备了相应权限, 才能读取快递上的包裹信息, 并且敏感信息内容用“*”代替。顺丰速运[2]提出的“隐址

件”服务, 在寄/收件的整个流程中寄件和收件双方的关键个人信息会进行隐藏。这可以解决上述情况中的(1)(2)(3)。但是对于情况(4)(5), 这些方法无法防止由于快递员和快递公司的原因造成的企业组织信息泄露。而刘亮等人[3]提出了一种基于时效控制加密的 QR 码物流隐私保护方案, 其加/解密模块控制在用户方, 可缓解(4)(5)中快递公司造成的快递隐私泄露。但是, 该方案生成的特制快递标签与原本快递标签存在差异, 容易引起攻击者的注意。严文博等人[10]提出了用 JPEG 图像隐写算法将敏感信息嵌入 QR 码中。在快递配送时, 需要快递员使用特定 APP 从云端获取具体的收件人信息, 这有效地防止了由于快递订单引发的隐私泄露。虽然能够避免引起攻击者注意, 但是要求快递人员配合安装特定的 APP, 难以推广。此外, 这种方法难以防止恶意的快递人员利用权限之便故意泄露快递的隐私信息。

3 物流隐私保护系统

本文提出了一种基于 QR 码隐写的物流隐私保护系统。该方法适用于对自身隐私要求级别高的企业组织。各成员企业组织使用 RSA 生成公钥和私钥, 并公布公钥。隐私信息使用 AES 加密, 加密的密钥传递使用收件方公钥进行加密通过密钥管理平台传递。密文的隐写是利用 QR 码纠错功能, 在不改变 QR 码本身功能的情况下把密文隐写到 QR 码中。

3.1 隐私保护方案框架

物流隐私保护方案框架主要有四个环节组成: QR 码生成环节, 物流快递配送环节以及隐写内容提取和快递分发环节。

下面具体介绍隐私保护方案的环节流程, 以图 2 中企业 A 的寄件用户 1 给企业 B 的收件用户 2 寄件为例。其中, 用户、部门和企业采用非对称加密算法生成公钥和私钥, 并预先在密钥管理平台上公开自己的公钥。

首先是 QR 码生成环节:

①企业 A 寄件用户 1 将寄件人的地址、联系电话等相关信息使用密钥k_1进行 AES 加密后得到密文 c_1, 并将收件人的地址、联系电话、密文 c_1、经过派生的密钥k_2、包裹信息和快递包裹移交给上级部门, 使用收件人公钥对密钥 k_1 和包裹编号加密后提交至密钥管理平台。

②上级部门收到寄件用户 1 提供的包裹和信息后, 对收件人、联系电话、以及密文 c_1 等信息采用密钥 k_2 进行加密得到密文 c_2, 并将收件部门、密文 c_2、经过派生的密钥 k_3、包裹信息和快递包裹移交给企业组织 A 的收发室, 同时使用收件部门的公钥对密钥 k_2 和包裹编号加密后提交至密钥管理平台。

③收发室在收到部门提供的快递包裹和提供的信息后, 使用密钥 k_3 对寄件部门、收件部门、密文 c_2 进行加密得到密文 c_3, 并使用收件企业 B 的公钥对密钥 k_3 和包裹编号加密后提交至密钥管理平台。然后, 企业 A 提供其统一的对外联系电话和对外公开地址以及邮寄包裹必需的信息提供给快递公司, 由快递公司生成 QR 码 Q。收发室对 QR 码 Q 进行解析, 获取 QR 码的配置信息, 包括版本(1 到 40), 容错级别(L, M, Q, H)以及掩模。根据 QR 码 Q 的配置, 再以相同的配置生成 QR 码的过程中, 使用 Reed-Solomon 编码对密文 c_3 生成纠错码, 再与数据字节进行异或处理, 最后生成含有隐写信息的 QR 码 Q', 并附于快递包裹上, 交付给快递员。

下一环节是物流配送环节:

在这一环节中, 由于本文提出的隐私保护方案采用的是将加密敏感数据隐写到物流运输的 QR 码中, 没有改变 QR 码的解码结果。因此对于本隐私保护方案, 物流快递公司无需额外调整其物流配送中的任何操作, 与正常的物流配送一致。这也是本方案能够广泛推广的优势所在。

下一步是隐写内容提取环节:

①企业组织 B 的收发室在收到快递包裹后, 扫描快递包裹上的 QR 码 Q', 获得 QR 码的文本信息 t 和配置信息, 包括版本(1 到 40), 容错级别(L, M, Q, H)以及掩模。

