面向指纹考勤机的数据取证与分析方法

顾红星[1]，上官梦轩[2]，丁锰[1,3]，何易临[1]，马万里[1]
（1. 中国人民公安大学侦查学院，北京 100038；
2. 温州市公安司法鉴定中心，浙江 温州 325101；
3. 中国人民公安大学公共安全行为科学实验室，北京 100038）

摘　要：指纹考勤机作为目前广泛应用的身份识别设备，其存储的指纹数据和考勤记录在法庭科学中具有重要的取证价值。然而，对其取证和分析面临以下3个问题：常规取证方法受限、指纹模板数据识别困难、缺少有针对性的关联分析方法。针对以上问题，提出了一种面向指纹考勤机的数据取证与分析方法。首先根据指纹考勤机的存储特性，研究直连提取方法，能够直接从存储芯片中读取底层数据；其次，针对考勤机中常见的3类指纹模板，提出对应的识别方法；最后，对于考勤机内部不同类型的数据，提出异构数据的关联分析方法。实验结果表明，该方法能够突破接口限制，可实现复杂条件下的数据提取，对于指纹模板数据可实现跨设备识别，在此基础上通过关联分析能够发现涉案人的行为模式，为法庭科学取证工作提供新的思路和方法。

关键词：数据取证与分析；法庭科学；取证介质；指纹模板；数据关联
中图分类号：TP391；TN06
文献标志码：A
doi: 10.11959/j.issn.1000−0801.2026103

Data forensics and analysis for fingerprint attendance machines

Gu Hongxing[1], Shangguan Mengxuan[2], Ding Meng[1,3], He Yilin[1], Ma Wanli[1]
1. Department of Criminal Investigation, People's Public Security University of China, Beijing 100038, China
2. Wenzhou Public Security Forensic Center, Wenzhou, Zhejiang 325101, China
3. Public Security Behavioral Science Lab, People's Public Security University of China, Beijing 100038, China

Abstract: As widely applied identity recognition devices, fingerprint attendance machines store fingerprint data and attendance records, which possess significant evidentiary value in forensic science. However, the forensic extraction and analysis of this data were challenged by three primary issues: the limitations of conventional extraction methods, the difficulty in identifying fingerprint template data, and the lack of targeted correlation analysis methods. To address the aforementioned issues, a data forensic and analysis method specifically for fingerprint attendance machines was

收稿日期：2014−12−01；修回日期：2015−01−01
通信作者：丁锰，dingmeng@ppsuc.edu.cn
基金项目：广东省证据材料司法鉴定（南天）工程技术研究中心开放课题基金项目（No. ETRC202403）
Foundation Item: The Opening Project of Guangdong Provincial Forensic Science of Evidence Materials (Nantian) Engineering Technology Research Center(No. ETRC202403)

proposed. Firstly, based on the storage characteristics of the attendance machines, a direct-connect extraction method was developed, by which low-level data could be read directly from the storage chip. Secondly, corresponding identification methods were proposed for the three common types of fingerprint templates found in these machines. Finally, for the different types of data within the machines, a correlation analysis method for heterogeneous data was put forward. Experimental results demonstrated that interface limitations could be overcome by this method, and data extraction was achieved under complex conditions. For fingerprint template data, cross-device identification was realized. Building on this, the behavioral patterns of individuals involved could be discovered through correlation analysis. Consequently, new approaches and methodologies were provided for forensic science work.

Key words: data forensics and analysis, forensic science, forensic medium, fingerprint template, data correlation

0 引言

近年来，指纹考勤机作为目前广泛应用的身份识别设备，在法庭科学中具有重要的取证价值。指纹考勤机内置存储芯片，其中存储了指纹模板和考勤记录[1]，因此是一种重要的取证介质。指纹考勤机通过采集和比对个人指纹信息来实现高效的考勤管理[2]，其不仅记录了个人的指纹特征信息，还详细记载了其工作时间、出勤规律以及组织结构等关键数据[3]。通过对考勤机的取证和分析，可以帮助公安机关准确识别犯罪嫌疑人身份，重建诈骗组织的人员架构，分析各个成员在犯罪活动中的具体角色和分工[4]。

目前，指纹考勤机的取证和分析主要存在3方面的困难。

（1）对指纹考勤机进行取证时，常规取证方法受限。文献[5]和文献[6]指出，对于采用球栅阵列[7]（ball grid array，BGA）封装或直接焊接于主板的芯片，常规手段难以无损拆卸。因BGA通常不预留标准接口，数据提取需先使用热风枪、植球工具等专业设备进行高温脱焊。同时，文献[8]及多项研究指出BGA焊点对温度极为敏感，在热循环条件下，焊点的最大应力集中在角焊点处。这就意味着温度控制稍有不当，极有可能导致关键焊点的应力集中和变形，进一步增加了数据取证的技术门槛。

（2）考勤机中的指纹模板数据识别困难。为保护信息安全，考勤机中的指纹数据往往会经过SHA-256[9]等不可逆加密算法处理，只保留数字特征模板而不存储原始图像，从源头防止生物信息泄露[10]。文献[11]指出仅凭存储的模板无法直接逆向还原原始指纹图像。此外，根据美国国家标准与技术研究院（national institute of standards and technology，NIST）的专项评估，厂商普遍采用私有指纹模板格式，这些格式只能被创建它们的软件解析[12]。由于缺乏统一的标准化规范，同一指纹在不同厂商设备中的加密模板差异极大，无法直接跨设备比对。这种不可逆性和设备间的不兼容性进一步增加了指纹模板识别的难度。

（3）考勤机底层数据差异很大，缺少有针对性的关联分析方法。文献[13]和文献[14]指出，目前尚缺少有效的技术方法能将涉案考勤机中来自不同功能模块的异构数据转化为具有关联性的证据信息。文献[15]通过比较NIST SP 800-86和ISO/IEC 27037标准在数字取证证据分析中的框架应用，强调了关联分析在数字取证领域的重要地位。文献[16]通过分析统一数字取证方法，强调对于复杂案情，侦查人员必须找到一种有效方法来将这些异构数据格式进行整合，实现数据去孤岛化，以便进行深层次的关联挖掘。因此，迫切需要研究适用于考勤机异构数据关联化分析方法，以提高法庭科学证据的完整性。

为解决上述问题，本文提出一种面向指纹考

勤机的数据取证与分析方法。本文的主要贡献在于：

（1）研究了一种面向串行外设接口闪存（serial peripheral interface flash，SPI Flash）的直连提取方法，能够直接从存储芯片中读取底层数据，该方法成功突破了接口限制，实现了在复杂条件下的稳定数据提取，拓展了现有取证方法的技术边界。

（2）针对考勤机中常见模板，提出了指纹模板识别方法，能够有效分析指纹信息以及跨设备指纹分析，提升了对生物特征信息的取证分析能力。

（3）对于考勤机内部不同类型的数据，提出异构数据的关联分析方法，通过字符映射和数据解密，建立基础关联关系，进而发现涉案人行为模式，为法庭科学取证工作提供新的思路和方法。

1　相关研究

1.1　指纹考勤机存储特性

在指纹考勤机中，SPI Flash作为关键存储组件，用于保存固件和操作系统[17]。这是一种基于或非型闪存（NOR Flash）架构的嵌入式非易失性存储器[18-20]，支持芯片内执行功能，允许主控芯片直接从Flash中运行固件而无须加载至随机存取存储器（random access memory，RAM）[21]，从而提高了系统的可靠性和启动效率。SPI Flash采用串行外设接口协议进行数据传输，其结构简单且易于集成到嵌入式设备中[22]。

SPI Flash其非易失性确保数据在断电后不会丢失，这对于存储敏感信息如考勤记录至关重要[23]。且SPI Flash具有接口简单、成本低廉、功耗较低等优点[24]，便于在取证过程中高效访问和提取数据。这些特性使SPI Flash在嵌入式系统中广泛用于可靠数据存储[25]，支持长期取证分析。

