保留格式加密技术在民航信息系统数据处理中的应用研究

刘 俊 李泽昊 苏国宇 李婧雯
(中国民航大学计算机科学与技术学院 天津 300300)

摘 要 采用保留格式加密技术加密航班信息、旅客信息、客票信息等民航信息系统数据,选取简单、有效的 FF1 算法和 FF3 算法以及更为灵活的混合格式加密算法 IFX,采取高位偏移、变长格式、数值间接映射、以数据库的 id 号为随机因子等策略优化扩展算法功能。在保证信息安全性、完整性和真实性的同时,保留民航信息系统数据的内在格式,支持在不解密数据的条件下实现对数据的统计分析,降低数据泄露的风险。

关键词 保留格式加密,混合格式加密算法,航班信息加密

中图法分类号 TP301 **文献标识码** A

Application of Reserved Format Encryption Technology in Information Processing of Civil Aviation Information System

LIU Jun LI Ze-hao SU Guo-yu LI Jing-wen
(College of Computer Science and Technology, Civil Aviation University of China, Tianjin 300300, China)

Abstract The usage of format-preserving encryption technology in the encryption of flight information, passenger information and ticket information was studied by choosing simple and effective algorithms such as FF1 algorithm, FF3 algorithm and a more flexible hybrid format encryption algorithm IFX. The functions of high bit offset, forward expansion, numerical indirect mapping, and using ID of database as random factors are used to optimize the function of the extended algorithm. While ensuring security, integrity and authenticity of the information, the internal format of flight data is kept, and the statistical analysis of civil aviation information is implemented without decryption, so as to reduce the risk of data leakage.

Keywords Format-preserving encryption, IFX, Flight information encryption

1 引言

在大数据时代,任何有用的数字信息都可能被用于数据挖掘,海量数据中时常还会包含一些敏感信息,因此保证大量信息数据的安全性,尤其是敏感信息的安全性,是大数据时代必须面对的挑战。互联网技术的迅速发展带动了全球电子商务的兴起,各行各业都在迅速发展电子商务,民航产业也不例外。庞大的民航信息网络带来了大量的用户数据,这些数据中不可避免地带有用户的隐私数据。中国互联网中心发布的《中国互联网发展状况统计报告》[1]表明,电子商务中,超过 52.26%的用户最关心的问题是用户自身数据的安全。采用传统的加密方式来确保数据不泄露时,密文数据的类型变化和长度变化,会对之后数据的利用以及存储带来巨大的麻烦。

其中民航信息系统的数据非常具有代表性,这些数据通常有严格的格式,例如乘客身份证号、航班代码等,如果采用常规的加密方式对这些数据进行加密后,需要花费巨大的开销来修改存储这些数据的数据库结构或者修改相关应用程序来适应密文的变化。当然,支付卡行业数据安全标准[2](Payment Card Industry Data Security Standard, PCI DSS)的提出与应用,以及消费者保护个人隐私意识的提高,促使人们探索发现新的技术来加密信用卡号等个人识别信息(Personally Identifiable Information, PII)。其中一类新的加密技术能够保证明文和密文具有相同的格式,这类加密技术被称为"保留格式加密"(Format-Preserving Encryption, FPE),这类加密技术为解决格式要求比较严格的数据保密带来了新的思路。

目前在信用卡号、社会保险号等敏感数据的数据库加密存储上,已经有公司付诸实践,应用保留格式加密算法完成了信用卡号、社会保险号等数据的加密存储。本文研究民航信息系统数据的特征,结合当前已经成熟的 FPE 算法,提出一种能够按照保留格式的要求加密民航信息系统相关数据的算法。

2 FPE 简介

2.1 FPE 的两种定义

定义 1(基本 FPE) FPE 可以简单描述为一个密码:

$E: K \times X \to X$

其中,K 为密钥空间,X 为消息空间。

定义 2(一般化 FPE) FPE 可以描述为一个密码:

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

其中,K 为密钥空间,Ω 为格式空间,T 为调整空间,X 为消息空间,所有空间都非空,且 $T \notin X$。

刘 俊(1996-),男,主要研究方向为网络安全、大数据及相关安全技术,E-mail:jliu_1996@163.com(通信作者);李泽昊(1996-),男,主要研究方向为密码学、区块链;苏国宇(1997-),男,主要研究方向为网络安全、可信计算;李婧雯(1997-),女,主要研究方向为网络安全、密码学。