②用与 QR 码 Q'相同的版本、容错级别、掩模和文本信息 t 生成 QR 码 Q_1, 此时 Q_1 理论上与寄件时的 QR 码 Q 一致。

③将 QR 码 Q_1 与 QR 码 Q'进行异或处理, 生成异或的结果 c_4, 此时 c_4 应与企业组织 A 的收发室嵌入的密文 c_3 一致。该提取过程可通过特定的手机或其他扫描设备上的应用程序自动化提取。

最后是快递分发环节:

①企业 B 从密钥管理平台下载加密的密钥, 用自身私钥进行解密后得到密钥 k_3 和编号, 并使用密钥 k_3 对具有相同编号包裹的 c_4 进行解密得到收件部门信息。然后, 将快递包裹分发给收件部门。

②收件部门在收到快递包裹后, 通过密码管理平台获取加密的密钥并用自身私钥解密得到密钥 k_2 和编号。对密钥 k_2 使用密钥派生函数可得到 k_3, 使用 k_3 对具有相同编号包裹的密文 c_4 进行解密得到密文 c_2, 再使用密钥 k_2 对密文 c_2 解密, 得到收件用户 2, 联系电话, 然后通知收件用户 2 取件。

③企业B的收件用户2收到取件通知后, 到收件部门取件。通过密钥管理平台获取加密的密钥并用自身私钥解密得到密钥k_1和编号。对密钥 k_1 使用密钥派生函数可得到 k_2, 对密钥 k_2 使用密钥派生函数可得到 k_3。然后使用 k_3 对具有相同编号包裹的密文 c_4 进行解密得到密文 c_2, 使用 k_2 对得到的密文 c_2 进行解密得到密文 c_1, 再使用密钥 k_1 对密文 c_1 解密, 从而得到了寄件人地址、联系电话等信息。值得注意的是, 在处理包裹包装时, 即使被他人获取到快递的 QR 码, 由于敏感隐私信息已被加密并且隐写, 他人也难以察觉 QR 码中的敏感隐私信息。这能够有效地保护企业组织邮寄包裹时的隐私安全, 并且能够防止物流公司泄露企业快递隐私信息。

由于密文任何一个比特改变都会导致无法解密, 考虑到物流过程和扫码过程很有可能误识别 QR 码, 因此在扫码阶段对进行多次采样, 避免偶然因素。此外, 在本物流隐私保护方案中 QR 码隐写过程中, 采用了 Reed-Solomon 编码对密文 c_3 生成纠错码, 以便对识别错误的密文进行纠错。

3.2 多级加密设计

本节对本文提出的隐私保护系统中密钥生成过程进行解释。本文使用密钥管理平台对密钥进行管理传递, 使用多级加密对敏感隐私信息进行保护, 其中隐私信息加密采用对称加密, 对称加密的密钥加密采用非对称加密的公钥加密, 密钥的派生采用哈希算法。

(1) 密钥管理平台

密钥管理平台中的每一位用户、部门和企业首先采用非对称加密算法生成公钥和私钥, 并在密钥管理平台上公开自己的公钥。在具体实现中选取 RSA 算法作为非对称加密算法来验证。寄件用户、部门和企业在邮寄包裹时, 通过密钥管理平台获取收件用户、部门和企业的公钥对对称加密的密钥进行加密并发送给相对应的收件用户、部门和企业。

(2) 多级加密

本文针对物流标识中的隐私信息进行分层保护, 采用多个密钥对物流单面的多种信息进行隐私保护, 其保护的内容和对应的隐私保护密钥关系如下表。

邮寄快递包裹时通常需要寄件人的姓名、联系电话、通讯地址等。而这些信息往往可能包含企业组织内部的敏感隐私信息, 例如人员组成、部门职责和详细地址等。在整个快递邮寄正常过程中, 通常不需要用到寄件方信息。当收件方需要退回快递包裹时, 才需要使用到寄件方信息。因此寄件方信息优先加密, 仅收件人可查看。在快递邮寄时, 提供收件方的对公的收发室地址即可满足快递公司方的运输快递的要求, 因此, 收件方的对公地址作为公开地址, 收件部门属于需要保护的信息, 仅在收发室需要分发包裹时, 可以解密提取。而收件人信息, 是在收件部门通知收件人取件时可解密提取。

表 1 多级加密的密钥配置

Table 1 Key configuration for multi-level encryption

单面信息	对应密钥	作用
寄件方的姓名、联系电话和通讯地址	k_1	保护发件方个人隐私
收件人的姓名和联系电话	k_2	保护收件方个人隐私
收件企业部门和该部门的联系方式	k_3	保护收件企业部门隐私

