软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2012,23(1):152−170 [doi: 10.3724/SP.J.1001.2012.04096]
©中国科学院软件研究所版权所有.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

保留格式加密技术研究*

刘哲理[1,2], 贾春福[1,2], 李经纬[1,2+]

[1](南开大学 信息技术科学学院 计算机与信息安全系,天津 300071)

[2](福建师范大学 网络安全与密码技术重点实验室,福建 福州 350007)

Research on the Format-Preserving Encryption Techniques

LIU Zhe-Li[1,2], JIA Chun-Fu[1,2], LI Jing-Wei[1,2+]

[1](Department of Computer and Information Security, College of Information Technical Science, Nankai University, Tianjin 300071, China)

[2](Key Laboratory of Network Security and Cryptology, Fujian Normal University, Fuzhou 350007, China)

+ Corresponding author: E-mail: lijw1987@gmail.com

Liu ZL, Jia CF, Li JW. Research on the format-preserving encryption techniques. *Journal of Software*, 2012, 23(1):152−170. http://www.jos.org.cn/1000-9825/4096.htm

Abstract: The paper reviews the current research situation of FPE (format-preserving encryption) including basic constructing methods, encryption modes and security. When describing the basic constructing methods, it introduces the basic principles of Prefix, Cycle-Walking and Generalized-Feistel and their application scopes. When explaining the encryption modes, it mainly analyzes the construction features of FPE modes or schemes, introduces the principles of three classical modes, summarizes the different types of Feistel networks and presents an overview of their applications in FPE. When talking about the security, it describes the security notions of FPE and their corresponding games, analyzing the relationship among them. In the end, it introduces the application scopes of FPE and points out that performance, integrity authentication and key problems of database encryption with FPE, such as making range query and arithmetic operation on encrypted data, are the major problems to be solved in the future. All these works will play a role in promoting research of format-preserving encryption.

Key words: format-preserving encryption; rank-then-encipher; Feistel network; security goal; order preserving encryption; privacy homomorphism

摘 要: 围绕基本构建方法、加密模型和安全性等方面,对保留格式加密(format-preserving encryption,简称 FPE)的研究现状进行了综述.在基本构建方法方面,介绍了 Prefix,Cycle-Walking 和 Generalized-Feistel 方法的工作原理及适用范围;在加密模型方面,分析了 FPE 模型或方案所呈现的构造特点,介绍了典型模型的工作原理,总结了 Feistel 网络的类型及其在 FPE 中的应用情况;在安全性方面,描述了保留格式加密的安全目标及相关的游戏模型,分析了各安全目标之间的关系.最后介绍了保留格式加密的应用领域,指出性能、完整性认证以及 FPE 在数据库加密应用中

* 基金项目: 国家自然科学基金(60973141); 天津市自然科学基金(09JCYBJ00300); 高等学校博士学科点专项科研基金(20100031110030); 网络安全与密码技术福建省高校重点实验室开放课题(2011004); 中央高校基本科研业务费专项资金

收稿时间: 2010-07-21; 修改时间: 2011-04-02; 定稿时间: 2011-07-21; jos 在线出版时间: 2011-09-09

CNKI 优先出版时间: 2011-09-08 17:19, http://www.cnki.net/kcms/detail/11.2560.TP.20110908.1719.003.html

© 中国科学院软件研究所 http://www.jos.org.cn

如何对密文进行范围查询、算术运算将是进一步需要解决的问题.这些研究工作将对保留格式加密的研究起到一定的促进作用.

关键词: 保留格式加密;排序后加密;Feistel 网络;安全目标;保序加密;秘密同态

中图法分类号: TP309 **文献标识码**: A

在实际应用中,对数据库中的信用卡号、身份证号等敏感数据进行加密非常必要,然而使用传统分组密码通常会扩展数据,使数据长度和类型发生变化,需要修改数据库结构或应用程序来适应这些变化,成本非常高.

此类加密要求密文与明文具有相同的格式,是一类新的加密问题,称为“保留格式加密(format-preserving encryption,简称 FPE)”.

