基于保形加密的民航旅客信息脱敏方法

杜宇浩，王保国，谭　玲，胡　楠
（中国航空结算有限责任公司，北京　101300）

摘　要：针对民航旅客数据量大、涉及个人敏感信息多、在特殊情况下需要用到原值的问题，研究了一种基于保形加密的民航旅客信息脱敏方法。首先，按照应用场景和敏感程度的不同，结合数据特点，筛选出合理字段并据此确定调整因子的值，然后利用保形加密框架加密敏感信息，从而避免脱敏结果单一。该方法在保证数据不失真的前提下，能有效减少各字段之间的关联性，增加密文破解难度，降低个人信息泄露风险；同时，在需要获取原值时可以高效还原数据，以满足各类业务场景需求。以民航电子客票数据为例，模拟相应数据进行实验，结果验证了所提算法的有效性和实用性。

关键词：电子客票；保形加密；民航信息；数据脱敏
中图分类号：TP309　　**文献标识码：**A　　**DOI：**10.19358/j.issn.2097-1788.2025.11.008
引用格式：杜宇浩，王保国，谭玲，等. 基于保形加密的民航旅客信息脱敏方法［J］. 网络安全与数据治理，2025，44(11)：45-51.

Civil aviation passenger data encryption method based on format-preserving encryption

Du Yuhao, Wang Baoguo, Tan Ling, Hu Nan
(Accounting Centre of China Aviation, Beijing 101300, China)

Abstract: In response to the large amount of passenger data in civil aviation, the involvement of sensitive personal information, and the need to use plain text in special situations, a method for desensitizing civil aviation passenger information based on Format-Preserving Encryption (FPE) was studied. Firstly, reasonable fields were selected according to different situations and sensitivity levels, combined with data characteristics, and the tweak value were determined accordingly. Then, sensitive information was encrypted using a FPE framework to avoid unique result. While ensuring that the data is not distorted, it can effectively reduce the correlation between fields, increase the difficulty of cipher text cracking, and reduce the risk of personal information leakage. At the same time, it can efficiently restore data when obtaining plain text to meet various situations requirements. This paper took civil aviation electronic ticket data as an example, simulated corresponding data for experiments, and the results verified the effectiveness and practicality of the proposed method.

Key words: electronic ticket; FPE; civil aviation data; data masking

0　引言

随着网络信息技术快速发展，各行业数字化建设不断推进，数据将逐步从信息化资产向生产要素转变，数据流转更加频繁。在数据开发、测试、生产、管理等过程中，数据传输链路长、参与人数多等客观情况给数据安全带来挑战。因此，如何降低信息安全风险，消除敏感信息泄露的威胁成为学者广泛关注的问题。

数据脱敏技术通过预设方法对敏感信息进行数据变形，是解决上述问题的有效方法。文献［1］将 RSA 非对称加密标准与传统 MD5 方法相结合，应用于云计算的完整性校验和加密过程，有效解决信息泄露问题，但破坏了原有数据结构，导致脱敏后数据难以直接利用。文献［2］采用同态加密策略，实现了对过程工业中重要参数的脱敏，一定程度上保留了数据的原始信息，但加解密过程存在对密文数量的限制。实际上，民航领域广泛存在需要保留原有数据格式的业务处理场景。

保形加密（Format-Preserving Encryption，FPE）的提出为解决上述问题提供了思路。保形加密后的密文具有与原文相同的格式特征，能在脱敏的同时尽可能满足数据分析的需求。文献［3］针对大数据场景，以 Hadoop 平台为基础，设计了基于传统 FPE 算法的数据脱敏系统。文献［4］针对数值型数据，在此基础上改进，替换 FPE 中常采用的 Cycle-Walking 策略，显著提升了算法性能，使其在大规模脱敏的场景中更具优势。文献［4］将 FPE 算法引入特征识别领域，用于生物特征身份验证系统中的数据保护，在保留原始数据结构的同时，有效降低信息泄露风险。但民航旅客数据量大，高频、涉敏字段多，且各字段之间存在关联，如果获取到大量数据，仍存在被逆推破解的风险。