1.2　指纹考勤机中的指纹模板

目前，指纹考勤机中常见的指纹模板有细节点模板、灰度图像模板、统计特征模板共3类模板。每类模板的特征如下所述：

（1）细节点模板由指纹中的端点和分叉点等关键特征构成[26]，这些特征通过坐标位置、方向角和类型等参数进行描述[27]。

（2）灰度图像模板直接存储指纹的像素级信息，包括原始灰度图像、经过增强处理的图像以及归一化的图像等形式[28]。

（3）统计特征模板通过提取指纹的统计特性来表示身份信息，主要包括Gabor特征、小波特征等[29]。Gabor滤波器是一个用于边缘提取的线性滤波器，能提供良好的方向选择性和尺度选择特性[30]。

1.3　关联分析算法

关联分析算法在指纹考勤机取证中的主要作用是从考勤机产生的大规模异构数据集中挖掘隐藏的变量间关联规则和潜在模式，通过发现考勤数据的内在联系和规律特征，为数字取证决策分析提供支撑[31]。其中，Hash Join算法是一种通过构建简单、高效的哈希表来优化连接性能的关联分析算法。在关联分析中，该方法是发现数据项之间关联关系的基础操作[32]。

聚类算法也可用于关联分析中[33]。其中，K-means聚类算法在关联分析中，通过将具有相似特征的数据点聚集，能够揭示数据中潜在的规律和异常[34]。

2　本文方法

针对法庭科学的实际需求，本文方法分为面向SPI Flash的取证方法和面向考勤机中指纹数据的取证方法2个部分，以此来覆盖取证的全流程。

2.1　面向SPI Flash的取证方法

针对指纹考勤机中SPI Flash存储芯片，本文研究应用SPI接口的直连提取方法（direct con-

nection extraction，DCE）。本文的DCE方法有3项关键技术，DCE关键技术见表1。（1）研究兼容方法，确保DCE与大多数考勤设备兼容，并且能满足从低端到高端考勤机的数据读取需求。（2）研究指令模拟方法，确保DCE能识别和模拟芯片厂商的专有指令集，即使是非标准化的存储芯片也能有效读取。（3）研究坏块处理方法，保证DCE规避坏块，同时将有效数据重组至低磨损的健康块中，以维持稳定的数据存储性能。

表1　DCE关键技术

方法名称	具体内容
兼容方法	支持1.8 V/3.3 V双电压模式，支持范围从1~512 MB及以上的存储芯片
指令模拟	逆向解析厂商私有指令集
坏块处理	结合出厂坏块标记（OOB区域）及动态磨损均衡算法重组有效数据

DCE方法通过一些关键指令实现了底层数据取证，这些指令使DCE能够与各类SPI Flash芯片建立可靠通信，实现数据的精确提取，DCE关键指令见表2。

表2　DCE关键指令

指令代码	指令名称	指令描述
03h	标准读指令	3字节地址周期
0Bh	快速读指令	含8个Dummy周期
20h	扇区擦除指令	4 KB粒度

为确保数据提取过程的准确性和完整性，本文还开发了基于DCE方法的辅助软件。辅助软件可以帮助识别芯片信号，读取存储内容的十六进制数据，并提供多种操作模式，同时支持对提取的数据进行实时预览和格式化显示，在提取完成后还支持数据完整性的校验。

2.2　面向考勤机中指纹数据的取证方法

为解决指纹考勤机只保留数字特征模板而不存储原始指纹图像问题，本文提出指纹模板识别方法。对目前考勤机中常见的3类指纹模板分别进行研究。针对只保存提取出的细节点坐标、方向、类型等抽象数值的情况，提出细节点识别方法；针对直接保存灰度的指纹图像的情况，提出灰度图像识别方法；针对保存图像局部块的统计数的情况，提出统计特征识别方法。

2.2.1　数据结构解析

在进行指纹模板识别前，为精确定位指纹编码存储区域，首先要采用数据结构解析方法。通过对头部信息的精细化解析，能准确识别关键数据元素，如标识符、时间戳、用户ID等。例如，对于格式如[Header]，[UserID]，[Timestamp]，[Fingerprint]，[Check-sum]这类复杂数据结构，能精确定位与提取指纹编码存储区域。该方法可克服传统方法处理非标准数据的局限。

2.2.2　识别方法构建

第一类是细节点识别方法。对于细节点模板数据，通过精确解析坐标、角度及类型信息，来还原指纹关键特征点分布。细节点模板T_{minutiae}计算式如下：

$$T_{\text{minutiae}}=\{M_i|i=1,2,\cdots,n\} \tag{1}$$

其中，$M_i \triangleq (x_i,y_i,\theta_i,t_i)$表示第$i$个细节点，$x_i$表示细节点的横坐标位置，$y_i$表示细节点的纵坐标位置，$\theta_i$表示细节点的方向角度，$t_i$表示细节点类型（分岔点或端点），$n$为细节点总数。

为还原指纹的关键点分布，编码格式表达式为Encode(M_i)=[Header][x_i][y_i][θ_i][t_i]。其中，[]表示数据字段，[Header]表示头部信息，例如，0x02。x_i占2字节，y_i占2字节，θ_i占1字节，t_i占1字节。

第二类是灰度图像识别模板。对于图像模板数据，本文实现小波标量量化（wavelet scalar quantization，WSQ）专业压缩格式的解码，从而根据压缩过的灰度图像进行推断，解码后的灰度数据$I(x,y)$计算式如下：

$$I(x,y)=\text{WSQ}_{\text{decode}}(C_{\text{compressed}}) \tag{2}$$

其中，$C_{\text{compressed}}$表示WSQ压缩数据。

对于其存储的数据，利用狄拉克函数进行还原，重构后的指纹图像$F_{\text{reconstructed}}(x,y)$计算式如下：

$$F_{\text{reconstructed}}(x,y)=\sum_{i=0}^{W-1}\sum_{j=0}^{H-1}I(i,j)\cdot\delta(x-i,y-j) \quad (3)$$

其中，W和H分别为图像的宽度和高度，$\delta(x-i,y-j)$为狄拉克函数。

第三类是统计特征识别模板。使用多个不同参数的Gabor滤波器构建滤波器组，每个滤波器捕获特定方向和频率的纹理信息，通过卷积操作获得图像在各个滤波器下的响应，第i个滤波器$R_i(x,y)$响应计算式如下：

$$R_i(x,y)=I(x,y)\otimes g_i(x,y),i=1,2,\cdots,k \quad (4)$$

其中，$I(x,y)$表示输入图像，$g_i(x,y)$表示第i个Gabor滤波器，$\otimes$表示卷积操作，k表示Gabor滤波器的总数。

通过统计特征实现从二维响应矩阵到低维特征向量的降维，统计特征向量$\boldsymbol{F}_{\text{statistical}}$的计算式如下：

$$\boldsymbol{F}_{\text{statistical}}=[\boldsymbol{R}_1,\boldsymbol{R}_2,\cdots\boldsymbol{R}_i,\cdots,\boldsymbol{R}_k] \quad (5)$$

其中，对于第i个滤波器响应$\boldsymbol{R}_i(x,y)$，其统计特征为$\boldsymbol{R}_i=[\mu_i,\sigma_i^2,E_i]$。均值$\mu_i$反应响应的整体强度水平，方差$\sigma_i^2$衡量响应的变化程度和纹理复杂度，能量$E_i$表示纹理的总体强度和显著性。

利用提取的统计特征计算每个滤波器的标量权重，避免直接使用高维响应数据，重构图像$\boldsymbol{F}_{\text{reconstructed}}(x,y)$计算式如下：

$$\boldsymbol{F}_{\text{reconstructed}}(x,y)=\sum_{i=1}^{k}w_i\cdot g_i(x,y) \quad (6)$$

其中，$w_i=\alpha_1\mu i+\alpha_2\sigma_i^2+\alpha_3E_i$，$w$为第$i$个滤波器的权重，$\alpha_1$为均值项的权重系数，$\alpha_2$为方差项的权重系数，$\alpha_3$为能量项的权重系数。