FPE 的初衷是为了解决数据库或者应用系统中敏感数据的加密问题,随着研究的进展,其应用并不仅限于此.比如:FPE 可以应用于数据遮蔽(data masking)[1]领域,通过克隆原始数据进行掩码转换,输出一个与原数据格式、关联一模一样的数据,用于解决从生产环境的数据向测试环境(或者开发环境)导入时可能产生的数据内容安全问题.此外,FPE 对于网络数据安全一样有用,可以使数据报在不改变格式的情况下在传输过程中受到保护.目前,国外已公开发表多篇有关 FPE 的理论研究成果,美国 Voltage 公司已将 FPE 技术应用到其安全产品 SecureData 中.

本文关注近年来保留格式加密研究的进展,描述了保留格式加密的研究历史,并主要围绕基本方法、加密模型以及安全性等方面对已有的研究成果进行综述,指出已有加密模型的构造特点,全面介绍了与 FPE 相关的安全目标和相应的游戏模型,以期对保留格式加密在国内的研究起到一定的推动作用.

1 保留格式加密

1.1 FPE问题

保留格式加密是一种对称密码,要求密文与明文具有相同的格式.

在 1997 年,Smith 和 Brightwell[2]认为,FPE 有助于增强数据库和数据仓库的安全性,并指出使用传统分组密码解决该问题的困难:“数据库中敏感数据加密后,密文要和明文具有大致的相似性.使用 DES 算法对一个社会保险号加密后,密文不仅不像社会保险号,甚至不包含任何数字,由于社会保险号在数据库中被定义为长度为 9 的字符型字段,因此无法存储 DES 算法加密后的数据,应用程序也无法正确读取和显示,除非在应用程序和数据库中进行大量的修改来适应加密后数据格式的变化”.

对于数据库敏感数据的保留格式加密,需要保证密文满足数据库对于数据格式的约束,主要包括:(1) 数据不能被扩充.例如,当加密 N 位的数字时,必须输出另外一个 N 位的数字;(2) 数据类型不能被改变;(3) 数据必须能被确定性地加密.对于数据库中作为主键或者索引字段的数据,被加密后将保留其所在的列作为主键或者索引的特性;(4) 加、解密过程可逆.

1.2 研究历史

FPE 问题自提出以来已有大约 30 年的研究历史,可以划分为两个阶段:早期的研究学者致力于探索可行的解决方法,最近的研究学者则致力于解决复杂消息空间上的保留格式加密问题和提出高效的加密模型.其中,具有代表性的有关 FPE 的文献及其所做的贡献可见表 1.

1.2.1 基本方法探索阶段

基本方法探索阶段主要致力于提出简单问题域上 FPE 问题的解决方法.自 1981 年提出 FPE 问题以来,直到 2002 年才提出了适用于整数集上的 3 种基本的 FPE 构建方法.

FPE 问题最早可以追溯到 1981 年,文献[3]描述了一种使用 DES 算法加密字符串的方法.与分组密码算法不同,该方法要求密文与明文具有相同的格式:假设密文和明文是由确定字母表 $chars$ 中字母组成的字符串,比如 $chars$={0,1,…,9},则加、解密必须在集合 $chars^n$(由 $chars$ 中字符组成的长度为 n 的字符串构成的集合)中完

© 中国科学院软件研究所 http://www.jos.org.cn

成,即对于每个明文 $x\in chars^n$,其将被映射成密文 $y\in chars^n$.然而随着 2005 年 5 月 FIPS 74 标准[3]的停止使用,该方法也被废弃.