针对传统 FPE 的安全性问题，文献［5］将改进的 FPE 应用于民航领域，考虑到字段之间的关联关系，计算出各字段之间的关联度，据此筛选相关字段脱敏，增加了由密文逆推明文的难度，一定程度上提高了安全性。从本质上讲，FPE 要求数据脱敏前后有唯一的对应关系，导致攻击者在没有准确 FPE 密钥的情况下，可以根据大量明密文对应数据将脱敏数据重标识化，这一特点限制了其安全性上限[6]。针对此问题，文献［7］提出一种基于泛化 FPE 的动态脱敏方法，打破了传统 FPE 中明密文唯一映射关系，有效提高民航旅客信息的安全性，然而对于需要解密的场景，由于同一密文可能对应多个原文，此方法无法准确还原数据。

综上，本文基于 FPE 技术，结合民航旅客信息的应用场景，从改变单一解密密钥角度出发，根据数据特点，筛选出非敏感的必要字段，并据此赋值调整因子 T，在保留原数据格式的前提下，建立数据明密文多对一的映射关系。在提升数据安全性的同时，根据需求能够有效还原数据原值，在民航领域具备更强场景适配能力。

1　保形加密

1.1　FPE 定义

保形加密最初用来解决有限字符集加密问题，称为基本 FPE，可描述为：

$$E:\ K\times X\to X \tag{1}$$

其中 K 为密钥空间，X 为消息空间，确保密文和明文属于相同的消息空间。

进一步，考虑到加密空间的复杂性，为了更完整描述 FPE 问题，将一般性 FPE 描述为：

$$E:\ K\times \Omega\times T\times X\to X\cup\{\perp\} \tag{2}$$

其中 K 为密钥空间，Ω 为格式空间，T 为调整因子空间，X 为消息空间，$\perp\notin X$ 且所有空间非空。即对于任意给定 $\omega\in\Omega$，可确定消息空间 X 的子空间 X_ω。给定密钥 $k\in K$ 和调整因子 $t\in T$，FPE 可定义为在 X_ω 上的置换 $E_k^{\omega,t}$。

一般化 FPE 算法具体步骤可描述为：

（1）初始化。选取相应的 FPE 模型，初始化算法基础参数，定义保留格式 ω 以及由 ω 确定的子空间 X_ω，初始化对称密钥 k，调整因子 t。

（2）加密。输入明文 x，经 $E_K^{\Omega,T}(X)=E(K,\ \Omega,\ T,\ X)$ 置换，若 $x\in X_\omega$，返回密文 $y=E_k^{\omega,t}(x)$；若 $x\notin X_\omega$，则返回 $\perp$。

（3）解密。输入密文 y，根据调整因子 t 和密钥 k，经步骤（2）逆运算 $D_K^{\Omega,T}(Y)=D(K,\ \Omega,\ T,\ Y)$，若 $y\in Y_\omega$，返回明文 $x=D_k^{\omega,t}(y)$；若 $y\notin X_\omega$，返回 $\perp$。

1.2　FPE 方法

Prefix、Cycle-Walking、Generalized-Feistel 构建方法[8]是解决整数域 Z_n 上 FPE 问题的三种基本方法。Prefix 通过对消息空间的加密和排序形成置换表，能快速实现整数小于 10^6 的加解密问题。Cycle-Walking 利用分组加密策略，但加密过程中频繁的迭代会导致算法存在性能问题。Generalized-Feistel 通过更灵活的方式，可自定义分组密码大小，能在解决较大消息空间的加密问题的同时保证加密效率，成为一种安全高效的 FPE 通用方案。

基于这些基本方法，RtE（Rank then Encipher）和 FFX（Format-preserving，Feistel-based，X 是参数选择等相关因素）模型将解决的问题域由 Z_n 扩展到 chars^n，逐渐成为解决一般化 FPE 问题的通用策略，其中 chars^n 为长度为 n 的字符串构成的集合。文献［9］将其归纳为编码后加密（Coding then Encipher，CtE）模型，其示意图如图 1 所示。

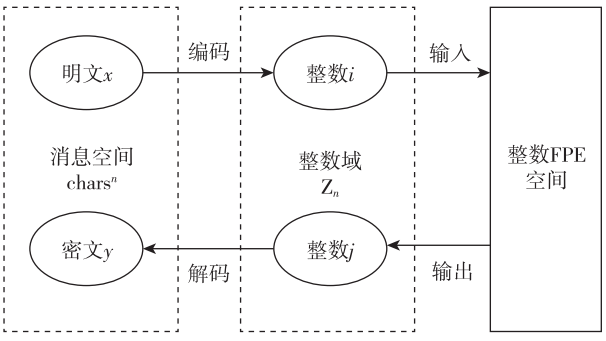


图 1　CtE 模型示意图

CtE 模型首先将字符域映射到整数域，建立两个域之间数据的唯一对应关系，转化为整数之后，再用 FPE 基本方法加密，最后将加密后的整数逆映射回原有问题域。CtE 模型间接地将加密类型从整数域扩展到字符域，广泛应用于加密定长编码字符型数据。