2.2 FPE 的加密思想和特点

2.2.1 FPE 的加密思想

FPE 算法使用编码后加密的思路，其具体操作如图 1 所示。在一次完整的加密过程中，先使用预置的编码表将明文编码为数值串 N，加密的核心操作是实现从合法 N 到合法 N 的映射，最后把加密后的 N 反编码为密文 Y。同理可以实现 FPE 的解密操作。

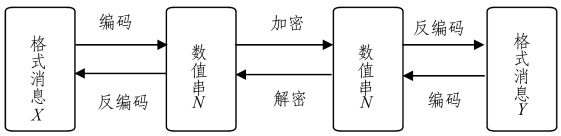

图 1 编码加密流图

2.2.2 FPE 的加密特点

FPE 用于处理具有格式要求的数据，主要用于对数据库敏感字段的加密，其保证加密字段满足数据库的约束性。一般类型 FPE 的加密特点如下：

$$FPE.Enc(K,X,T)=Y \tag{1}$$

$$FPE.\mathrm{Dec}(K,FPE.Enc(K,X,T),T)=X \tag{2}$$

$$length(X)=length(Y) \tag{3}$$

$$type(X)\in\nabla,type(Y)\in\nabla \tag{4}$$

X 为带格式的明文，Y 为带格式的密文；扭转因子 T 为一种随机因子，用于提高加密的安全性；K 为分组加密算法的密钥；∇ 为格式空间；type 函数用于提取消息格式，length 函数用于获取消息的有效长度。式(3)强调加密不会有数据位的扩展。式(4)表示明密文满足相同的格式约束，比如有相同的数据类型，属于相同的正则空间等。

2.3 问题分析及算法概述

2.3.1 字段格式约束分析

为了不失一般性，同时避免冗余，在表 1 中展示了民航信息系统数据库中具有代表性的字段及其需要满足的格式约束。现归纳在该场景中的 FPE 算法需要满足以下要求：1)既可用于 Z_n 空间也可用于 $char^n$ 空间；2)既可用于单个值空间 Z 也可用于串空间；3)易于实现数据库合理性约束；4)加密格式易于从一般格式空间扩展到复杂正则空间；5)算法本身的效率、稳定性、安全性要有保证。

表 1 字段名及字段特点

字段名	类型	问题域	合法性校验	正则空间	字段特点
航空公司	char[3]	$char^n$	三字码校验	[A-Z0-9]{3}	单一格式串/值
起/降机场	varchar[4]	$char^n$	—	[A-Z]{3,4}	单一格式串
航班号	char[6]	$char^n$	—	[A-Z0-9]{2}|[0-9]{4}	混合格式串
起/降时间	date	特殊域	—	—	特殊格式
旅客姓名	varchar[20]	$char^n$	—	—	单一格式串
手机号	int[11]	Z_n	格式校验	[0-9]{11}	单一格式串
年龄	int	Z	[0,999]	—	值
身份证号	char[18]	$char^n$	格式校验	[0-9]{17}|[0-9X]	单一格式串
票价	double	$char^n$	[0,9999]	—	特殊格式

注：Z_n 即串类型的数值，$char^n$ 即串类型的特殊字符，单个值 Z 是 Z_n 的串长退化为 1 的特殊情况；本文后续提到的值类型问题即处理 Z 空间上保留格式加密，串类型问题即处理 Z_n 和 $char^n$ 空间上的保留格式加密

2.3.2 算法选取

针对 2.3.1 节的问题，现选取合适的 FPE 算法。在 FFX 的模型框架下，2016 年，Dworkin[4] 提出了两种具有代表性的 FPE 算法——FF1 和 FF3。这些算法将加密类型从 Z_n 扩展到一般的类型字符空间 $Char^n$，同时扩展了加密的范围。在安全性方面得到了一系列的保证[5-7]。

FF1 和 FF3 可以实现单一格式的 FPE，比如姓名、年龄、手机号等。所谓的单一格式即这些字段的正则空间不需要组合。但对于航班号这种组合模式的正则空间则需要使用混合模式的 FPE。

在 2016 年 12 月，Johnson 提出了用于混合格式的加密模型 IFX[13]。IFX 模型在 FF1/FF3 算法上进行了进一步改进，其加密特点如图 2 所示。图中串的每个字符 x_i 映射为了对应位置上的 y_i，可以看出该加密进一步保留了字段中各字符的位置和类型信息，该模型可以实现航班号等字段的 FPE。