为找到最佳的权重组合，使重构对象与原始图像的误差最小采用最小二乘准则优化重构质量，最佳权重组合$\{\alpha_1,\alpha_2,\alpha_3\}$计算式如下：

$$\{\alpha_1,\alpha_2,\alpha_3\}=\arg\min\iint|I(x,y)-\sum_{i=1}^{k}w_i\cdot g_i(x,y)|^2dxdy \quad (7)$$

其中，arg min为使目标函数最小的参数值，约束条件为$\sum_{i=1}^{k}w_i=1$和$w_i\geqslant 0$。

2.2.3　最小变化对比分析法

通过识别方法的构建，可以初步明确常见指纹模板的存储格式。由于仅凭存储的指纹信息无法直接逆向还原原始指纹图像，本文在传统差异比较方法的基础上，提出了专门针对指纹模板识别的最小变化对比分析法（minimal change contrastive analysis，MCCA）。针对考勤机存储指纹模板格式高度专有和法庭科学取证的实际情况，此方法作为一种非侵入式的黑盒方法，不试图破解指纹模板的完整结构，而是通过施加受控的输入扰动，来逆向探测和表征考勤机的内容特征提取算法的响应规律。

MCCA是一种基于增量数据比较的逆向工程分析方法。与传统差异分析方法不同，该方法通过控制变量的逐步增加，利用差分模板映射算法来对比分析指纹模板在不同状态下的数据变化模式，从而识别和验证底层数据结构的编码机制和存储规律。MCCA的流程分为3部分，首先进行基线标定，选取基准指纹图像I_{base}，通过目标设备$\mathcal{F}$获取基线模板T_{base}，其关系表示为：

$$T_{\text{base}}=\mathcal{F}(I_{\text{base}}) \quad (8)$$

然后对I_{base}施加一个已知的、最小的图像扰动Δw（例如引入一个端点），生成扰动图像I_{pert}，其关系表示为：

$$I_{\text{pert}}=I_{\text{base}}\oplus\Delta w \quad (9)$$

其中，$\oplus$代表图像合成算子。

最后将I_{pert}输入到同一设备，获取扰动模板T_{pert}如下：

$$T_{\text{pert}}=\mathcal{F}(I_{\text{pert}}) \quad (10)$$

该方法的核心在于量化扰动和解释扰动Δw

与输出模板 $\Delta T = T_{\text{pert}} \ominus T_{\text{base}}$ 之间的因果关系。

为实现上述差分分析，本文设计了差分模板映射算法。该算法通过逐字节的比较，能自动识别有意义的数据变化边界，避免了固定块大小分析的局限性。当检测到存储数据发生变化时，算法会记录变化区域的起始位置、长度以及具体的数据内容，并通过特征匹配函数将这些变化与新录入的变量进行关联，从而建立完整的特征-存储映射表。

差分模板映射算法的伪代码见算法1。其中，data_i 表示第 i 次存储数据，data_i+1 表示第 $i+1$ 次存储数据，$\text{fingerprint}_{\text{features}}$ 表示第 i 次录入指纹特征信息，$\text{feature}_{\text{mapping}}$ 表示特征-存储映射关系数组，pos 表示当前数据比较的指针，start 表示数据变化区域的起始位置偏移量，mapping 表示单个映射关系记录，$\text{storage}_{\text{offset}}$ 表示存储变化的起始偏移地址，$\text{storage}_{\text{length}}$ 表示存储变化区域的字节长度，$\text{changed}_{\text{data}}$ 表示发生变化的具体数据内容，$\text{matched}_{\text{feature}}$ 表示与存储变化对应的指纹特征，MIN 函数为取最小值函数，MATCH_TO_FEATURE 函数为特征匹配函数，将变化的存储数据与输入的指纹特征进行关联分析，返回最匹配的特征类型，$\text{data}_i[\text{start:pos}]$ 表示数组切片操作，$\text{feature}_{\text{mapping}}.\text{append}(\text{mapping})$ 表示添加新的映射关系记录。

算法1：差分模板映射

输入：data_i, data_i+1, $\text{fingerprint}_{\text{features}}$

输出：$\text{feature}_{\text{mapping}}$

$\text{feature}_{\text{mapping}} \leftarrow []$

$\text{pos} \leftarrow 0$

while $\text{pos} < \text{MIN}(\text{length}(\text{data}_i), \text{length}(\text{data}_i+1))$ **do**

  if $\text{data}_i[\text{pos}] \neq \text{data}_i+1[\text{pos}]$ **then**

    $\text{start} \leftarrow \text{pos}$

    while $\text{pos} < \text{MIN}(\text{length}(\text{data}_i), \text{length}(\text{data}_i+1))$ AND $\text{data}_i[\text{pos}] \neq \text{data}_i+1[\text{pos}]$ **do**

      $\text{pos} \leftarrow \text{pos} + 1$

    end **while**

    mapping←{$\text{storage}_{\text{offset}}$:start, $\text{storage}_{\text{length}}$: $\text{pos} - \text{start}$, $\text{changed}_{\text{data}}$: $\text{data}_i+1[\text{start:pos}]$, $\text{matched}_{\text{feature}}$: MATCH_TO_FEATURE($\text{data}_i+1[\text{start:pos}]$, $\text{fingerprint}_{\text{features}}$)}

    $\text{feature}_{\text{mapping}}.\text{append}(\text{mapping})$

  else

    $\text{pos} \leftarrow \text{pos} + 1$

  end if

end while

return $\text{feature}_{\text{mapping}}$

2.3 异构数据的关联分析方法

为解决考勤机数据缺少有针对性的关联分析方法问题，本文提出异构数据的关联分析方法。首先对底层数据进行预处理。针对考勤机中底层数据固定长度分段存储的特点，设计字符映射机制。该机制首先通过分析数据的固定长度模式，自动识别数据的分段边界。假设原始数据流为 $D_{\text{raw}} = c_1 c_2 \cdots c_L$，其固定长度模式可表示为 M 个数据字段 F_i 的序列：

$$D_{\text{raw}} = F_1 \oplus F_2 \oplus \cdots \oplus F_M = \bigoplus_{i=1}^{M} F_i \tag{11}$$

其中，$\oplus$ 为字符串拼接，每个字段 F_i 具有固定长度 $l_i = \text{length}(F_i)$。

然后根据数据特征动态尝试不同的字符编码映射长度，通过建立长度与内容对应关系表，实现对不同字符类型（如用户ID、时间戳、权限标识等）的字符映射。将此对应关系表定义为一个类型映射函数 Φ，它将字段长度 l_i 映射到其对应的数据类型 τ_i：

$$\tau_i = \Phi(l_i) \tag{12}$$

每种数据类型 τ 均存在一个特定的解码函数 D_τ，该方法避免了传统的固定长度映射可能导致的数据错位问题。对于其中存在加密的密码信息，

采用“黑盒分析+白盒验证”的解码方法，该方法引入了编码一致性检验，通过多组已知样本的交叉验证，确保推导出的解码算法具有普适性。

设P为原始密码明文，E为其在考勤机中对应的存储数据（密文）。首先，通过创建已知密码样本构建验证集S_{known}：

$$S_{\text{known}}=\{(P_1,E_1),\cdots,(P_N,E_N)\} \tag{13}$$

其中，N为样本数量。解码目标是寻找映射函数$f:E\to P$，使其满足编码一致性检验。

结合密码长度限制（$P\in[0,10^6-1]$）与底层数据变化模式，定义原子操作集$\mathcal{G}$（包含数值变换、位操作、偏移）。设$\mathcal{F}$为由$\mathcal{G}$中元素通过有限次复合运算生成的函数全集。构建候选解码算法的搜索空间H。该空间H定义如下：

$$H=\{f\in\mathcal{F}\,|\,\forall(P_i,E_i)\in S_{\text{known}};f(E_i)=P_i\} \tag{14}$$

其中，E_i代表第i个已知的存储数据，P_i代表第i个密码明文。

通过测试H中不同的数值变换规则，并利用已知密码和对应的存储数据进行验证，逐步收敛至正确的解码算法f_{correct}。该收敛过程可表示为：

$$f_{\text{correct}}=\arg\min_{f\in H}\sum_{j=1}^{N}\delta(f(E_j),P_j) \tag{15}$$