1.2.1 RtE 模型

RtE 模型通过对消息空间内元素按照一定规则进行排序，形成一个指定的有序队列，建立明密文和索引表的唯一对应关系，通过算法中定义的 rank 和 unrank 过程将 $chars^n$ 加密问题转换到整数域，成为一种更通用的方法。其原理如图 2 所示。

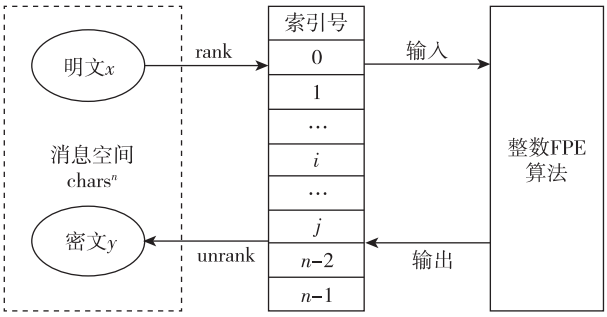


图 2　RtE 模型示意图

RtE 模型已经应用于解决中文 FPE 问题。首先将待加密中文的 Unicode 码与常用中文范围的第一个汉字 Unicode 码作差，完成 rank 过程，即得到索引号，然后在该索引范围内进行整数 FPE 加密，最后将输出结果通过 unrank 过程映射回中文域，即得到加密后的中文。

1.2.2 FFX 模型

NIST[10] 提出的 FFX 模型采用非平衡的 Feistel 网络，通过建立固定字符表与索引表的双射关系，将字符串中的每个字符编码为数字串，对数字串应用 Feistel 加密方法，最后将数字串解码，即实现了对消息空间 $chars^n$ 的加密，其原理如图 3 所示。首先将编码后的数字串按照长度不均等分割，作为左右两部分输入，然后将其中一部分数据经轮函数 F_K 运算，运算后的输出形成局部子密钥作为下一轮的输入，交替进行，将最后一轮运算出的输出与另一部分合并，形成加密结果。根据运算轮数的不同，符合安全性的 FFX 模型主要有 FF1 和 FF3，分别对应 10 轮和 8 轮运算。

现有的 FPE 算法大多在 Feistel 网络基础上进行改进[11]，实现保留格式加密目的。特别的，为了加强算法安全性，尤其有效防止字典攻击，FFX 模型在轮函数 F_K 运算过程中新增对调整因子 T 的支持，增加了子密钥的复杂程度。调整因子 T 的取值十分灵活，可采用自定义的公开数据。

2　民航旅客数据应用及脱敏场景

民航业作为客票电子化起步较早、数字应用较成熟的行业，系统中已广泛存在多种类型的数据，例如旅客订座记录（Passenger Name Record，PNR）、电子收费凭证

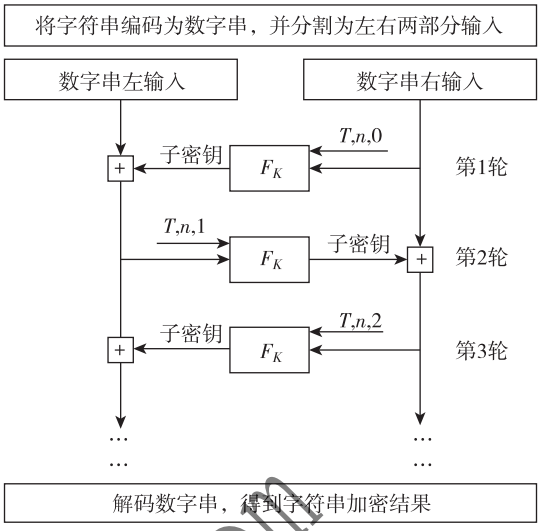


图 3　NIST FPE 示意图

（Electronic Miscellaneous Document，EMD）、民航电子客票（Electronic Ticket，ET）等，这些数据根据业务需要，应用于订票、离港、值机、结算等系统中，但不同数据仍具有相互关联、系统间共享的特性[12]。下面将以民航电子客票为例进行阐述。不失一般性，其他类型数据也存在相似字段和相关应用，在多数场景下同样适用。