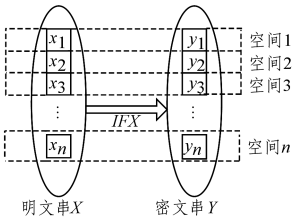

图 2 IFX 加密效果展示

2.3.3 算法概述

现简要阐述选取到的 FF1，FF3，IFX 算法的加密流程。这 3 种算法都使用了交互式的 Feistel 结构[3]，该 Feistel 结构的特点如下：1)支持对输入分组的不均等划分；2)迭代轮次的奇偶性影响每一轮次的操作；3)一次性吸收整个加密字段；4)轮函数可以同时实现消息的扩展和压缩；5)使用模运算和相应的取整策略控制输出消息的长度。

考虑到算法 FF1，FF3，IFX 之间具有的相似性，现将这些算法的加密概图统一用图 3 表示，加密流程也可以统一概括为如下步骤：

1)将字符串 X 中的每个字符 x 映射到该字符在编码表中编排的序号 i，$char^n$ 空间上的问题转化到了 Z_n 空间上。

$$N=Encode(X)$$

2)将 Z_n 空间上的数值分组 N 压榨为数值 b 输入到 Feistel 网络。

$$b=Num_{radix}(N)$$

3)轮函数函数 $f_k(n,T,i)$ 或 $f_k(n,T,i,W)$ 依靠分组密码算法生成伪随机偏移量 y。

FF1 和 FF3：$y=\mathrm{Num}(f_{\mathrm{CHIP128}}(K,b,T,i,n))$

IFX：$y=\mathrm{Num}(f_{AES}(K,b,T,i,W,n))$

4)将另一分组 a_i 与偏移量 y 执行混合相加。转步骤 2)反复迭代多轮。

$$a_{i+1}=(a_i+y)\bmod d$$

5）将得到的值 c 映射回 Z_n 空间，最后映射回串空间，得到结果密文 Y：

$Y=Decode(STR^{n}_{radix}(c))$

以上步骤中 $Encode()$ 与 $Decode()$ 分别对应图 1 中的编码操作和反编码操作，IFX 的轮函数中引入的是用于描述混合格式信息的数组，$radix$ 表示编码表中编码元素的个数。其他函数的具体实现和其他符号的作用可参照文献[4,13]中的相关描述。

分析这些算法的具体实现，*IFX* 之所以可以用于混合格式，在于其与 *FF*1 和 *FF*3 有如下差异：1）*IFX* 使用多表编码，而 *FF*1 和 *FF*3 使用单表编码。2）*IFX* 算法额外输入描述混合格式信息的数组。进一步，*IFX* 与 *FF*1 和 *FF*3 还有其他方面的差异：1）*IFX* 使用可变轮次的 *Feistel* 网络，*FF*1 和 *FF*3 使用固定轮次的 *Feistel* 网络；2）*IFX* 需要运行预处理函数以初始化 *Feistel* 网络中一些必要的配置信息；3）*FF*1 和 *FF*3 使用 128 *bit* 的分组密码算法构建伪随机函数，*IFX* 使用 128，192，256 *bit* 的 *AES* 算法构建伪随机函数；4）*IFX* 可用于处理单一格式的问题，而 *FF*1/*FF*3 不能解决混合格式的问题。

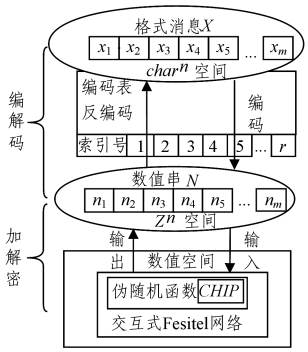

图 3　编码加密概图

2.4　适用于民航信息系统的 FPE

2.4.1　编码方案的选择

执行加密的首要操作是编码，编码将会直接影响后续加密过程。不同编码方案的选择可以实现不同的加密特性，如表 2 所列。

表 2　编码表

编码	字符集	特异性
ASCII	简单	—
UTF-8	复杂	变长格式 FPE 问题
Base64	简单	变长格式 FPE 问题
UTF-16，GBK2312，IOS-8859-1	复杂	—
GBK	复杂	中文域的 FPE 问题
自定义	简单	特定域的 FPE 问题

对简单字符集的编码可以使用 ASCII 表；对复杂字符集编码可以选用 UTF-8 或 UTF-16。UTF-8 的变长编码的特点可以混淆字段的长度，用于处理变长的 FPE 问题（详见 2.5.2节）等。注意，这些现成码表在使用前可能需要做适当的修正，比如 ASCII 表的前 32 个不可打印字符往往需要剔除，那么在实际的编解码操作中，可以加减 32 个偏移量以跳过这些字符。上述编码方式使用了现成的编码集，这种编码方式的优点是可以实现字符的按秩索引，加快编解码效率，同时可以省去构建编码表的时间。但是处理特定域上的 FPE 问题则需要手动构建编码表，这样的编码方式适合处理有针对性且字符集简单的 FPE 问题。