根据上述场景, 物流隐私保护方案的多层加密需满足基本的单向安全性, 即上层密钥能够解密下层密钥加密的内容, 但是下层密钥无法解密上层密钥加密的内容。例如, 隐私保护密钥k_2是k_1经过密钥派生方法计算得到的二层密钥, 隐私保护密钥k_3是k_2经过密钥派生方法计算得到的三层密钥:

$$k_2 = KDF(k_1) \tag{1}$$

$$k_3 = KDF(k_2) \tag{2}$$

持有密钥k_1可以解密用密钥k_2或k_3加密的密文, 但是密钥k_2和k_3无法解密用k_1加密的密文。这要求派生KDF函数具有难以反向计算的特性, 在具体实现中选取低碰撞率和难逆向的MD5算法作为KDF函数来验证。由于多级加密的设计, 所属的不同级别无法查看上级的加密信息, 这保证了敏感隐私信息的不扩散和不外泄。

(3) 物流信息加密

本文采用 AES 加密作为物流隐私保护方案中物流信息对称加密算法, 实现信息加密封装, AES 加密与 DES 和 3DES 加密相比, 具有更快的运算速度和更高的安全性。针对上述要求, 多级加密可设计为如式(1)所示。

$$\begin{aligned} c_1 &= E_{AES}(m_1, k_1) \\ c_2 &= E_{AES}(c_1 + m_2, k_2) \\ c_3 &= E_{AES}(c_2 + m_3, k_3) \end{aligned} \tag{3}$$

其中m_1中对应寄件方的姓名、联系电话和通讯地址, m_2对应收件人的姓名和联系电话, m_3对应收件部门及其联系方式, k_2 由 k_1 经过哈希得来, 而 k_3 由 k_2 经过哈希得来。因此已知 k_1 能够快速生成 k_2 并解密由 k_2 加密的数据。由于哈希的不可逆性, 因此已知 k_2 难以计算出 k_1 进行解密, 保证了由 k_1 加密数据的安全性。经过 AES 对称加密信息后, 使用 Reed-Solomon 编码的c_3即为待嵌入的加密隐写数据。

在本物流隐私保护方案中, 各企业组织的公钥和私钥是预先生成, 而用于 AES 加密的 k_1 是在邮寄

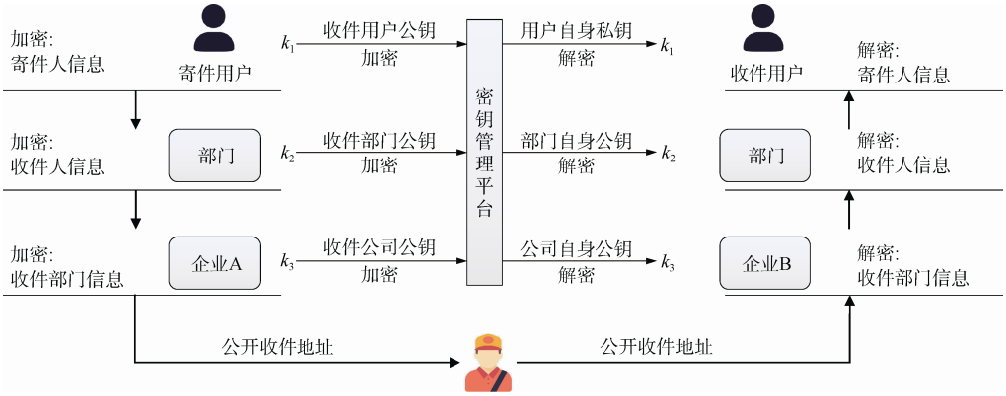

图 2　物流隐私保护框架

Figure 2　The framework of logistics privacy protection system

包裹时寄件用户 1 临时随机生成, 密钥 k_2 和 k_3 由 k_1 派生生成的。本文采用 MD5 算法作为密钥派生算法。该算法可以为每个输入到哈希算法的数据产生出一个 128 位(16 字节)的散列值。在实际的使用中, MD5 的输出结果会把 16 字节的散列值, 例如“0xecd9 0a314ce703839458db9296140010”的形式, 表示成“ecd90a314ce703839458db9296140010”的可见字符串形式, 所以 MD5 的输出结果又可看作 32 字节的字符串。为了保持 k_1、k_2 和 k_3 的长度一致以便于哈希计算, 本文中 k_1 的长度设置为 128 位, 用 32 字节可见字符串保存。