2.1　民航电子客票

民航电子客票是以电子形式反映客票销售记录、客票使用状态以及相关服务情况等信息的票据。民航电子客票为民航领域中的关键数据，按照内容分类，可分为票面信息、票联信息、旅客个人信息、历史操作记录等内容。

票面信息主要包含票号、出票航司、客票状态、出票日期等，是客票的基本信息，根据业务需求，选取常用相关字段入库存储，见表 1。

表 1　电子客票票面信息相关字段

字段名	数据组成	数据类型	样例数据
票号	航司代码 3 + 票号 10	Varchar（13）	7881234567890
出票航司	航司二字码（英文）	Varchar（2）	CA
客票状态	状态（英文）	Varchar（1）	R
出票日期	年月日（数字）	Varchar（8）	20120101
起始城市	城市三字码（英文）	Varchar（3）	LXA
目的地城市	城市三字码（英文）	Varchar（3）	SHE

票联信息主要包含各航段承运航司、航班号、起/降日期、起/降时间、起/降机场、客票金额、税费等，描述的是一张客票多航段的详细记录，常用存储字段见表 2。

表 2　电子客票票联信息相关字段

字段名	数据组成	数据类型	样例数据
航段编号	数字	Int	1
承运航司	字母	Varchar（2）	CA
航班号	数字	Varchar（4）	8888
起/降日期	数字	Varchar（8）	20120101
起/降时间	数字 2：数字 2	Varchar（5）	18：00
起/降机场	英文	Varchar（3）	PKX
常客卡号	航司二字码(2) + 卡号	Varchar（8）	MU12345678
含税总价	数字金额	Varchar（40）	733. 58
货币类型	英文	Varchar（3）	CNY

旅客个人信息主要包含姓名、身份证号、护照号、出生日期、常客卡号、信用卡号等敏感信息，涉及旅客隐私，常用存储字段见表 3，为保护隐私，用#隐去真实值。

表 3　电子客票旅客个人信息相关字段

字段名	数据组成	数据类型	样例数据
中文姓名	姓(中文) + 名(中文)	Varchar（40）	王一#
英文姓	姓（英文）	Varchar（40）	WANG
英文名	名（英文）	Varchar（40）	YI ###
身份证号	地址码 + 生日 + 顺序码 + 校验码（数字）	Varchar（18）	142300########2771
护照号	字母 + 数字	Varchar（9）	E18###321
出生日期	年月日（数字）	Varchar（8）	20120101
常客卡号	航司二字码(英文) + 卡号(数字)	Varchar（10）	CA12####78
信用卡号	信用卡号（数字）	Varchar（40）	4381########8598

历史操作记录主要包含历史操作类型、操作日期、操作时间、操作代理人号/IATA 号等，记录了本张客票从产生到接收时刻的每一步详细操作信息，方便追溯客票操作流程，常用储存字段见表 4。

表 4　电子客票历史操作记录相关字段

字段名	数据组成	数据类型	样例数据
历史记录顺序号	数字	Int	1
历史操作类型	英文	Varchar（4）	NFMT
操作日期	年月日（数字）	Varchar（8）	20230101
操作时间	数字 2：数字 2	Varchar（5）	09：01
操作代理人号/IATA 号	数字	Varchar（10）	8888
历史记录文本	自由文本	Varchar（100）	L/F FLOWN

2. 2　民航旅客数据脱敏场景

电子客票通过详细的记录以实现客票从销售、使用、结算各环节的全流程应用。在实际中会涉及多部门、多公司间的数据交换，甚至有可能共享给其他国家。

根据民航局 MH/T 5057—2021《智慧民航数据治理规范数据安全》、国标 GB/T 35273—2020《信息安全技术 个人信息安全规范》，为了有效保障数据安全，达到民航数据安全分级分类、全生命周期分级防护的目标，同时结合实际业务需要，定义相应的脱敏方式为：（1）对于英文姓名，整个字段全部脱敏；（2）对于身份证号，保留前 6 位和后 4 位数字，对 7 ~ 10 位的生日进行脱敏；（3）对于护照号，保留第 1 位英文，之后数据进行脱敏；（4）对于常客卡号，保留前两位航司代码和后两位数字，其余部分进行脱敏；（5）对于生日，全部进行脱敏。具体方式及样例数据见表 5，* 代表脱敏后的结果。