其中，$\arg\min$为参数最小化操作，P_j代表第j个已知的密码，$\delta(\cdot,\cdot)$为损失函数用于量化不匹配程度。

在对考勤机底层数据进行预处理后，首先需要关联多维度的考勤特征（如个人信息、考勤记录等）。传统的哈希连接（如$h(k)=k\bmod q$）虽然实现简单，但存在易碰撞的问题。当数据分布不均时，例如，考勤数据在特定时间中高度集中或表长选择不合适时，容易发生哈希碰撞，这会导致连接性能从$O(N)$退化到接近$O(N^2)$，降低关联的速度。为解决此问题，使得哈希值能均匀分布，在期望上保证较低的碰撞率，即使面对严重的数据倾斜也能保持稳定，采用的关联策略$h_{a,b}(k)$如下：

$$h_{a,b}(k)=((ak+b)\bmod p)\bmod q \tag{16}$$

其中，k为输入的键值（如个人ID），p是一个大于所有k的大素数，a和b是在$[0,p-1]$区间内随机选择的整数（$a\neq0$），q为哈希表的大小，mod表示模运算。

建立基础关联关系后，为发现其中的考勤模式，以及为异常检测和行为分析提供数据驱动的依据，本文采用聚类分析。传统的K-means聚类算法在处理考勤数据时存在对异常值问题，K-means的目标函数J使用了平方欧式距离，但考勤数据中不可避免地出现偶然的极端打卡记录（如系统错误或忘记打卡导致的0点或23：59），这会导致这些异常值对均值μ_i产生过度的拉扯，使聚类中心严重偏离正常模式。针对上述问题，本文的聚类分析不再使用均值μ_i作为中心，而是使用中心点即一次真实的考勤记录，将目标函数J修改为基于对异常值更鲁棒的L_1范数（曼哈顿距离）：

$$J_{\text{robust}}=\sum_{i=1}^{k}\sum_{x\in C_i}\|\boldsymbol{x}-\boldsymbol{m}_i\|1 \tag{17}$$

其中，k为聚类数量，C_i为第i个聚类，$\boldsymbol{m}_i$为第i个中心点，$\boldsymbol{x}$为数据点，$\|\cdots\|1$为L_1范数即曼哈顿距离。

由于中心点必须是真实数据点，更新规则也不再是求平均。而是在C_i内部进行搜索，找到能使类内L_1距离总和最小的真实数据点$\boldsymbol{x}^*$，并将其设为新的中心点$\boldsymbol{m}_i$：

$$\boldsymbol{m}_i^{(t+1)}=\arg\min_{x\in C_i^{(t)}}\sum_{y\in C_i^{(t)}}\|\boldsymbol{y}-\boldsymbol{x}\|_1 \tag{18}$$

其中，$\arg\min$为参数最小化操作，$C_i^{(t)}$为第t次迭代时第i个类所包含的所有数据点的集合，$\boldsymbol{x}$和$\boldsymbol{y}$为$C_i^{(t)}$中的2个数据点，$\boldsymbol{y}-\boldsymbol{x}$为2个数据点之间的差向量。

3 实验与结果分析

本文通过实验展示所提方法的具体应用流程。实验过程包括数据提取和数据分析2个部分。提取过程中所用硬件包括取证介质（指纹考勤机）、外壳拆卸工具（螺丝刀、撬片等）、水平仪，软件为数据分析辅助软件X-Ways Forensics。

实验流程图如图1所示。取证介质为指纹考勤机，其品牌为得力（deli），型号为3960S，序列号为64 439 803 693。取证数据包括指纹信息、个人信息、出勤规律和组织结构等，数据大小为8 192 KB。

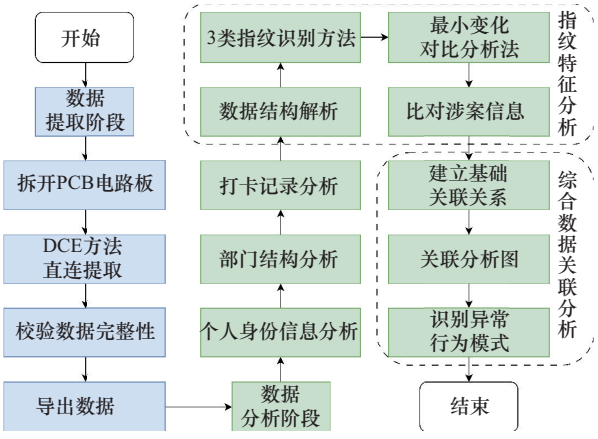


图1 实验流程图

3.1 数据提取

依据法庭科学无损取证原则，为防止通电可能引发的数据复写或状态变更，保证取证介质的状态不被改变，在数据提取过程中未对指纹考勤机进行任何通电操作。

取出存储芯片，存储芯片照片如图2所示。为确定存储芯片的完整技术参数，需综合考虑芯片型号、芯片外围电路功能。最终确定存储芯片的类型为SPI FLASH，品牌为winbond，型号为W25Q64JV，容量为8 MB（8 192 KB），封装及规格为SOP8-208mil，可以适用上述DCE方法。

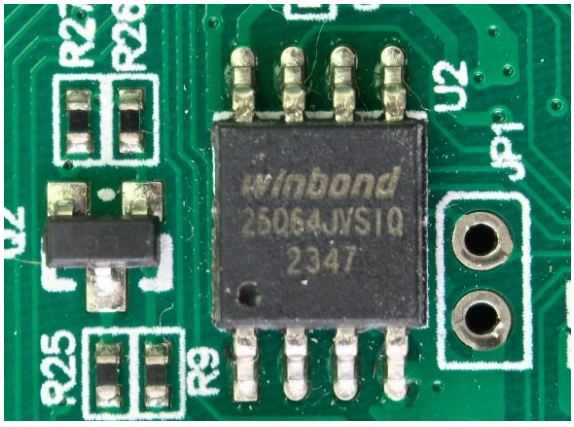


图2 存储芯片照片

确认存储芯片的参数后，使用DCE方法和DCE辅助软件进行数据提取。

下面是本文DCE方法的具体步骤：

（1）获取芯片技术参数。下载W25Q64JV芯片的DataSheet（技术规格说明书）文档（https://www.alldatasheet.com/），读取其中的芯片参数。工作电压为2.7 V至3.6 V，引脚间距为0.65 mm，最大工作电流为25 mA。

（2）选择合适的连接设备。根据芯片的规格参数，选择并更换匹配的探针或芯片适配器，其中，探针是用于与芯片引脚建立电气连接的器件，适配器是将芯片固定并连接到读取设备的转接装置。

（3）芯片定位和电路板校准。确定芯片的1号引脚（芯片编号的参考基准）位置，使用水平仪确保PCB（Printed Circuit Board）电路板水平放置以避免连接不良。

（4）精确对位和连接。调整PCB电路板的位置，使探针的1号脚对准芯片的1号引脚，然后缓慢降低探针，直到探针与芯片引脚完全接触。

（5）验证连接状态。如连接失败，要重新调整探针的压力和角度，重复操作直至成功连接芯片。

（6）数据读取和检验。执行芯片数据读取操作。读取完毕后，进行数据完整性校验，通过双重验证机制，CRC3校验和BCH ECC纠错来保证提取的数据与芯片中数据完全一致，数据完整性

校验机制见表3。

（7）数据保存和归档。将读取的数据导出并保存为文件，文件命名与对应的案件名称保持一致，以便后续取证工作的追溯和管理。

表3　数据完整性校验机制

验证类型	校验机制	功能
CRC3校验	页级校验，采用多项式0x04C11DB7	检测传输错误
BCH ECC纠错	块级纠错，可修复8 bit/512 B错误	确保数据完整性