3.3　隐写密文设计

为了提高敏感隐私信息的隐蔽性, 本文提出隐写密文的设计。它利用 QR 码的纠错能力将数据隐写到 QR 码中。隐写后的 QR 码的使用和数据的扫描读取与正常的 QR 码无差别, 通过工具扫描结果与未隐写数据的 QR 码一致。QR 码生成过程如下所示:

首先, 将输入的文本进行转换生成数据码(Data Code), 若未达到该 QR 码版本的数据上限, 还会生成补齐码(Padding Bytes)进行填充。然后根据生成的数据码和设定的容错级别和生成纠错码(Error Correction Code)。

第二步, 把数据码和纠错码穿插放置, 并生成最终码(Final Bits)。QR 码的最终码排布如图 3 所示。

图 3 中是版本为 8, 容错级别为 H 的 QR 码数据排布图。在生成 QR 码时, 输入的文本被转成二进制流, 然后按照 8 比特为一组从右下角按照图中色块的排列方式开始排列, 其中每一个色块包含 8 比特的数据。

第三步, 为了防止由于最终码的原因导致 QR 码出现大面积的空白或黑块而增加扫描识别的困难, QR 码的生成过程中还会对 QR 码计算掩模, 然后将掩模和 QR 码进行异或处理。

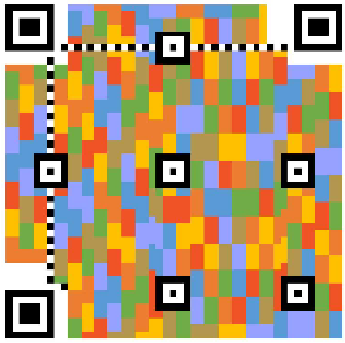

图 3　QR 码数据流布局

Figure 3　The layout of QR code data flow

第四步, 将 QR 码的配置格式信息, 包括版本、容错级别和掩模编号等信息都嵌入到 QR 码固定位置, 便于 QR 码解码。最终得到常规的 QR 码。

给定任意 QR 码, 密文隐写到 QR 码过程如下:

首先是提取提定 QR 码的文本信息和配置格式信息, 包括版本, 容错级别, 掩模编号。

然后重复生成 QR 码的前两步, 直到生成最终码。由于最终码已包含纠错码, 修改阈值以内的最终码实现来密文隐写不会影响 QR 码的正常解码。本文采用将最终码与待嵌入的加密隐写数据 c_2 进行异或处理来实现隐写数据的嵌入。此时, 隐写的长度小于最大纠错阈值时, QR 码即可正常解码。其中版本为 40, 容错级别为 H 的 QR 码的最大纠错阈值为 9720 比特。

再将隐写了密文的最终码按照 QR 码默认规则进行排布, 并与给定编号的掩模进行异或处理。

最后将 QR 码的配置格式信息, 包括版本、容错级别和掩模编号等信息都嵌入到规定位置, 得到含有隐写密文的 QR 码。

本文中, 从含有隐写数据 QR 码中提取隐写数据

如下:

首先正常解码 QR 码 Q, 得到 QR 码的版本 v、容错级别 e, 掩模 *mask* 以及文本 t。

然后, 以版本 v、容错级别 e, 掩模 *mask* 作为配置信息, 文本 t 作为输入, 按照正常流程生成 QR 码 Q'作为对比 QR 码, 该 QR 码 Q'不包含错误码。

将 Q'和 Q 进行比对判断是否含有隐写信息, 如果存在隐写内容, 将 Q'和 Q 进行异或处理得到隐写的加密数据 c_2。最后, 利用提供的密钥进行解密即可得到隐写加密数据的原文。

4 实验设计与结果分析

4.1 实验设置

为了验证本文设计的基于 QR 码隐写的物流隐私保护方案的性能, 本文使用 Google 开源库 ZXing[19]实现 QR 码图像的编码和解码, 对称密钥的哈希算法采用 MD5, 并按照第三章流程编写多级加密隐写和解密部分。

4.2 快递信息的隐私保护效果

利用多级加密能够有效地对隐写的敏感隐私数据实现分级的加密访问控制。低权限级别的密钥无法解密高级别密文, 这样可以最大限度地保护用户的快递寄件时的隐私信息。表 1 是对邮寄快递包裹信息进行加密的过程。寄件方: 湖南省长沙市国家技术中心研发部密码教研室, 寄件人: 张三, 联系电话: 13007311234; 收件方: 安徽省合肥市高性能研究所电子工程部精密仪器教研室, 收件人: 李四, 联系电话: 13005511234。