表 5　敏感信息脱敏方式

字段	脱敏方式	脱敏前数据	脱敏后数据
中文名	全部	王##	***
英文名	全部	WANG ### ###	**** *** ***
身份证号	保留前 6 位和后 4 位	142300########2771	142300 ******** 2771
护照号	保留第 1 位	E18###321	E ********
出生日期	全部	201201##	********
常客卡号	保留前 2 和后 2 位	CA12####78	CA ****** 78

从表 5 可以看出，由于民航旅客数据具有多维信息，大多字段都可以作为个人唯一标识，且敏感信息各字段间存在关联，特别的，身份证号和出生日期，中文名和英文名，具有直接对应关系。然而，为保证数据可用性，这些信息在不同系统中通常需要保持脱敏一致性。这一特性导致脱敏数据泄露风险高，且存在较大破解风险。

3　基于 FPE 的民航旅客信息脱敏方法

3. 1　脱敏算法

在民航旅客数据的流转使用中，包括但不限于电子客票的票号、PNR 数据的 PNR 号以及 EMD 的 EMD 号按照相关规则在数据产生时同步生成，作为各自数据的最基础、最重要的标识。本文结合民航数据旅客信息的应用场景，利用 FPE 方法对数据加密前后的结构保留能力，减少脱敏对数据可用性的影响。考虑到数据各字段间具有关联性，导致易被破解这一特点，从突破传统 FPE 明密文唯一映射角度出发，利用非敏感必要字段的多样性与随机性，提出基于保形加密的民航旅客信息脱敏（Civ-

il Aviation passenger Data encryption based on FPE，CADFPE）方法，图 4 为 CADFPE 方法的算法示意图。

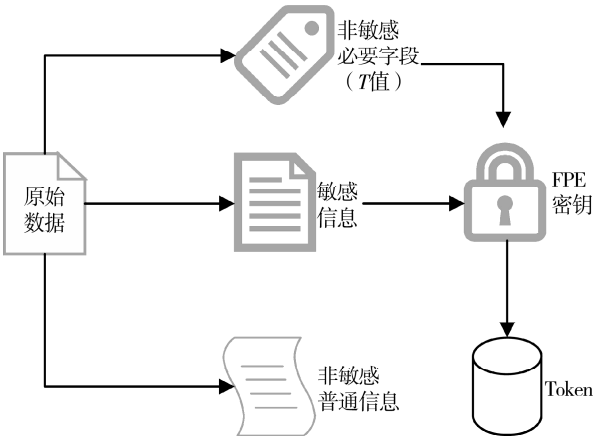


图 4　CADFPE 方法示意图

如图 4 所示，CADFPE 方法需要分析原始数据，预先选择非敏感必要字段 T，选取不同系统中一致且共享的业务字段作为 T 值。然后将原始数据中的敏感字段和 T 值组合配对，与 FPE 静态密钥参与 FPE 过程最终得到与原始字段格式相同的加密值，即 Token。

结合图 4 与相关方法理论的描述，CADFPE 方法步骤如下：

算法加密步骤：

输入：敏感字段 x，非敏感必要字段 T，轮次数 r，密钥 K，轮函数 F_K（·）

输出：密文 y

（1）建立敏感字段 $x \in \text{chars}^n$ 与整数 $z \in Z_n$ 的映射关系，即 $x \rightarrow z$；

（2）根据 T、r、n 以及密钥 K 确定轮函数 $F_{K,n,T,r}(x)$，将 z 代入得到输出 $z' = F_{K,n,T,r}(z)$；

（3）将 $z' = F_{K,n,T,r}(z)$ 映射回 chars^n，得到密文 y。

解密方法为加密的逆过程。

以电子客票为例，结合民航数据应用场景，对其加密过程具体为：

（1）分析电子客票数据，根据业务拆分的票面、票联、旅客个人信息均可由同一票号串联，因此选取票号作为非敏感必要字段 T，旅客个人信息作为敏感字段 x；

（2）将各字段 x_i 映射到整数域 Z_{n_i}，对于同一客票，有确定的 $F_{K,n,T,r}(Z_n)$，即得到唯一 $Z_{n_i}' = F_{K,n,T,r}(Z_{n_i})$；