DCE辅助软件界面如图3所示，软件中央区域实时显示正在读取的原始数据，并以十六进制格式呈现，数据区左侧ADDR（Address）列显示当前读取的内存地址位置，帮助定位数据在芯片中的具体位置。左上角区域提供案件信息，包括检材编号、送检单位等。左下角的芯片连接示意图显示当前连接状态，包括引脚对应关系。右上角区域显示取证介质的器件信息。右侧操作栏区域提供了多项关键操作功能，为芯片数据提取提供全面的操作支持。其中，自动烧录（F9）功能能够自动识别芯片类型并执行标准化的数据写入流程；擦除芯片（F4）功能可安全擦除芯片内的所有数据；手动烧录（F6）允许操作人员手动指定芯片型号和烧录参数；数据校验（F8）确保数据的完整性和准确性；读出芯片（F7）是核心功能，执行完整的数据提取操作，软件会按照预设的读取策略逐扇区读取芯片内容；芯片查空（F3）功能用于检测芯片的空白区域；芯片检测（F5）提供引脚导通性测试和电压检测，确保硬件连接的稳定性；批量模式（F2）支持多个芯片的连续处理。

3.2　数据分析

通过本文的DCE方法和辅助软件成功提取存储芯片数据之后，对十六进制数据进行字符映射来得到原始信息，针对其中加密部分进行数据解密，最后对异构数据进行综合分析。提取的数据包括个人信息、考勤信息和组织结构等，数据大小为8 192 KB。

3.2.1　个人身份信息分析

个人身份信息是考勤系统中的核心数据，为发现个人信息在存储芯片中的存储规律和数据结构特征，这里先以文件第3 984扇区为例，个人信息如图4所示。通过对数据内容进行标注，可

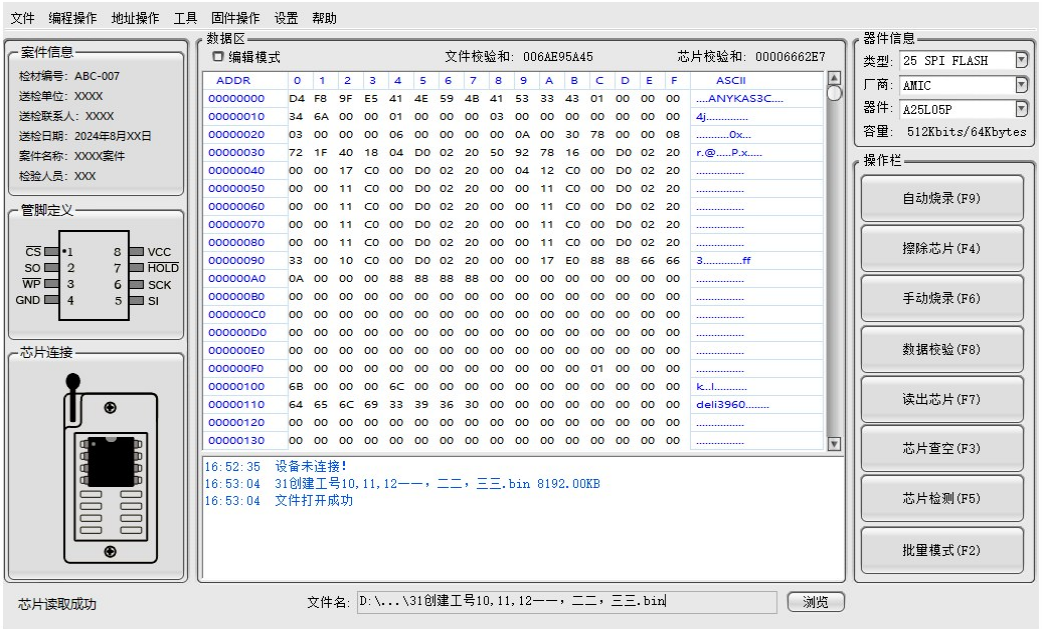


图3　DCE辅助软件界面

图4　个人信息

以识别出个人信息记录的完整结构，每条记录包括工号、指纹组数量、密码状态、指纹状态、密码、最后设置时间。

进一步分析数据的存储格式，发现个人身份信息采用固定长度的分段存储方式。每20 bytes构成一组完整的用户记录，4 bytes为一段分别对应不同的信息字段，十六进制的个人身份信息见表4。

表4　十六进制的个人身份信息

第一段	第二段	第三段	第四段	第五段
01000000	01001011	00000000	9A785600	D189502E
02000000	01001011	00000000	03000000	DE89502E
03000000	15001011	00000000	04000000	0A8A502E
04000000	15001011	00000000	32040000	CF9F4F2E
05000000	05000011	00000000	00000000	2881502E
06000000	05001011	00000000	32040000	6C81502E
07000000	00000001	00000000	00000000	00000000

可根据表4中的数据得出以下信息：

（1）工号：存储位置为0x00-0x03，如01000000对应工号1。

（2）指纹数量：存储位置为0x04，通过特定的位编码方式表示设置的指纹组数。其中01代表一组，05（二进制0101）代表2组，15（二进制010101）代表3组。

（3）密码状态：存储位置为0x06，用于标识是否启用密码验证功能，10代表设置密码打卡，00未设置密码打卡。

（4）指纹状态：存储位置为0x07，用于标识指纹注册状态，11代表设置指纹、01代表未设置。

（5）密码：存储位置为0x0C~0x0E，密码长度限制为十进制的六位，采用特殊的编码算法进行存储。

（6）时间戳信息：存储位置为0x10-0x13，记录最后一次设置指纹或密码的操作时间，采用标准Unix时间戳，表示自1970年1月1日00:00:00 UTC以来经过的秒数。

针对加密信息的解析问题，在黑盒分析阶段，首先建立包含不同位数和数值特征的密码

样本集，包括3位数密码（如123）和6位数密码（如987 654），通过向考勤机输入这些已知密码，记录并分析其数据变化。观察结果显示，存储数据与原始密码之间存在明显的数学变换关系。

基于黑盒分析的观察结果，构建候选解码算法的搜索空间。该搜索空间涵盖了常见的数据变换方式，包括进制转换、位操作、数值偏移以及序列重排等可能的编码策略。通过算法测试和样本匹配，逐步缩小候选算法范围。

经过反复的算法验证和优化，最终确定了密码的解码算法：将存储的十六进制数据转换为二进制表示，然后按每4位进行分组处理，对每组的数据执行减1操作，最后按照从右向左的顺序重新排列各组，即可还原出原始密码，该算法的数学表达可描述为：对于二进制表示$B=b_1b_2b_3\cdots b_n$，其中b表示二进制位，n表示二进制位的总数，按4位分组得到$[G_1][G_2]\cdots[G_m]$，则原始密码$P=(G_m-1)(G_m-1-1)\cdots(G_1-1)$，其中$G$表示分组，$m$表示分组的组数。

为验证解码算法，在白盒验证阶段采用了编码一致性检验方法。通过多组独立的已知样本进行交叉验证，结果表明该解码算法对所有测试样本均能实现100%的准确解码。典型样本的解码过程见表5。

表5　典型样本的解码过程

十六进制	二进制	对应密码
0X432000	0100 0011 0010	123
0X56789A	0101 0110 0111 1000 1001 1010	987654

3.2.2　部门结构分析

为全面还原考勤系统中的组织架构信息，需要对部门名称和人员归属关系进行分析。首先识别出各个部门名称和“部门”标识字符在存储芯片中的编码格式和存储位置。部门名如图5所示，通过定位“部门”二字对应的十六进制数值，该数值可以作为识别部门信息块的关键标识符。然后根据定位的位置还原出各部门的中文名称。