Table 1 History of format-preserving encryption researches

表 1 保留格式加密的研究历史

时间	论文	方法或贡献	解决的问题域
1981	Guidelines for implementing and using the NBS data encryption standard[3]	一种基于 DES 的加密方案	$chars^n$
1997	Using datatype-preserving encryption to enhance data warehouse security[2]	一种基于 DES 的 CFB 模式的加密方案	$chars^n$
2002	Ciphers with arbitrary finite domains[4]	Prefix	整数集
		Cycle-Walking	
		Generalized-Feistel	
2004	Security of random Feistel schemes with 5 or more rounds[6]	证明了 Feistel 网络的轮数与由其安全性的关系	
2008	Voltage security whitepaper: Format preserving encryption[7]	提出使用 FPE 来保护数据库中的个人识别信息	
2008	Feistel finite set encryption mode[8]	FFSEM	整数集
2009	Format-Preserving encryption[9]	Rank-then-Encipher	任意正则语言
		一种基于非平衡 Feistel 网络的整数 FPE 方案	整数集
2009	The FFX mode of operation for format-preserving encryption[10]	FFX	$chars^n$
2009	How to encipher messages on a small domain—Deterministic encryption and the thorp shuffle[17]	一种基于 Thorp Shuffle 的解决方案	小型整数集
2010	BPS: A format-preserving encryption proposal[13]	BPS	$chars^n$
2010	A new FPE scheme based on Feistel network[18]	基于 Type-2 Feistel 网络的整数 FPE 方案	整数集
2010	Format-Preserving encryption for DateTime[19]	基于随机基准值的加密方案	数据库中日期时间型字段

注:*chars* 为一个有限字符表

1997 年,文献[2]考虑了更一般的情况,描述了对数据库中特定类型字段进行加、解密的困难,试图寻找一种保留数据类型不变的加、解密方法,将其定义为保留数据类型加密(datatype-preserving encryption,简称 DPE),并指出传统的分组密码不适合解决该问题.文献[2]的作者提供了可参考的解决方法:使用 DES 算法基于 CFB (cipher-feedback)模式来产生一个偏移量向量,并与明文字符串的索引向量求和,得到密文字符串的索引向量,最后再将其转化为密文字符串.然而,该方法在固定密钥的情况下,当两个明文字符串具有相同的前缀时,所得的密文也具有相同的前缀,这会造成明文部分信息的泄露.

2002 年,文献[4]首次从密码学角度研究了 FPE 问题,认为核心是设计某密码 $E:K\times X\to X$.其中,K 是密钥空间,X 是有限的消息空间.文献[4]关注整数集 $X=\mathbb{Z}_n=\{0,1,\ldots,n-1\}$上的 FPE 问题,提出了 3 种 FPE 构建方法:Prefix, Cycle-Walking 和 Generalized-Feistel.Prefix 密码在内存中建立一个伪随机置换,利用该置换来加密数据,仅适合于较小整数集(6 位以内十进制整数).Cycle-Walking 方法能够使用传统分组密码(分组大小为 $2m$)在小于但接近于 2^{2m} 的整数集上产生一个安全置换,然而,如果待加密整数集的大小远远小于所采用密码分组的长度时,则 Cycle-Walking 会消耗更高的代价.实验证明,采用 64 位加密算法,适合加密的范围为二进制位数为 54~64 位的整数集(16~19 位十进制整数).Generalized-Feistel 方法利用 Feistel 网络[5]构造分组长度为 $2m$ 的分组密码(这里使 2^{2m} 略大于待加密整数集的大小),并结合 Cycle-Walking 方法,理论上可对任意大小的整数集进行保留格式加密.实验证明,该方法较适合二进制位数为 40~240 位的整数集(12~80 位十进制整数).文献[4]还证明,当攻击者拥有的明文密文对少于 $2^{m/2}$ 时,Generalized-Feistel 方法具有足够的安全性.2004 年,Patarin[6]简单扩展了 Black 和 Rogaway 的结论,用更多轮次的 Feistel 运算替代原来的 3 轮运算,证明了攻击者至少需要拥有 2^m 的明文密文对时才可能破译密码,使得 Feistel 网络能够在更大范围内得以应用.

© 中国科学院软件研究所 http://www.jos.org.cn

1.2.2 加密模型研究阶段

2008 年,美国 Voltage 公司公布了其安全产品中所采用的 FPE 技术的白皮书[7],介绍了已有的 FPE 研究成果,提出了社会保险号和信用卡号的 FPE 方案.自此,FPE 得到了更多关注,研究学者们基于已提出的简单问题域上的 FPE 方法,试图解决更复杂问题域上的 FPE 问题,并提出高效的加密模型.