表 2 多级加密明文和密钥

Table 2 Plaintext and keys for multi-level encryption

符号	文本信息	密钥
m_1	湖南省长沙市国家技术中心研发部密码教研室, 张三, 联系电话: 13007311234	k_1= ecd90a314ce703839458db9296140010
m_2	李四, 电话: 13005511234	k_2= e755e2d9f9325d06a18c1c0cf7e07287
m_3	安徽省合肥市高性能研究所电子工程部精密仪器教研室	k_3= 0387809b34d8221026034de1882b8e68
明文	寄件: 湖南省长沙市开福区福元路 1 号 电话: 0731-1234567 收件: 安徽省合肥市蜀山区黄山路 460 号 电话: 0551-1234567	无

(a) 原QR码 (b) 载有密文的QR码

(a) Original QR code (b) QR code with secret data

图 4 物流运输时 QR 码标签

Figure 4 An example of QR code with secret data

张三首先将寄件人信息 m_1 用密钥k_1 加密提交给上级部门; 上级部门将收件人具体信息m_2用k_2(k_2 =MD5(k_1))进行加密并移交给企业收发室; 收发室将收件部门信息。

m_3用k_3(k_3=MD5(k_2))加密。表 1 例子中, 密文c_3的大小为 64 字节, 加入 Reed-Solomon 编码纠错码(纠错率 10%)为 80 字节。将快递包裹与公开明文信息交给快递公司, 得到 QR 码标签如图 4 (a)所示, 其扫描结果即为明文信息“寄件: 湖南省长沙市开福区福元路 1 号, 电话: 0731-1234567; 收件: 安徽省合肥市蜀山区黄山路 460 号, 电话: 0551-1234 567”; 收发室根据 QR 码嵌入密文, 得到 QR 码如图 4(b)所示。

与原始收件信息“安徽省合肥市国防科技大学电子工程部精密仪器教研室, 收件人: 李四, 联系电话: 13005511234”相比, 经过处理后的收件信息“安徽省合肥市蜀山区黄山路 460 号, 电话: 0551-1234567”在满足正常物流运输的情况下, 敏感隐私更加安全, 更不容易泄露敏感信息。而图 4 中原始 QR 码图 4(a)和隐写加密数据后的 QR 码图 4(b)常规扫码结果, 均为“寄件: 湖南省长沙市开福区福元路 1 号, 电话: 0731-1234567; 收件: 安徽省合肥市蜀山区黄山路 460 号, 电话: 0551-1234567”, 无差异。

收件单位收发室收到包裹后, 利用本文提供的工具从含有隐写信息的 QR 码(图 4(b))提取密文c_4, 并使用收件单位的私钥对密码管理平台获得密钥密文解密得到的密钥k_3对c_3解密, 得到解密信息“安徽省合肥市高性能研究所电子工程部精密仪器教研室”, 并将包裹分发给相应部门。相应部门使用部门私钥对密码管理平台获得密钥密文解密得到的密钥k_2对c_4解密得到解密信息“李四, 电话: 13005511234”,

表 3　载密 QR 码密文提取信息

Table 3　The privacy data extracted from the QR code

密文	密文内容	文本信息
c_3	eb02ba55ce144d7055539057611b64a6df3d9492fc28035abe5a7f77a038eb5f27224ae024f09f333bd81422f9349bbe80c11dfb4ae7f26bf89d73a590b70a29	安徽省合肥市高性能研究所电子工程部精密仪器教研室
c_2	71a1b617ba746b4b09ed6f4036bddb238ceeca803821fdcdbd04c45371b16208449247efead5ba196cb4010616e497b0	李四, 电话: 13005511234
c_1	b3fa5b1897189e62d4482b9f0fd914ba662cc878b3d8ed1d1d3b62ccd9571318	湖南省长沙市国家技术中心研发部密码教研室, 张三, 联系电话: 13007311234

然后联系收件人取件。收件人收到包裹后, 使用用户个人私钥对密码管理平台获得密钥密文解密得到的密钥k_1对c_4得到寄件人信息“湖南省长沙市国家技术中心研发部密码教研室, 张三, 联系电话: 13007311234”。

由于 MD5 哈希的不可逆性, k_i可以解密$k_j(j>i)$加密的信息, 而k_j无法解密k_i加密的信息, 在本实验中李四可以通过k_1解密密文c_4所有加密信息, 持有k_2只能解密k_2和k_3加密的信息, 而无密钥人员仅能扫描出 QR 码上非敏感信息。这有效地防止企业组合和个人的隐私外泄。