进一步分析个人信息和部门信息之间的关联信息，个人–部门关联信息如图6所示，通过数据结构分析可以识别人名与对应部门名之间的映射关系以及数据块之间的关联关系。

根据关联信息，可确定存储规律，人名–部

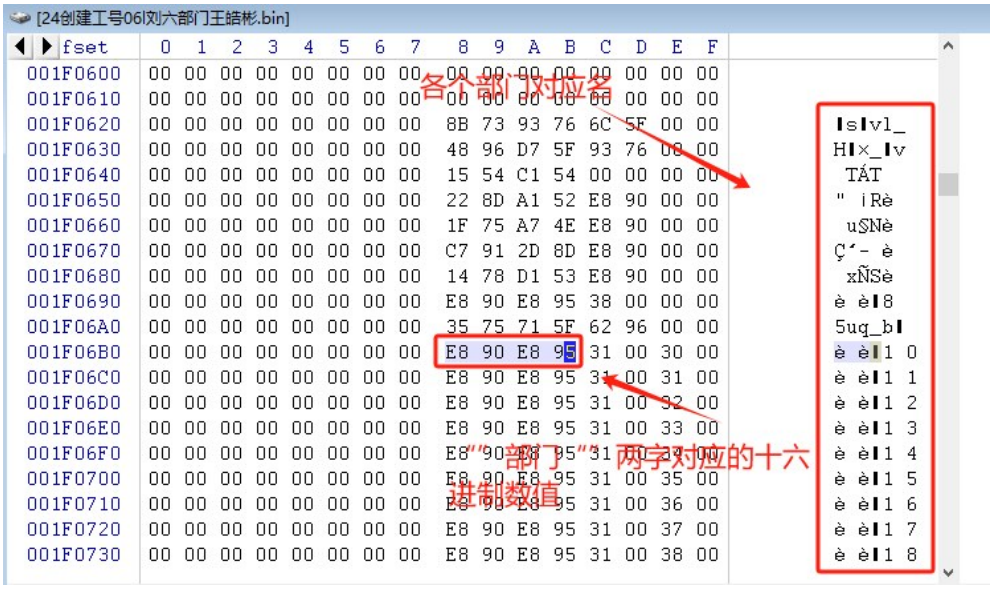


图5　部门名

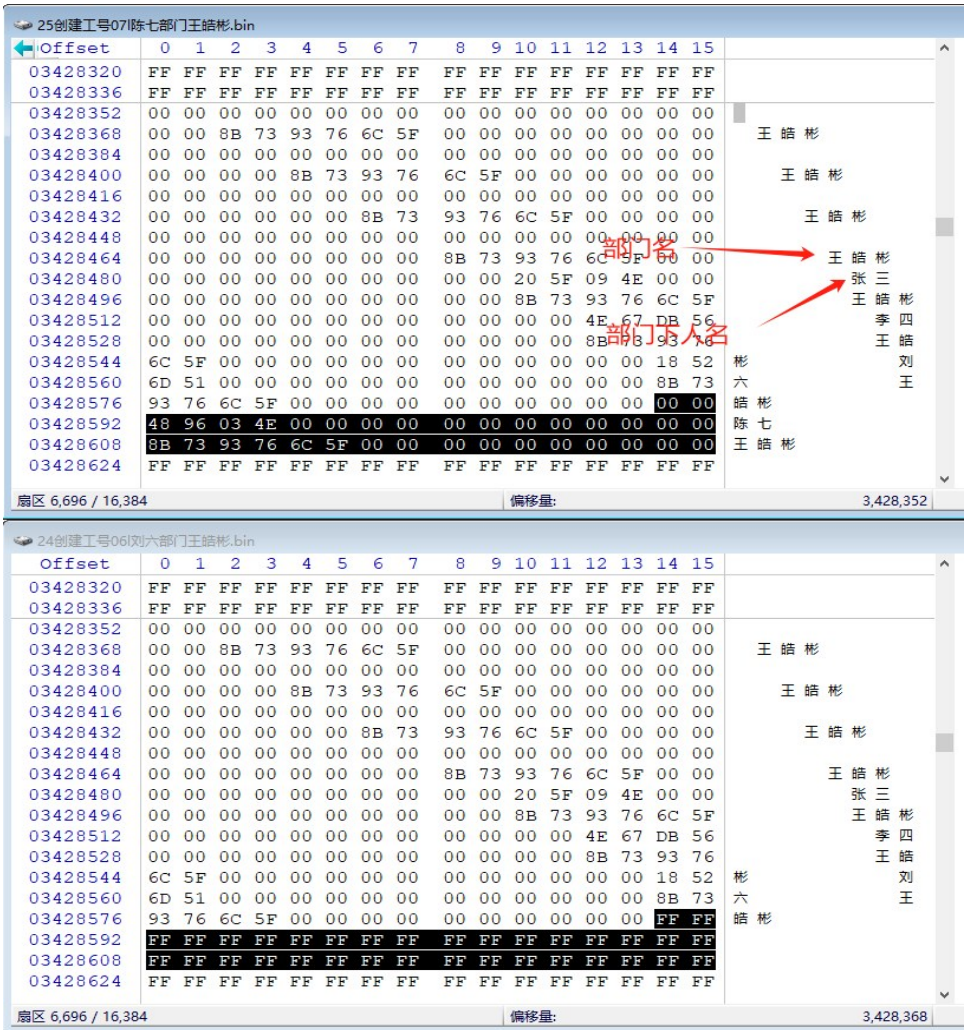


图6　个人–部门关联信息

门名存储规律见表6。部门名和人名按照0x22的固定间隔排列，但数据长度存在差异，这表明不同长度名字的存储需求。

表6　人名–部门名存储规律

信息类型	记录地址	偏移	数据长度/字节	间隔距离
部门名	3428368	0x12-0x17	6	0x22
	3428400	0x34-0x39	6	
	3428432	0x56-0x5B	6	
	3428464	0x78-0x7D	6	
人名	3428368	0x8A-0x8D	4	0x22
		0xAC-0xAF	4	
		0xCE-0xD3	6	
		0xF0-0xF3	4	

3.2.3　个人身份信息分析

打卡记录是反映组织人员活动规律的重要数据源。打卡记录如图7所示，打卡记录按照时间顺序进行存储，呈现出规律性的数据结构特征。

以图7中文件第7 768扇区为例进行分析，每打卡一次，增加一条记录，记录长度为8 bytes，8个bytes为一组，前4个bytes为时间，第5个bytes是工号。从3 977 216开始，时间偏移位置为0x00-0x03, 0x08-0x0B，工号偏移位置为0x04。

3.2.4　指纹特征分析

指纹信息是考勤系统中重要的生物特征数

图7　打卡记录

据。通过对头部信息的解析，来识别出指纹数据，从如图8所示的数据分布可以看出，指纹数据呈现出高度规律化的存储模式，主要表现为密集的十六进制字节序列。

图8　指纹信息

对目前考勤机中常见的3类指纹模板分别进行研究，针对不同模板类型选择相应的解析方法。需要明确的是，由于缺乏原始指纹图像，仅依靠指纹模板数据无法完全还原原始生物特征，但可以通过解析模板数据获取关键特征信息用于比对分析。具体而言：对细节点模板，分析坐标、角度和类型信息的编码规律；对图像模板，进行WSQ压缩形式的解码；对统计特征模板，解析基于Gabor滤波器的特征向量编码方式。

然而，由于不同厂商普遍采用私有指纹模板格式，通用解析算法往往无法直接适用。因此，需要在基础解析算法基础上，结合MCCA调整并优化针对各厂商私有格式的解码参数。具体流程如图9所示。

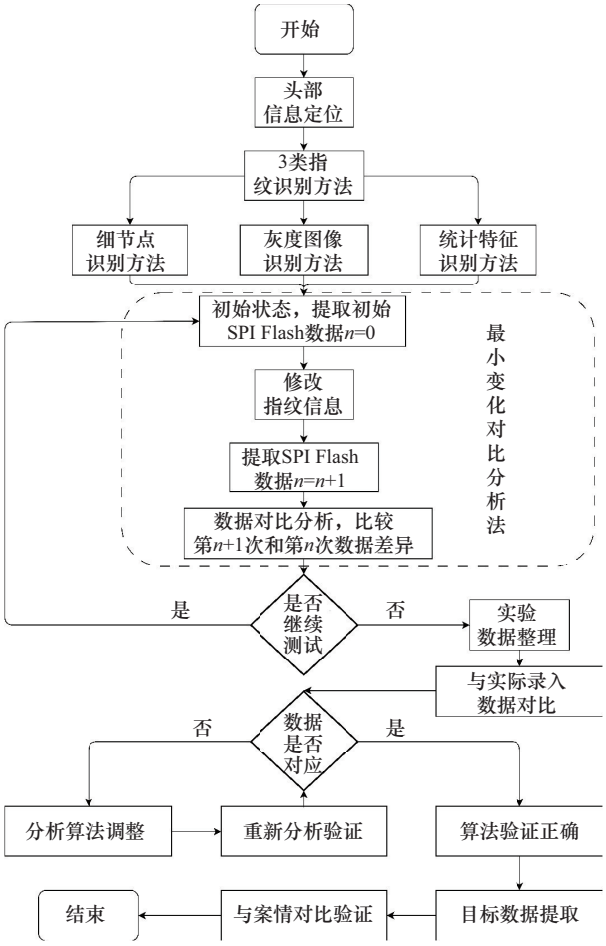


图9　跨设备指纹模板分析流程

该方法的具体实验步骤为：在同一考勤机上依次录入第1条、第2条直至第n条指纹样本，每次录入后立即使用DCE方法提取SPI Flash数据快照。通过差分模板映射算法精确定位第$i+1$和第i次数据的变化区域，结合已知的指纹特征信息，建立“输入特征–存储变化”的映射关系。为提高准确率，采用多样本交叉验证策略，通过不同的指纹样本重复实验，确保识别出的编码规律具有普适性。