（3）将 Z_{n_i}' 映射回 chars^n，得到密文 y。

当解密时，除要求传统 FPE 所需的 K、n、r 之外，必须同时有票号信息，才能确定 $F_{K,n,T,r}$（·），得到正确的原文。

3.2　民航旅客信息脱敏模型

结合民航信息的实际使用场景，从数据安全等级角度，可将应用系统分为数据安全治理系统、业务处理系统。数据安全治理系统存储 FPE 密钥、用户访问权限信息，承担敏感信息的脱敏、还原、用户权限管理等功能。业务处理系统负责后续业务逻辑处理，保存业务过程数据、非敏感信息、经过加密后的旅客信息。民航旅客信息脱敏模型如图 5 所示。

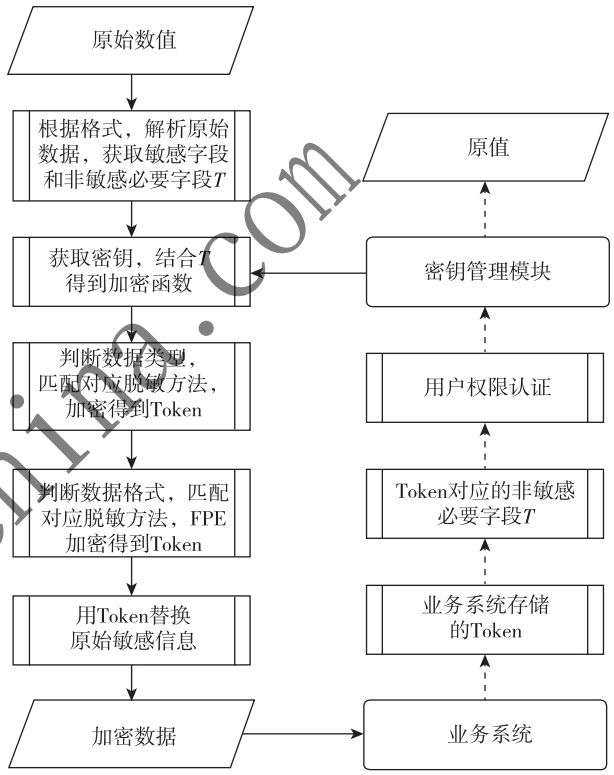


图 5　民航旅客信息脱敏模型

民航旅客信息脱敏模型包括密钥管理、数据解析、数据加密、数据解密、用户权限管理 5 个模块。密钥管理模块负责保存和对外提供 FPE 静态密钥；数据解析承担按定义规则拆分原始数据、将 Token 替换回原位置的功能；数据加密和数据解密模块通过调用密钥管理模块实现数据加解密；用户权限管理负责权限认证，保证数据安全。

当数据接入时，任何可能涉及敏感信息的原始文件须先经数据安全治理系统处理，根据不同数据类型中的字段划分规则，将数据解析为非敏感普通字段 N、非敏感必要字段 T、敏感字段 X，然后应用 CADFPE 方法加密敏感字段生成 Token，再将 Token 替换到原始数据中原敏感信息的相同位置，保证下游业务系统正常处理数据。数据安全治理系统与其他系统相互隔离，当业务系统需要查询

敏感信息原值时，需提供 Token 与其对应的非敏感必要字段，经用户权限验证之后，方可调用数据解密模块。

4 实验及结果分析

4.1 实验数据

按照民航电子客票格式随机生成 21 370 条模拟数据作为实验数据，主要包括票号、票面金额等业务信息和中英文姓名、身份证号、护照号、常客卡号、生日等个人敏感信息，选取部分模拟电子客票数据如表 6 所示。

4.2 实验结果

为了验证 CADFPE 算法在民航领域的有效性和实用性，将随机生成的模拟电子客票数据经 CADFPE 算法加密，其中选取表 6 相对应的结果如表 7 所示。

表 6 电子客票模拟数据

序号	票号	中文名	英文名	身份证号	护照号	出生日期	常客卡号
1	9995428350954	贾彬	JIA BIN	612702196706075077	E683034185	19670607	CA61858342
2	9998762416893	崔凯	CUI KAI	511122198503152827	E124620163	19850315	CA59520024
3	9992132514567	李欣怡	LI XIN YI	420184200506060037	E242000832	20050606	CA54220587
4	9995428542445	李亮	LI LIANG	441981200307292657	E264263570	20030729	CA04264357
5	9995482283666	刘晗日	LIU HAN RI	610114198205167598	E164550241	19820516	CA47153011
6	9995449787586	美玲	MEI LING	152315201109305266	E174804694	20110930	CA71366025
…	…	…	…	…	…	…	…
7628	9998762305342	吴语涵	WU YU HAN	140122196509206031	E291954289	19650920	CA66584045
…	…	…	…	…	…	…	…
21370	9992321342458	顾涛	GU TAO	440182200303190023	E438357708	20030319	CA90935771