4.3　容量测试

作为用于物流的隐私保护系统, 其需要隐写的敏感信息可能涉及多方面内容, 因此要求方案满足能够在打印图像上隐写较大的数据容量。Tancik 等人[18]提出的 Stegastamp 虽然能够在打印图像上实现数据隐写, 但是数据容量上难以满足需求。本文对提出的物流隐私保护系统进行了容量测试实验。

实验分别对 L、M、Q 和 H 四个纠错级别, 从 1 到 40 不同版本的 QR 码进行隐写数据测试。为了确保实验的可靠性, 实验中用随机生成的数据进行隐写。图 5 显示了本文提出的物流隐私保护系统的容量测试结果。

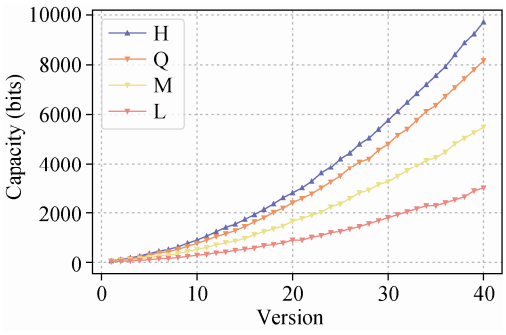

图 5　容量测试

Figure 5　The results of capacity experiment

图 5 可以看出不同的纠错级别的 QR 码隐写内容的容量不同。其中纠错级别为 H, 版本 30 的 QR 码隐写容量最高可达 5760bits, 版本为 10 的 QR 码隐写容量为 896bits。不同的纠错级别同样会影响 QR 码隐写的容量。版本为 40, 纠错级别为 H 的 QR 码最高可隐写 9720bits 的数据; 即使是纠错级别 L 隐写内容依然可达 3000bits。

通过容量测试实验, 可知本文提出的物流隐私保护系统在打印图像上的隐写数据容量上远高于 Tancik 等人[18]提出的 Stegastamp, 能够满足物流中寄/收件方复杂的隐写内容需求。

4.4　载密 QR 码物理鲁棒性实验

在快递物流运送分发过程中, 快递 QR 码标签可能会破损或褶皱。因此物流隐私保护系统需要满足在多种复杂环境中都能正常工作的苛刻要求。

根据严文博等人[10]在实验中提及采用标准测试图像 Lena[20]作为纹理生成的自然图像来模拟载密 QR 码破损或褶皱的情况, 生成如图 5 所示。图 5(a)为含有隐私密文的 QR 码, 图 5(b)则是图 5(a)用 Lena 图像作为自然纹理叠加后的图像。

使用峰值信噪比(PSNR) 对图 5 中图像(a)和(b)进行失真评价。PSNR 的计算方法如式(3)所示:

$$PSNR = 10\lg\frac{(2^n-1)^2}{MSE} \tag{4}$$

其中 *MSE* 是图像之间像素的均方误差, 像素的比特位深度 n 取 8。PSNR 表示图的失真量, 取值在 0 到 100dB。PSNR 的值越小, 说明图像失真越大。经过计算, PSNR 的值为 27.40dB, 说明图 5(b)与图 5(a)存在较大失真, 但是依然能够正常解码 QR 码和提取并解密 QR 码携带的加密隐私数据。这说明, 本文提出的基于 QR 码隐写的物流隐私保护系统具有较强的鲁棒性, 能够在复杂的物流环境中正常工作。

此外, 在实际操作中, 扫描人员通常难以保证以固定角度(与法线夹角 0°), 固定距离扫描 QR 码。因此, 为了模拟实际使用场景, 实验还采用不同角度和距离的打印图片进行解密测量鲁棒性。在本实验中, 我们以版本为 20, 不同纠错级别(ECL), 不同数量隐写内容, 不同角度, 和不同距离的 QR 码进行测

(a) 载密QR码
(a) QR code with secret data

(b) 纹理增强的载密QR码
(b) Natural texture-enhanced QR code

图 6　自然纹理对载密 QR 码图像的影响

Figure 6　The impact of natural texture on the QR code

量实验。根据第 **4.3** 节中容量测试可知, 版本 20 的 QR 码, 纠错级别 L 可容纳 896bits, M 可容纳 1664bits, Q 可容纳 2400bits, H 可容纳 2800bits 的隐写内容。