基于MCCA调整后的解析算法，首先在此指纹考勤机上进行验证测试，确认算法能够解析该设备的指纹模板数据，验证成功后，将解析得到的指纹特征数据与实际案情中的指纹证据进行比对验证，通过比对结果的一致性来进一步确认解析算法的有效性和准确性。

为满足涉案取证中可能遇到的跨设备分析需求，建立基于设备识别的解析策略：对于已知型号设备，直接应用相应的经过案情验证的解析算法；对于未知设备，首先尝试厂商已验证的算法进行参数调整，解析成功后同样需要进行案情比对验证，若仍无法解析，则重新执行MCCA。通过逐步积累不同厂商设备的解析经验和案情验证结果，形成经过实战检验的涉案指纹设备解析能力库，最终实现对主流考勤设备指纹模板的可靠解析，为跨设备的指纹证据关联分析提供支持。

3.2.5　异构数据关联分析

单独的数据类型只能提供片面信息，而通过系统性的关联分析，能够构建出完整的涉案人员活动轨迹和行为模式。

将提取的异构数据进行关联分析，关联数据分析如图10所示。关联分析界面分为五个主要区域：部门设置区域用于管理组织架构信息，个人信息区域展示员工基础信息，相应部门区域建立人员归属关系，指纹特征信息区域显示解析得到的生物特征数据，打卡信息区域显示考勤行为数据。这种分区设计便于数据的分类管理和快速检索。

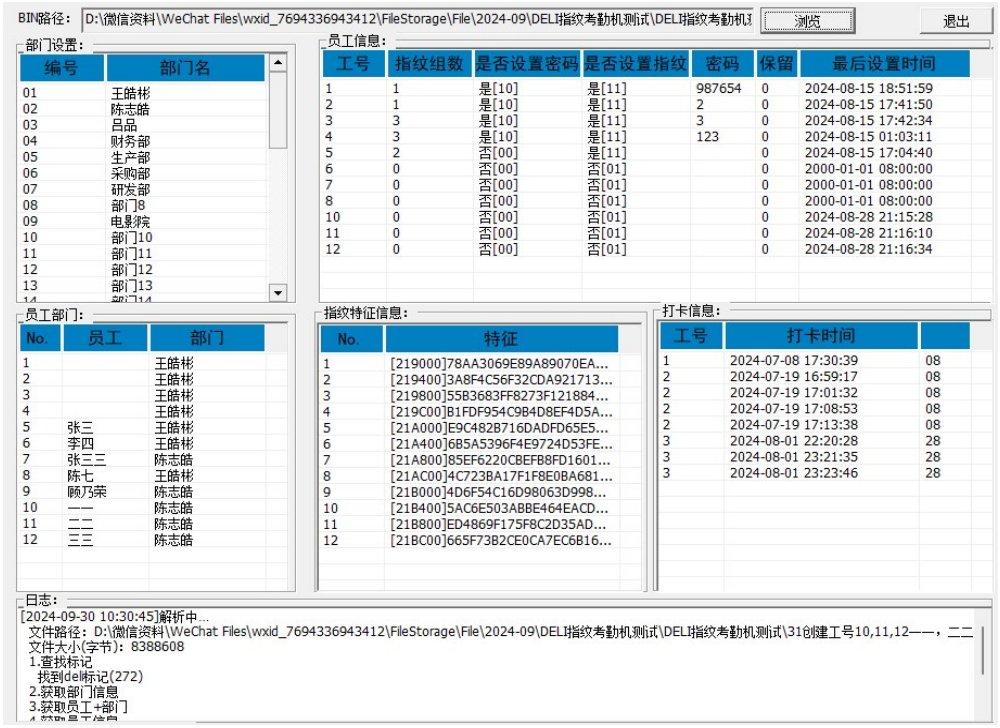


图10　关联数据分析

为实现上述多维数据的有效整合，采用Hash Join算法建立高效的数据关联关系。首先构建哈希索引表，然后建立3层递进的关联关系：第1层为部门–员工关联，通过部门编号快速定位每个部门下的具体人员信息，实现组织架构的清晰映射；第2层为员工–打卡记录关联，以工号作为连接键，识别出各员工的打卡记录，建立人员和行为数据的关联；第3层为员工–指纹特征关联，通过个人ID将员工身份与解析得到的生物特征关联，实现身份验证的数据支撑。通过这种层次化的关联关系，原本分散存储的异构数据得以有机整合，形成完整的考勤关联信息框架。

在数据关联的基础上，进一步采用K-means聚类算法对整合后的员工考勤行为进行模式识别。通过分析关联数据中的时间特征，能够有效识别出3类不同的行为模式：正常考勤模式表现为大部分员工的打卡时间集中在正常工作时段（如17:00~18:00），将这类行为归为正常考勤聚类；加班工作模式体现为部分员工存在22:20~23:23的晚间打卡记录，反映延时工作行为特征；异常行为模式则通过识别2000-01-01等异常的时间戳来发现，这类数据可能反映系统故障、数据篡改或其他异常情况，为案件调查提供重要的异常行为线索。

基于上述关联分析和模式识别结果，采用时空序列分析方法进行证据链构建和行为轨迹重构。以整合后的考勤打卡记录为时间基准线，结合指纹识别记录中的设备位置信息和人员身份信息，从而能够重构特定人员在特定时间段的活动轨迹。通过分析打卡时间序列的连续性、考勤状态的变化规律以及异常行为的分布特征，还能够识别出潜在的异常行为节点，为法庭科学取证工作提供数字化证据支撑。

4　结束语

为了解决常规取证方法受限、指纹模板数据识别困难、缺少有针对性的关联分析方法3个问题，提出了一种面向指纹考勤机的数据取证与分析方法。实验结果表明，该方法能够突破接口限

制，实现了复杂条件下的数据提取，更实现了指纹模板数据的跨设备识别，在此基础上通过关联分析能够有效发现涉案人的行为模式，为法庭科学取证工作提供新的思路和方法。

然而，本研究也存在一些局限性，所提出的方法虽然在一定程度上拓展了法庭科学取证方法，但取证介质类别较为单一，尚未充分覆盖物联网生态中多样化设备。未来的工作将集中于扩充取证介质库，覆盖智能家居、可穿戴设备以及车联网等多场景下的异构数据取证与融合分析，以应对更广泛的实战需求。

参考文献：

[1] Rahkoyo E, Yangdol S, Kaur B, et al. IoT-based fingerprint attendance system: enhancing efficiency and security in educational and organizational settings[C]//Proceedings of the 2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE). Piscataway: IEEE Press, 2024: 1-7.

[2] Kanagamalliga S, Rajalingam S, Karthikeyan M, et al. Arduino-powered fingerprint authentication for door access control[C]//Proceedings of the 2024 5th International Conference on Electronics and Sustainable Communication Systems (ICESC). Piscataway: IEEE Press, 2024: 131-135.

[3] Malabi Y, Hani'ah M, Noprianto, et al. Efficient employee attendance system integrating RFID and Android-based face recognition with liveness detection[C]//Proceedings of the 2024 International Conference on Electrical and Information Technology (IEIT). Piscataway: IEEE Press, 2024: 163-168.