2.4.2　对值类型的处理。

处理字段：年龄、票价、航空公司

编码方案：自定义或者 ASCII 表

问题分析：1）年龄加密是典型的值问题，但必须注意 FF1，FF3，IFX 算法有对输入串长的约束，使其无法直接用于对单个值的加密，因此值类型在实现加解密时需要将值扩展成对应的数值串，随后使用一般方式的 FF1，FF3，IFX 算法加密即可（考虑到值扩展为数值串方法的多样性，具体的扩展方法和值类型加密的细节将在 2.5.1 节中单独说明）。2）类型的票价格式为 XXXX.YY，可以将 XXXX 与 YY 分别使用值类型的加密方法加密后拼接回类型。3）航空公司三字码在物理上是串类型，但数据库做三字码校验是以字段为整体的，因此应该将所有合法的三字码统一编码为一个数值以满足三字码校验，适合使用值类型的加密。

结论：逻辑上不可分的多个字符应该统一编码为一个数值。

2.4.3　对单一串类型的处理。

处理字段：机场代码、旅客姓名、手机号、身份证号

编码方案：旅客姓名用 GBK 或 UTF-8 等；其他类型使用自定义或者 ASCII 表

问题分析：这些字段是典型的单一格式字段，使用 FF1，FF3，IFX 即可。但是需要注意手机号开头 3 位为运营商号码，这样的信息不属于机密信息，同时还常用于作为合法手机号的校验方式，于是手机号只需要对后 8 位进行加密。同样地，身份证号有地区位和校验位等，这样的信息也不用加密。公开信息不加密既可以方便实现数据库的不解密索引，又可以进行字段校验。

结论：不具机密性且具有逻辑意义的字段部分不必加密。

问题扩展：单一格式的算法可扩展到混合格式。文献[8]提到先将混合格式切分为若干单一格式，对每部分单独使用单一格式加密，最后再拼接回混合格式的方法（类似于 4.6.2 节中类型的加密）。但对原本统一的字段进行分割使得加密无法实现雪崩效应[9]，同时割裂使得每部分消息密文空间变小，安全性降低。因此单一格式的 FPE 不适用于混合格式的字段。

2.4.4　对混合串类型的处理。

处理字段：航班号

编码方案：自定义或者 ASCII 表

问题分析：航班号前两位是航空公司二字字码，后若干位是数字，是典型的混合格式字段。加密保留加密字段中各字符的位置和类型信息。使用 IFX 即可实现。

2.4.5　对时间类型的处理。

处理字段：航班起/降时间

编码方案：自定义或者 ASCII 表

问题分析：FPE 处理时间这种复杂格式是十分吃力的，比如 XXXX-YY-DD：HH-MM-SS 这种格式，其每部分都有相应的约束，直接加密十分有难度。文献[10]提到了用于时间的加密方法，不过这种方法较为繁琐。考虑到在计算机中时间实际上是以数值存放的，即选的一个时间点为基准，用给定时间到该基准所有经过的毫秒数之和来表示给定时间。那么可以完全绕开时间的复杂格式，直接对表示时间的数值进行 FPE，这样时间类型的 FPE 问题就变成了数值类型的 FPE 问题。使用 FF1，FF3，IFX 皆可解决。

2.5 对 FF1/FF3/IFX 的修正

2.5.1 值类型问题修正为串类型问题

问题分析：在 2.4.2 节中提到了值类型的加密必须要先扩展到数值串，本部分将讨论扩展为数值串的两种方法。方法 1 使用直接拆分法，比如给定数值 $x=84\,310$，那么可以直接扩展为数值串 $N_x=\{8,4,3,1,0\}$输入到 FF1，FF3，IFX 中，并将编码表设定为集合$\{0,1,2,3,4,5,6,7,8,9\}$，通过加密，N_x 的每个元素将会在$[0,9]$上映射得到另一个数值串 N_y。但这意味着 N_y 的首位可能被加密为 0，导致 N_y 转为加密数值 y 时最高位直接丢失，使加解密不可逆。于是需要在 N_y 的首位添加大小为 1 的偏移量，使得首位从$[0,9]$的空间转到$[1,10]$，如此可以顺利完成加密。方法 2 使用前向填充拆分法，如图 4 所示。为了避免发生方法 1 中的高位丢失，在加解密前，可以在 $N_x=\{2,2,1,0,9\}$与 $N_y=\{8,4,6,5,2,9,5\}$前填充若干个 0 以达到一个事先约定好的串长上界，当 N_x 转为 x 和 N_y 转为 y 时，丢弃的 0 又被补回，实现可逆的加解密。