如表 4 所示, 实验分别从角度、距离对打印 QR 码进行测试。其中角度是与 QR 码平面法线所成角度,

表 4　关于角度和距离的鲁棒性实验结果

Table 4　Robustness test results on angles and distances

QR 码		角度		距离	
		0°	15°	10 cm	30cm
版本 20 ECL: H	隐写 0bits	能	能	能	能
	隐写 512bits	能	能	能	能
	隐写 1024bits	能	能	能	能
版本 20 ECL: Q	隐写 0bits	能	能	能	能
	隐写 512bits	能	能	能	能
	隐写 1024bits	能	能	能	能
版本 20 ECL: M	隐写 0bits	能	能	能	能
	隐写 512bits	能	能	能	能
	隐写 1024bits	能	能	能	能
版本 20 ECL: L	隐写 0bits	能	能	能	能
	隐写 512bits	能	能	能	能
	隐写 1024bits	否	否	否	否

距离是指与 QR 码平面的距离。实验结果显示在 QR 码的隐写容量范围内, QR 码内容和隐写数据在不同角度和不同距离可正常提取, 满足快递物流的实际的使用环境需求。版本 20, ECL 为 L 的 QR 码隐写内容 1024bits 无法解密, 是因为其超过了最大对应版本和纠错级别的最大容量(896bits)。实验可知, 本文提出的方法在隐写数据小于最大隐写容量时, 对角度和距离具有鲁棒性, 满足当前快递物流的使用需求。

4.5　与现有工作对比

本系统与顺丰速运提出了“隐址件”服务相比, 优势在于本系统能够将隐私数据掌握在自身企业组织手里, 能够有效地防止收件/寄件企业组织的隐私信息因快递环节或快递公司的原因泄露。另外, 本文提出的物流隐私保护系统可直接应用于各类快递公司运输, 快递公司无需修改其本身原有的快递流程, 便于隐私保护系统的推广同时也避免了快递运输选择的单一性。

与刘亮等人[3]提出的方案和严文博等人[10]提出的方案相比, 本方案最大的优势在于无需对快递公司和快递运输环节进行任何的更改。这为本方案的推广提供巨大的便捷。本方案利用 QR 码本身的纠错功能在不改变 QR 码扫描结果的条件下, 隐写加密的隐私数据。并且原 QR 码与载有密文的 QR 码差异不大, 隐写加密数据不易被发现, 具有良好的隐私性和保密性。

此外, 本方案能够在不同角度, 不同距离的实际环境中均能正常识别和解密, 这有利于本方案应用于扫描环境复杂多变快递运输环节中。

5　结束语

本文针对现有物流环境中企业组织的敏感隐私数据和个人隐私数据泄露问题和现有提出的快递隐私保护方案的不足, 提出了一种基于 QR 码隐写的物流隐私保护方案, 并详细描述了本方案的系统架构、各环节组成和实际应用步骤。本方案利用 QR 码自身的纠错功能, 实现了在不改变 QR 码扫描结果和使用功能的条件下, 隐写加密隐私数据到物流 QR 码上。隐写加密数据后的 QR 码与正常 QR 码相差不大, 且不易被发现。此外, 本方案采用 MD5 哈希与 AES 加密结合, 实现多级加密, 限制隐私的访问权限, 防止隐私的向下扩散问题; 使用 RSA 非对称加密算法传递密钥, 实现加密信息的仅能收件方解密的功能。实验部分详细验证了本方案对因自然纹路、角度、距离、光强而图像失真的鲁棒性。

本方案首次为当前对自身隐私保密要求高的企业组织快递物流的隐私保护问题提出了有效的解决方案。为推动快递物流隐私保护的发展, 本方案代码已开源于 Gitee①。

① https://gitee.com/allen-runner/QRCodeWithSecret

参考文献

[1] State Post Bureau of the People's Republic of China. 2020 China Express Development Index Report [R]. 2021.
(中华人民共和国国家邮政局. 2020 年中国快递发展指数报告[R]. 2021.)

[2] SF Express. https://www.sf-freight.com/. Jun. 2021.

[3] Liu L, Guo W B, Yang Y W, et al. Research on QR Code Logistics Privacy Based on Segmented Encryption and Time-Limited Control[J]. *Chinese Journal of Network and Information Security*, 2019, 5(4): 63-70.
(刘亮, 郭文博, 杨昱威, 等. 基于分段加密和时效控制的QR码物流隐私保护方案[J]. *网络与信息安全学报*, 2019, 5(4): 63-70.)