[4] 梁广俊, 辛建芳, 王群, 等. 物联网取证综述[J]. 计算机工程与应用, 2022, 58(8): 12-32.
Liang G J, Xin J F, Wang Q, et al. Survey of IoT forensics[J]. Computer Engineering and Applications, 2022, 58(8): 12-32.

[5] Sang T, Gharaibeh M A, Wentlent L, et al. Drop and impact reliability investigation of BGA and LGA interconnects[J]. Soldering & Surface Mount Technology, 2023, 35(4): 244-254.

[6] Song K X, Gao J C, Flowers G T, et al. Modeling and analysis of signal integrity of ball grid array packages with failed ground solder balls[J]. IEEE Transactions on Components, Packaging and Manufacturing Technology, 2022, 12(2): 306-315.

[7] Hung H H, Cheng Y C, Hwang S J, et al. Analysis of flip-chip ball grid array underfill flow process[J]. The International Journal of Advanced Manufacturing Technology, 2024, 134(9): 4851-4870.

[8] Yang H, Xu Z Y. Comparison of nano-silver solder joints and SAC305 based on ANSYS simulation and life prediction[J]. AIP Advances, 2023, 13(6): 065025.

[9] Singh D, Kaur H, Verma C, et al. A novel 3-D image encryption algorithm based on SHA-256 and chaos theory[J]. Alexandria Engineering Journal, 2025, 122: 564-577.

[10] Abdullahi S M, Sun S F, Wang B, et al. Biometric template attacks and recent protection mechanisms: a survey[J]. Information Fusion, 2024, 103: 102144.

[11] Kaur P, Kumar N, Singh M. Biometric cryptosystems: a comprehensive survey[J]. Multimedia Tools and Applications, 2023, 82(11): 16635-16690.

[12] National Institute of Standards and Technology. Proprietary fingerprint template (PFT) III evaluation: NIST-ITL-IAD-2022 [R]. Gaithersburg: NIST, 2022.

[13] Fakhouri H N, AlSharaiah M A, Alhwaitat A K, et al. Overview of challenges faced by digital forensic[C]//Proceedings of the 2024 2nd International Conference on Cyber Resilience (ICCR). Piscataway: IEEE Press, 2024: 1-8.

[14] Klasén L, Fock N, Forchheimer R. The invisible evidence: Digital forensics as key to solving crimes in the digital age[J]. Forensic Science International, 2024, 362: 112133.

[15] Faizal A, Luthfi A. Comparison study of NIST SP 800-86 and ISO/IEC 27037 standards as a framework for digital forensic evidence analysis[J]. Journal of Information Systems and Informatics, 2024, 6(2): 701-718.

[16] Alshumrani A, Clarke N, Ghita B. A unified forensics analysis approach to digital investigation[C]//Proceedings of the 18th International Conference on Cyber Warfare and Security (ICCWS 2023). Reading: Academic Conferences International Press, 2023: 466-475.

[17] Önel İ G, Arbac S, Korçak Ö. Flash memory based performance analysis on PROFINET devices[C]//Proceedings of the 2025 24th International Symposium INFOTEH-JAHORINA (INFOTEH). Piscataway: IEEE Press, 2025: 1-7.

[18] Tong G G, Huang H S, Li X. Flash controller verification for dual-core IoT chip[C]//Proceedings of the 2025 5th International Conference on Artificial Intelligence and Industrial Technology Applications (AIITA). Piscataway: IEEE Press, 2025: 1509-1512.

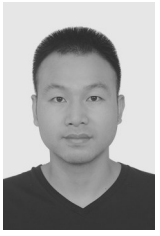
[19] Li Y H, Nie Y J, Huang Z H, et al. Research on firmware upgrade method of PCIE board based on FPGA[C]//Proceedings of the 2024 IEEE 4th International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA). Piscataway: IEEE Press,2024: 1340-1344.

[20] Huang H, Pan Y Q, Xia W, et al. Simplifying and accelerating NOR flash I/O stack for RAM-restricted microcontrollers[C]//Proceedings of the 30th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2. New York: ACM Press, 2025: 1076-1090.

[21] 孙鉴, 武晓晓, 谢开斌, 等. 面向闪存的树型索引综述[J]. 计算机工程与应用, 2022, 58(22): 30-40.
Sun J, Wu X X, Xie K B, et al. Review of tree index based on flash memory[J]. Computer Engineering and Applications, 2022, 58(22): 30-40.
[22] Ben Moussa N, Chetoui M, Amairi M. Optimization-based robust fractional-order controllers design for multiple-input single-output systems with time delays[J]. Measurement and Control, 2025, 58(10): 1383-1398.
[23] Suresh A, Pruthviraj N N, Srivastava S, et al. Reliable and expandable SPI interface for multichannel ADC[C]//Proceedings of the SoutheastCon 2024. Piscataway: IEEE Press, 2024: 989-994.
[24] Shaila C K, Manoj G, Divya P S, et al. Functional verification of SPI protocol using UVM based on AMBA architecture for flash memory applications[C]//Proceedings of the 2023 4th International Conference on Signal Processing and Communication (ICSPC). Piscataway: IEEE Press, 2023: 311-315.
[25] Gao J C, Zhang G H, Wu Y F, et al. Research on the reliable loading technology of OS-based space-borne embedded systems[C]//Proceedings of the 2024 IEEE 6th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC). Piscataway: IEEE Press, 2024: 1365-1370.
[26] Li S M, Xu H, Wang J K, et al. Hierarchical perceptual noise injection for social media fingerprint privacy protection[J]. IEEE Transactions on Image Processing, 2024, 33: 2714-2729.
[27] Bouhamed S A. Possibilistic modeling for fingerprint image quality assessment in Automatic Identification Systems[J]. Applied Soft Computing, 2025, 171: 112790.
[28] Jia Z X, Huang C W, Wang Z, et al. Finger recovery transformer: toward better incomplete fingerprint identification[J]. IEEE Transactions on Information Forensics and Security, 2024, 19: 8860-8874.
[29] Yang W C, Wang S, Hu J K, et al. Feature extraction and learning approaches for cancellable biometrics: a survey[J]. CAAI Transactions on Intelligence Technology, 2024, 9(1): 4-25.
[30] 李思远, 陈文燕, 朱文龙. 基于被动矩阵有机发光二极管屏的光学指纹采集系统[J]. 科学技术与工程, 2021, 21(12): 5005-5010.
Li S Y, Chen W Y, Zhu W L. Optical fingerprint acquisition system based on passive matrix OLED screen[J]. Science Technology and Engineering, 2021, 21(12): 5005-5010.
[31] Dunsin D, Ghanem M C, Ouazzane K, et al. A comprehensive analysis of the role of artificial intelligence and machine learning in modern digital forensics and incident response[J]. Forensic Science International: Digital Investigation, 2024, 48: 301675.
[32] Birler A, Schmidt T, Fent P, et al. Simple, efficient, and robust hash tables for join processing[C]//Proceedings of the 20th International Workshop on Data Management on New Hardware. New York: ACM, 2024: 1-9.
[33] Tanović A, Mezei I. Embedded parallel K-means algorithm evaluation on ESP32 across various memory allocations[C]//Proceedings of the 2024 IEEE East-West Design & Test Symposium (EWDTS). Piscataway: IEEE Press, 2025: 1-7.
[34] Zubair M, Iqbal M A, Shil A, et al. An improved K-means clustering algorithm towards an efficient data-driven modeling[J]. Annals of Data Science, 2024, 11(5): 1525-1544.

[作者简介]

顾红星（2002－），男，中国人民公安大学硕士生，主要研究方向为电子数据取证。



上官梦轩（1992－），男，温州市公安司法鉴定中心中级职称，主要研究方向为电子数据取证。

丁锰（1980－），男，中国人民公安大学副教授，主要研究方向为电子数据取证。

何易临（2003－），男，中国人民公安大学硕士生，主要研究方向为电子数据取证。

马万里（2002－），男，中国人民公安大学硕士生，主要研究方向为电子数据取证。