问题扩展：两种方法在使用上各有优劣，方法 1 不适合加密对数值位数敏感高的数据，比如航空公司年度收益额。因为对于 10 位的数值，通过方法 1 加密后仍是 10 位或者 11 位，这会暴露收益额的数量级。方法 2 可以实现变长数量级的加密，但加密前需要指明串长上界，这会降低加密的灵活性。

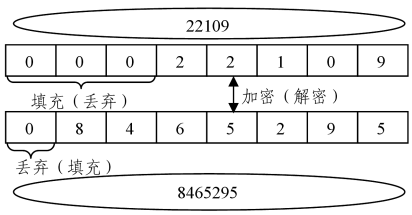

图 4 前向填充拆分图

2.5.2 定长格式到变长格式的修正

在 2.2.2 节中分析了保留格式加密的特点，其中 FPE 默认是保留长度信息的，即默认长度的信息是公开的，这使得 FPE 不适合加密长度敏感的数据信息，比如用户口令，因为用户口令通常不等长，加密的结果会暴露出口令的长度。

文献[11]中提到了对变长数据的处理，即采用变长编码 UTF-8。这种方法有一定局限性，因为一类字符（比如同为数字或字母）往往都编码为相同长度的字节数，这使得变长的特性在加密一类字符时体现不出。于是这里提出最大填充的解决方法。

假定数据库定义 $varchar[n]$类型存放变长数。算法步骤如下：1）根据消息长度生成若干位的长度标识信息，如图 5 所示，若消息长度为 m，则需要使用 t 位的长度标识，$t=\lfloor \log_{radix} m \rfloor$；2）将长度标识信息附加到原始消息的末尾，同时在中间填充 c 位字符。使得填充后的消息为字段最大长度，即 $c=n-m-t$。

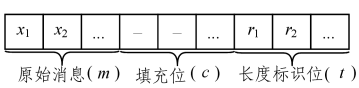

图 5 变长格式填充图

通过这种方法隐匿了消息的长度信息，同时解密时可以通过长度标识信息提取出有效位，使得 FPE 算法可以处理变长的格式。

2.6 扭转因子 T 的选择

扭转因子 T 的概念最早在分组密码中提出[11]，其作用类似于分组加密算法运行模式中使用的 IV 或者 hash 算法中的 $solt$，这一概念随后被规范化[12]，其具体实例可以参照文献[4]。

T 对 FPE 加密至关重要，但 T 的选择十分灵活。可行的方案有：1）用字段本身携带的公开信息，比如上文提到的手机号的前 3 位，身份证号对应的区号等；2）使用随机函数生成，这种方法即 $solt$ 与 IV 的生成方法，但这需要数据库分配专门的字段去存储或者使用专门的文件存储，该方法不灵活且需要不断维护信息；3）使用数据库记录的 ID 号，数据库中常使用自动递增的 ID 作为主码，这样的字段无实际意义，不用加密且适合作为 T。

3 算法性能分析

加密性能是每种加密算法必须考虑的问题，下面将对 FF1，FF3，IFX 的加密性能做测量分析，但考虑到 FF1，FF3，IFX 算法的性能受多个输入参数的影响，于是还将进一步讨论编码表大小与加密字段长度对加密性能的影响。

本实验的操作系统为 Microsoft Windows 10 专业版（64 位），CPU 为 Intel(R) Core(TM) i7-5500U CPU @ 2.40GHz (2394 MHz)，内存为 12GB，运行环境为 Java(TM) SE Runtime Environment (build 1.8.0_131-b11)。