[4] Artz D. Digital Steganography: Hiding Data within Data[J]. *IEEE Internet Computing*, 2001, 5(3): 75-80.

[5] Liu T Y, Tsai W H. A New Steganographic Method for Data Hiding in Microsoft Word Documents by a Change Tracking Technique[J]. *IEEE Transactions on Information Forensics and Security*, 2007, 2(1): 24-30.

[6] Binny A, Koilakuntla M. Hiding Secret Information Using LSB Based Audio Steganography[C]. *2014 International Conference on Soft Computing and Machine Intelligence*, 2015: 56-59.

[7] Hamid N, Yahya A, Ahmad R B, et al. Image steganography techniques: an overview[J]. *International Journal of Computer Science and Security*, 2012, 6(3): 168-187.

[8] Kadhim I J, Premaratne P, Vial P J, et al. Comprehensive Survey of Image Steganography: Techniques, Evaluations, and Trends in Future Research[J]. *Neurocomputing*, 2019, 335: 299-326.

[9] Macrakis K, Bell E K, Perry D L, et al. Invisible Ink Revealed: Concept, Context, and Chemical Principles of "Cold War" Writing[J]. *Journal of Chemical Education*, 2012, 89(4): 529-532.

[10] Yan W B, Yao Y Z, Zhang W M, et al. Privacy-Preserving Scheme for Logistics Systems Based on 2D Code and Information Hiding[J]. *Chinese Journal of Network and Information Security*, 2017, 3(11): 22-28.
(严文博, 姚远志, 张卫明, 等. 基于二维码和信息隐藏的物流系统隐私保护方案[J]. *网络与信息安全学报*, 2017, 3(11): 22-28.)

[11] Liu Z M, Lin J J. A Watermarking Algorithm for Print-Scan Process on QR Code[J]. *Journal of East China University of Science and Technology* (*Natural Science Edition*), 2018, 44(1): 97-103.
(刘子鸣, 林家骏. 一种适用于 QR 码的抗打印扫描的水印算法[J]. *华东理工大学学报(自然科学版)*, 2018, 44(1): 97-103.)

[12] Soon T J. QR Code[J]. *Synthesis Journal*, 2008, 2008: 59-78.

[13] Wicker S B, Bhargava V K. Reed-Solomon Codes and Their Applications[M]. Hoboken: Wiley-IEEE Press [Imprint], 1999.

[14] Macrakis K. Prisoners, Lovers, and Spies: The Story of Invisible Ink from Herodotus to al-Qaeda[M]. Yale University Press, 2014.

[15] Lu T C, Tseng C Y, Wu J H. Dual Imaging-Based Reversible Hiding Technique Using LSB Matching[J]. *Signal Processing*, 2015, 108: 77-89.

[16] Tkachenko I, Puech W, Destruel C, et al. Two-Level QR Code for Private Message Sharing and Document Authentication[J]. *IEEE Transactions on Information Forensics and Security*, 2016, 11(3): 571-583.

[17] Lin P Y, Chen Y H, Lu E J L, et al. Secret Hiding Mechanism Using QR Barcode[C]. *2013 International Conference on SignalImage Technology & Internet-Based Systems*, 2014: 22-25.

[18] Tancik M, Mildenhall B, Ng R. StegaStamp: Invisible Hyperlinks in Physical Photographs[C]. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 2114-2123.

[19] ZXing. https://github.com/zxing/zxing

[20] Standard Test Images. https://www.ece.rice.edu/~wakin/image s/.

陶静 于 1995 年在国防科技大学获得计算机科学与技术硕士学位, 现任国防科技大学计算机学院副研究员, 研究领域为网络空间安全, 研究兴趣包括网络安全、网络应用等。Email: ellen5702@aliyun.com

王宝生 于 2005 年在国防科技大学获得计算机科学与技术博士学位, 现任国防科技大学计算机学院研究员, 研究领域为网络空间安全, 研究兴趣包括计算机网络体系结构、网络安全、漏洞挖掘等。Email: bswang@nudt.edu.cn

罗振豪 于 2019 年在国防科技大学网络空间安全专业获得硕士学位。现在国防科技大学网络空间安全专业攻读博士学位。研究领域为软件安全和系统安全。Email: zh.luo@nudt.edu.cn

邢倩倩 于 2018 年在国防科技大学计算机科学与技术专业获得博士学位。现任国防科技大学网络空间安全系助理研究员。研究领域为网络安全与密码应用。研究兴趣包括新型可信网络、路由安全、公钥密码基础设施安全。Email: xingqianqian12@nudt.edu.cn