表 7 电子客票模拟数据脱敏结果

序号	票号	中文名	英文名	身份证号	护照号	出生日期	常客卡号
1	9995428350954	犹礅	JAV VNJ	612702938626975077	E800808730	19780415	CA85259342
2	9998762416893	耪潇	SFF EOZ	511122980009462827	E376835059	19730414	CA71471624
3	9992132514567	嵌枹葙	OE WXR KW	420184291310300037	E171800232	19870517	CA59618587
4	9995428542445	醜俳	YY YNJYA	441981269524692657	E324804119	19730414	CA37719257
5	9995482283666	薮縓应	LYU EFO SK	610114983406477598	E023910141	19761029	CA24992211
6	9995449787586	陧翻	RAI RCBC	152315457403035266	E947028690	20020531	CA58982825
…	…	…	…	…	…	…	…
7628	9998762305342	猩迪锤	GZ HC ELB	140122702025386031	E023910141	19850129	CA21888545
…	…	…	…	…	…	…	…
21370	9992321342458	畀支	JP PFA	440182828461390023	E299395571	19860814	CA09050271

根据表 7 的结果可以看出，第 2 条和第 4 条出生日期字段原值虽然不同，但是经 CADFPE 算法处理后得到相同的加密结果“19730414”，第 5 条和第 7628 条的护照号原值不同，加密结果同为“E023910141”，第 3 条和第 4 条的英文姓氏均为“LI”，但加密之后分别为“OE”和“YY”。在此情况下，通过获取大量明密文对应关系而逆推明文的难度显著增加。

4.3 实验结果分析

相关系数（Correlation Coefficient）常用来描述数据间的关联程度。不同变量的相关系数绝对值越大，字段之间关联度越高，反之关联度越小，见式（3），其中 cov（X，Y）为变量间的协方差，σ 为方差。

$$r(X,\ Y)=\frac{\mathrm{cov}(X,\ Y)}{\sqrt{\sigma_X\sigma_Y}} \tag{3}$$

本文选择三组关联字段，分别计算出数据原值、经过 FPE 算法处理和经过 CADFPE 算法处理后的相关系数，作为密文破解难度的评价指标。对比结果如表 8 所示。

根据表 8 中三组数据的实验结果，传统 FPE 算法处理后字段间相关系数与数据原值的相关系数相近，CADFPE 加密后各字段间的相关系数变小，可以看出 CADFPE 算法有效降低了数据字段间的关联程度，增加了密文破解难度。

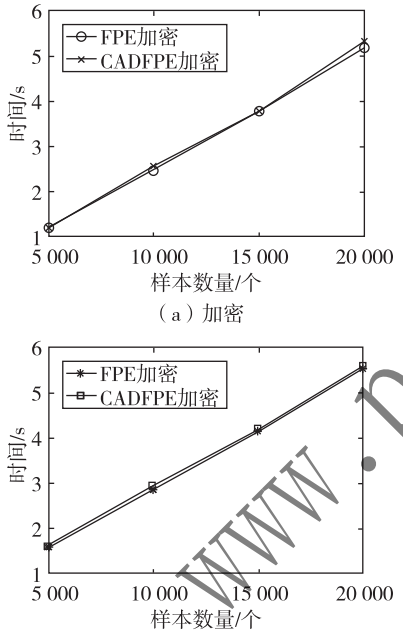
表 8　不同方法字段间相关系数对比结果

Item	数据原值	FPE	CADFPE
身份证号和生日	0.70	0.58	0.31
中文名和英文名	0.39	0.33	0.15
常客卡号和身份证号	0.41	0.47	0.19