2008 年,文献[8]基于平衡 Feistel 网络,以截断基础分组密码 AES 输出的方式构造所需的伪随机函数,提出了 FFSEM(Feistel finite set encryption mode)模型.该模型使用 Cycle-Walking 方法,理论上可以解决任意整数集上的 FPE 问题,已被 NIST(美国国家标准和技术协会)所采纳.

2009 年,Bellare 在文献[9]中对 FPE 问题进行了更深入的研究,充分考虑了消息空间的复杂性,完整地定义了 FPE 的密码学概念及安全目标,提出了一种通用的 FPE 构建方法“Rank-then-Encipher”,简称 RtE 方法.该方法主要包括两个步骤,即排序和加密,能否在待加密消息空间中找到高效的排序算法是该方法的关键.然而,并不是所有消息空间都存在高效的排序算法,Bellare 等人将此作为开放性问题保留下来,即在实际问题中寻找一个不需要使用排序算法的高效 FPE 方法.

同年,Bellare 在消息空间、Feistel 网络类型和轮运算等方面对 FFSEM 进行了扩展,提出了 FFX(format-preserving,Feistel-based,X 是参数选择等相关因素)模型[10].该模型使用非平衡 Feistel 网络[11],可以解决长度为 n 的字符串构成的消息空间 $chars^n$(*chars* 为某固定字符表)上的 FPE 问题.与 FFSEM 相比,FFX 所解决的 FPE 问题的消息空间更加复杂,加入了对调整因子(tweak)[12]的支持,支持非二进制字母表和非平衡划分,并且在加密信用卡号、社会保险号等信息时,可以避免使用 Cycle-Walking.

2010 年,文献[13]对 FFX 进行了改进,详细描述了调整因子的使用方法,提出了 BPS 模型.在该模型中,可以使用 TDES[14],AES[15]或 SHA-2[16]构建内部分组加密算法,通过 CBC 模式(cipher-block chaining mode)可以加密任意长度的字符串并保留其格式.

除了上述典型的加密模型外,在 2009 年,文献[17]针对文献[4]中所提出的方法(Prefix,Cycle-Walking 和 Generalized-Feistel)不能解决的较小整数集(比如 6~12 位十进制整数)内的 FPE 问题,提出了基于 Thorp Shuffle 的解决方案,并详细分析了其安全性.在 2010 年,文献[18]基于 k-分割的 Type-2 Feistel 网络提出了新的整数 FPE 方案.同年,文献[19]针对 RtE 方法在对日期型数据进行 FPE 时效率较低的问题,提出了基于随机基准值的日期类型的解决方案,基本思想是“随机选取一个日期作为基准值,通过对待加密的日期相对于该基准值的偏移量进行加、解密的方法,将日期类型的 FPE 问题转移为整数域上的 FPE 问题”,极大地提高了效率.

1.3 两种定义

基于 FPE 已有的研究成果,从两个角度对 FPE 进行了定义:基本 FPE 和一般化 FPE.基本 FPE 描述了 FPE 要解决的问题,即确保密文属于明文所在的消息空间;一般化 FPE 则强调 FPE 问题的复杂性在于待加密消息空间的复杂性.

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.

基本 FPE 强调明文和密文处于相同的消息空间,因此具有相同的格式.以 n 位信用卡号的保留格式加密为例,密文要求与明文一样,都是由十进制数字组成的长度为 n 的字符串,即两者均为消息空间 $\{0,\ldots,9\}^n$ 内的元素.根据基本 FPE 的定义,分组密码也是一种特殊的 FPE,它是由分组长度 n 决定的 $\{0,1\}^n$ 字符串集合上的置换.然而,FPE 要处理的消息空间远比分组密码复杂的多,比如格式为“YYYY-MM-DD”的日期型消息空间,不仅有长度为 10 的字符串长度限制,还需要满足特定位置是字符‘-’、年、月、日在合理范围内等格式要求.