3.1 *radix*（编码表大小）对加密性能的影响

使用随机数生成 10 万条字段长度为 10 的数值串，使用不同的加密算法加密并测量运行时间。得到如图 6 所示的柱状图。

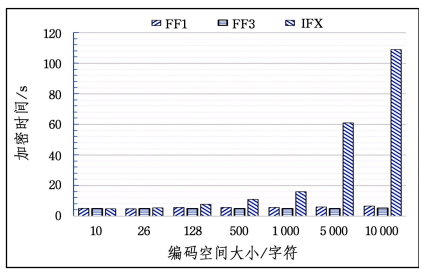

图 6 编码表大小对加密性能的影响

10 的字符空间对应 10 进制数类型，26 的字符空间对应字母类型，128 的字符表对应常用的字符集。

结论 1 $radix$ 的大小对 FF1，FF3 的影响是不显著的。

结论 2 IFX 算法在 $radix$ 小于一定值时影响是不显著的，大于该值后会使效率显著下降。

分析：注意结论 2 中极限点的取值是动态变化的，会受到字段长度等因素的影响，不过可以保守地取为 128。

3.2 字段长度对加密性能的影响

使用随机数生成 10 万条 $radix$ 为 10 的字符串，使用不同的加密算法加密并测量运行时间，得到如图 7 所示的效率图。

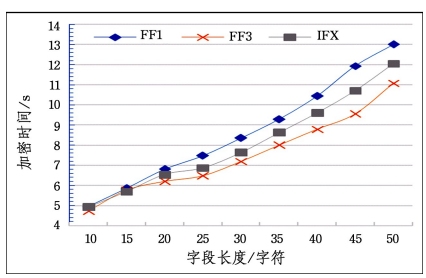

图 7 字段长度对加密性能的影响图

分析图 7 的效率曲线，可得出结论 3。

结论 3　在加密字段长度一定值时（可以保守取 50），FF1 的性能，也优于 IFX 优于 FF3。

分析：FF1，FF3，IFX 算法的效率受字段长度的影响都是显著的，这是因为这些算法的 Feistel 网络一次吸收整个加密消息，这使得这些算法不适合用于加密长文件。当 FF1，FF3，IFX 算法用于处理大长度的数据时，IFX 的性能将会变得最差。比如在这个实验的前提上，若字段长度扩展到 1000，则 FF1 加密 10 万条数据将会用大约 12 min，而 IFX 大约会花 43 min。结合上一个实验，无论长度过长或者 *radix* 过大，IFX 的性能都会大幅下降，可以得出下结论 4。

结论 4　IFX 不适合处理较大的格式空间上的加密。

注意 FF3 的长度限制 $length(X)\in[2,2\lfloor\log_{radix}(2^{96})\rfloor]$，在给定的 *radix* 下，字段的最大长度也就确定了。比如在这个实验中，$radix=10$，*FF3* 支持的最大加密长度是 58。这就是说 FF3 也无法用于处理大长度的数据。

3.3　真实民航数据的加密效率测试

本节使用相应的加密方法处理真实的民航数据，包括航班数据与乘客数据。测定加解密相应字段所需要的时间。

表 3 为使用相应算法加解密 4143 条航班信息的相应字段的用时信息，其加密后的具体效果可以参照第 4 节的表 6。

表 3　航班数据效率测试

（单位：ms）

字段名	加密算法	加密用时	解密用时
机场三字码	FF3	1638.0	1644.4
航班号	IFX	1371.8	1415.2
离港时间	FF1	1662.2	1730.4
离港日期	FF1	1799.4	1796.0

表 4 为使用相应算法加解密 2027 条乘客信息的对应字段的用时信息，其具体的加密效果可以参照第 4 节的表 8。

表 4　乘客数据效率测试

（单位：ms）

字段名	加密算法	加密用时	解密用时
姓名	IFX	735.0	754.6
年龄	FF3	756.0	769.0
身份证号	FF1	825.8	849.6
座位号	IFX	705.4	716.4

4　加密效果

通过从民航信息系统数据库中选取一些字段用于实现保留格式加密算法的效果，其中对于航班信息表，本文提取了机场、航班号、时间、机型、日期等字段，并列举了几条航班数据与旅客数据，其中航班数据如表 5 所列。

表 5　加密前的航班数据

机场	航班代码	时间	机型	日期
AAT	CZ6684	11：10：00	AT7	2017-09-05
AAT	CZ6842	21：10：00	AT7	2017-09-05
AKU	CZ6862	11：45：00	AT7	2017-09-05
AKU	CZ6864	16：45：00	AT7	2017-09-05

然后对表中数据执行 FPE 加密算法，加密后的效果如表 6 所列。

表 6　加密后的航班数据

机场	航班代码	时间	机型	日期
YKY	IA9057	01：29：44	OX7	2226-12-16
CNS	JQ3876	00：55：04	YO6	2033-09-11
GUB	KR1951	04：29：31	AB0	2152-12-30
VLR	RQ7427	12：32：03	FT7	2237-09-10

对于旅客信息表，其加密前的数据如表 7 所列。加密后”数据如表 8 所列。

表 7　加密前的旅客数据

姓名	年龄	身份证号	航班号	座位号
李毅	29	330100198302230492	105	7L
刘尔	30	330100198604182254	30	2B
王山	29	330100199206162834	100	2E
李坤	20	330100198210081821	35	34