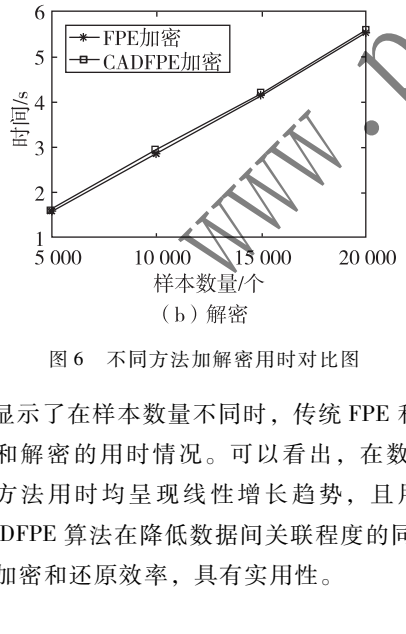
4.4　算法效率分析

实验操作系统为 Windows 10 专业版（64 位），处理器为 Intel（R）Xeon（R）Gold 5120 CPU @2.20 GHz，内存（RAM）8 GB，编程语言使用 Java 8。

为直观展现 CADFPE 算法效率，选用身份证号作为加解密验证字段。根据样本数量设置四组实验，分别对应数据量为 5 000、10 000、15 000、20 000。采用传统 FPE 方法和 CADFPE 方法对每组数据进行加解密处理，其用时统计结果如图 6 所示。



（a）加密



（b）解密

图 6　不同方法加解密用时对比图

图 6 显示了在样本数量不同时，传统 FPE 和 CADFPE 方法加密和解密的用时情况。可以看出，在数据量增加时，两种方法用时均呈现线性增长趋势，且用时接近。因此，CADFPE 算法在降低数据间关联程度的同时，不会降低数据加密和还原效率，具有实用性。

5　结论

本文针对民航业数据高度敏感、涉及业务量大，结合行业特性，引入基于保形加密的民航旅客信息脱敏（CADFPE）方法，通过选取非敏感必要字段赋值动态密钥，改变了传统 FPE 中明密文唯一对应关系，在保证民航业务正常运行的前提下，有效降低字段之间的关联关系，加大了密文破解难度，减少敏感数据泄密风险。并且针对民航数据特定应用场景，设计了民航旅客脱敏模型，利用模型中数据解析、加解密、用户访问管理等模块实现民航数据有效加密和传输，使算法具备实际应用价值。最后通过对民航客票模拟数据的实验结果表明，CADFPE 算法提升了 FPE 算法的安全性，有效降低数据破解风险。此外，CADFPE 算法具备一定的泛化性，下一步工作将继续深入研究，为拓展到其他应用场景提供思路。

参考文献

[1] NICHOLAS K, WILSON C, MUTHONI A. Enhancing confidentiality and integrity in cloud computing using RSA encryption standard and MD5 hashing algorithm [J]. International Journal of Computer Applications, 2018, 181 (14): 23 - 27.

[2] CHOKPAROVA Z, URBAS L. Application of multiplicative homomorphic encryption in process industries [M]. Computer Aided Chemical Engineering, 2022, 51: 1267 - 1272.

[3] 卞超轶，朱少敏，周涛. 一种基于保形加密的大数据脱敏系统实现及评估 [J]. 电信科学，2017，33（3）：119 - 125.

[4] 王浩，张永平，李同寒，等. 一种数值型的保留格式加密算法 [J]. 信息安全研究，2023，9（8）：745 - 753.

[5] 顾兆军，蔡畅，王明. 基于改进保留格式加密的民航旅客数据脱敏方法 [J]. 信息网络安全，2020，21（5）：39 - 47.

[6] BANSAL V, GARG S. A cancelable biometric identification scheme based on bloom filter and format-preserving encryption [J]. Journal of King Saud University-Computer and Information Sciences, 2022, 34 (8): 5810 - 5821.

[7] 丁建立，陈盼，马勇. 基于泛化 FPE 加密的民航旅客信息动态脱敏方法研究 [J]. 信息网络安全，2021，21（2）：45 - 52.

[8] BLACK J, ROGAWAY P. Ciphers with arbitrary finite domains [C]// Cryptographers Track at the RSA Conference on Topics in Cryptology. Springer-Verlag, 2002: 103 - 114.

[9] 刘哲理，贾春福，李经纬. 保留格式加密模型研究 [J]. 通信学报，2011，32（6）：184 - 190.

[10] BELLARE M. Format-preserving encryption [C]//DBLP, 2009: 295 - 312.

[11] 李敏. 保留格式加密技术应用研究 [D]. 天津：南开大学，2012.

[12] TAPLIN K. South Africa's PNR regime: privacy and data protection-ScienceDirect [J]. Computer Law & Security Review, 2021: 40 - 43.

（收稿日期：2025 - 08 - 28）

作者简介：

杜宇浩（1994 - ），男，硕士，工程师，主要研究方向：数据治理、大数据。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部