表 8　加密后后旅客数据

姓名	年龄	身份证号	航班号	座位号
蝳銧	686	330107938517969353	900	5K
庁虓	698	3301077336365245OX	669	33
墋奮	128	330106927442277124	824	4H
汽吩	702	330107004380861755	714	2X

可以看出，加密后数据的原始格式并没有发生变化，只有具体内容的改变，例如航班号中，加密前数据的格式是 2 个英文字母与 4 个数字组合，加密后依然是 2 个英文字母与 4 个数据的组合。此外，使用航班 ID 作为扭转因子，使得在不同条目中相同的数据加密后的密文也不相同，例如在表中相同的日期信息在加密后也各不相同。

结束语　将保留格式加密技术应用于加密民航信息系统数据的优点是：符合当前民航航班信息系统的需求，对当前航班信息泄露问题的解决有重大意义，是近几年民航信息系统数据库的保密技术研究的热点，有广阔的发展前景；可保证信息的安全性、完整性、真实性和可搜索性，保留航班数据的内在格式；另外，保留格式加密技术还可以直接用于民航信息系统的数据遮蔽[13]。因为在民航信息系统的开发、测试和使用人员培训的过程中，可能需要由第三方接手真实的信息数据。为保证信息的机密性和对真实信息的使用需求，可以使用 FPE 的加密方法仿真出看似真实可靠的数据以满足开发、测试、培训的需求。当然，基于保留格式加密技术的民航信息系统目前也有亟需解决的问题：比如如何进一步提高加密性能以处理海量数据。如何在不解密的情况下做出更复杂多样的查询统计等。这些问题都是 FPE 在大数据时代下所必须面临的严峻挑战。

参 考 文 献

[1] 么轶堃. 浅谈民航电子商务及其信息安全[J]. 民航经济与技术，2000(6)：60-62.

[2] PCI Security Standards Council. Payment Card Industry Data Security Standard[S]. 2006.

[3] 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.

[4] DWORKIN M. NIST Special Publication 800-38G，Recommendation for Block Cipher Modes of Operation：Methods for Format-Preserving Encryption[OL]. https://nvlpubs. nist. gov/nistpubs/SpecialPublications/NIST. SP. 800-38G. pdf.

[5] BELLARE M，RISTENPART T，ROGAWAY P. Format-Preserving encryption[C]// Proc. of the Selected Areas in Cryptog-

raphy (SAC 2009). LNCS 5867, Calgary: Springer-Verlag, 2009:295-312.
[6] BLACK J, ROGAWAY P. Ciphers with arbitrary finite domains [C]// Proc. of the Topics in Cryptology—CT-RSA 2002. LNCS 2271, San Jose: Springer-Verlag, 2002:114-130.
[7] HOANG V T, ROGAWAY P. On generalized Feistel networks [C]// Advances in Cryptology—CRYPTO 2010. LNCS 6223, Santa Barbara: Springer-Verlag, 2010:613-630.
[8] 王鹏. 多类型数据保留格式加密技术的研究与实现[D]. 北京: 北京邮电大学, 2017.
[9] 刘华宁. 一类伪随机二进制数列的碰撞与雪崩效应[J]. 数学学报, 2013, 56(6):907-914.
[10] LIU Z L, JIA C F, LI J W, et al. Format-Preserving encryption for datetime[C]// Proc. of the Intelligent Computing and Intelligent Systems 2010. Xiamen: IEEE Press, 2010:201-205.
[11] SCHROEPPEL R. Hasty Pudding Cipher specification [OL]. http://richard.schroeppel.name:8015/hpc/hpc-spec.
[12] LISKOV M, RIVEST R, WAGNER D. Tweakable block ciphers [C]// Advances in Cryptology—CRYPTO2002, Lecture Notes in Computer Science 2442. Berlin: Springer, 2002:31-46.
[13] 姚远. 信息系统数据遮蔽实现方式[J]. 软件导刊, 2014, 13(5):156-158.

(上接第 554 页)

课堂座位分析与反馈子单元利用人脸分割与人脸检测模型对数据存储模块中的课堂图像进行识别,将学生的座位位置与学生个人的学习成绩联系在一起,根据学习成绩分为优秀、良好以及普通的学生,用 3 种不同颜色的点来表示,绘制如图 11 所示的散点图,两坐标代表位置,最后将散点图更新到学生数据以及教师反馈信息中,教师可以根据该数据全面地、有针对性地照顾到更多的同学。

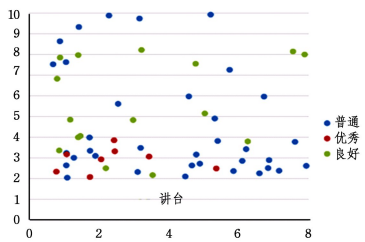

图 11 各类学生的座位分布图

该线下课堂质量双向评估系统获取信息的渠道方便快捷,而且能够减少大量的时间成本,有效提高教师的教学效率与学生的学习效率。

结束语 随着网络教学平台的兴起,在线教育逐渐走进各大高校,出现线上线下同步混合教育的模式。

文章首先对现有的智能教学系统和相关技术进行了研究,分析了现有的教学系统的不足之处,开拓线上线下相结合的设计思想,提出了一种基于深度学习的智能教学系统。基于在线学习行为分析为学生提供了个性化的学习内容推荐,基于课堂信息为学生提供课堂学习情况的分析以及反馈信息以便学生对自己的学习行为进行全面地反思和改进,促进学习效率的提高,还能为教师提供课堂教学质量的反馈信息以便了解学生对知识的掌握情况和对教学方法的评估以便改进不足的地方,从而提高教学效率。该系统对在线学习效率的提升以及课堂质量的提高都具有深刻的意义和价值。

实验显示出了目前方法的一些不足之处。在后续的研究中,将继续研究实现在长视频序列中高效地检测微表情技术,完善抽帧操作的时间间隔定义,进一步提升识别的速度和效率。

参 考 文 献

[1] COATES H. The value of student engagement for higher education quality assurance[J]. Quality in Higher Education, 2005, 11(1):25-36.
[2] ALLY M. Foundations of educational theory for online learning [J]. Theory and Practice of Online Learning, 2004, 2(5):15-44.
[3] LESTA L, YACEF K. An intelligent teaching assistant system for logic[C]// International Conference on Intelligent Tutoring Systems. Springer, Berlin, Heidelberg: IEEE Press, 2002: 421-431.
[4] FENGMEI C. Design of Intelligent Teaching Analysis System [J]. Journal of Advanced Oxidation Technologies, 2018, 21(2): 12-17.
[5] 胡峰, 赵俊博, 焦瑞莉. 基于 ZigBee 的互动教学系统学生端设计[J]. 测控技术, 2017, 36(5):152-155.
[6] 张新明, 何文涛. 支持翻转课堂的网络教学系统模型探究[J]. 现代教育技术, 2013, 23(8):21-25.
[7] 王永明, 徐继存. 论在线课程教学系统的建构[J]. 中国电化教育, 2018, 2(3):66-73.
[8] 贾积有, 张必兰, 颜泽忠, 等. 在线数学教学系统设计及其应用效果研究[J]. 中国远程教育, 2017, 1(3):37-44.
[9] HUNG J L, ZHANG K. Revealing Online Learning Behaviors and Activity Patterns and Making Predictions with Data Mining Techniques in Online Teaching[J]. Journal of Online Learning & Teaching, 2008, 4(4):426-436.
[10] RATNAPALA I P, RAGEL R G, DEEGALLA S. Students behavioural analysis in an online learning environment using data mining[C]// International Conference on Information and Automation for Sustainability. San Francisco: IEEE Press, 2015:132-139.
[11] 马国富, 王子贤, 刘太行, 等. 大数据时代下的线上线下混合教学模式研究[J]. 教育文化论坛, 2017, 9(2):22-24.
[12] 张策, 徐晓飞, 张龙, 等. 利用 MOOC 优势重塑教学实现线上线下混合式教学新模式[J]. 中国大学教学, 2018, 5(3):7-11.
[13] SADEGHI B H M. A BP-neural network predictor model for plastic injection molding process[J]. Journal of Materials Processing Technology, 2000, 103(3):411-416.
[14] KONTOYIANNIS I, ALGOET P H, SUHOV Y M, et al. Nonparametric entropy estimation for stationary processes and random fields, with applications to English text[J]. IEEE Transactions on Information Theory, 1998, 44(3):1319-1327.
[15] SARODE N, BHATIA S. Facial expression recognition[J]. International Journal on Computer Science and Engineering, 2010, 2(5):1552-1557.
[16] GIRSHICK R. Fast r-cnn[C]// Proceedings of the IEEE International Conference on Computer Vision. Boston: IEEE Press 2015:1440-1448.
[17] SCHROFF F, KALENICHENKO D, PHILBIN J. Facenet: A unified embedding for face recognition and clustering[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE Press, 2015:815-823.