为了更准确地描述 FPE 问题,定义集合 Ω 为格式空间,任意一个格式 $\omega\in\Omega$,可确定消息空间的一个与格式 ω 相关的子空间 X_ω.FPE 与集合 $\{X_\omega\}_{\omega\in\Omega}$ 有关,称 X_ω 为由格式 ω 确定的消息空间的一个分片,每个分片都是一个有限集.当给定密钥 k、格式 ω 和调整因子 t 后,FPE 就是一个定义在 X_ω 上的置换 $E_k^{\omega,t}$.

定义 2(一般化 FPE). FPE 可以描述为一个密码 $E:K\times\Omega\times T\times X\to X\cup\{\perp\}$,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 $\perp\notin X$.

© 中国科学院软件研究所 http://www.jos.org.cn

为了有效地研究分析加密模型,可通过算法三元组$\mathcal{E}_{\mathrm{FPE}}$=(Gen,Enc,Dec)来描述一般化 FPE,其中:

- 算法 Gen:初始化系统参数 *params*.不同 FPE 模型需要初始化的系统参数也有所差异,通常包含 3 部分:
  1) 初始化具有足够安全性的对称加密算法所需的参数,比如 Feistel 网络所需要的轮次数、轮函数和分组长度等;
  2) 初始化待解决的问题域,包括明确格式ω及由其确定的分片 X_ω;
  3) 初始化用于加解密的对称密钥 k,该对称密钥需要安全存储,不对外公开.
- 算法 Enc:输入为调整因子 t 和明文 x,返回在分片 X_ω内的密文 y 或者⊥.该算法执行 $E_K^{\Omega,T}(X)=E(K,\Omega,T,X)$过程, $E_K^{\Omega,T}(\cdot)$ 是 X_Ω上的一个置换.如果 $x\in X_\omega$,则返回 $y=E_k^{\omega,t}(x)$;否则,返回⊥.
- 算法 Dec:输入为调整因子 t 和密文 y,返回相同分片 X_ω内的明文 x 或者⊥.该算法是算法 Enc 的逆运算,定义如下:如果 $y\in X_\omega$,则返回 $x=D_k^{\omega,t}(y)$;否则,返回⊥.

1.4 标准安全目标

保留格式加密是一种特殊的对称密码,基础模块是分组密码和伪随机函数.由于安全性通常可以归约到基础模块的安全性上,因此,保留格式加密的一个重要的安全目标是伪随机性.

2002 年,Black 和 Rogaway[4]首次描述了保留格式加密的安全性,认为标准的安全目标就是伪随机置换(pseudorandom permutation,简称 PRP)安全.

根据基本 FPE 的定义,对任意 $k\in K$,$E(k,\cdot)=E_k(\cdot)$是消息空间 X 上由对称密钥 k 决定的一种置换.设 $Perm_k(\cdot)$表示消息空间 X 上所有置换的集合, $P\xleftarrow{\$}Perm(X)$ 表示从 $Perm_k(\cdot)$中随机抽取一个置换 P.设 A^f是一个可以查询预言机 f 的攻击者,f 要么是加密预言机 $E_k(\cdot)$,要么是一个随机置换预言机 $P(\cdot)$.假定攻击者从不执行消息空间之外的查询,而且不重复相同的查询,这样的攻击者 A 可以认为是保留格式加密方案$\mathcal{E}_{\mathrm{FPE}}$ 的 PRP 攻击者,并且定义其在攻击中可获得的优势为

$$\mathrm{Adv}^{\mathrm{PRP}}_{\mathcal{E}_{\mathrm{FPE}}}(A)\overset{def}{=}|Pr[k\xleftarrow{\$}K:A^{E_k(\cdot)}=1]-Pr[P\xleftarrow{\$}Perm_k(\cdot):A^{P(\cdot)}=1]|.$$

上式度量了攻击者 A 区分保留格式加密和随机置换的概率优势.

定义 3(PRP 安全). 令 $\mathrm{Adv}^{\mathrm{PRP}}_{\mathcal{E}_{\mathrm{FPE}}}(t,q)\triangleq\max_A \mathrm{Adv}^{\mathrm{PRP}}_{\mathcal{E}_{\mathrm{FPE}}}(A)$,其中,$t$ 为攻击者执行破解算法的时间,q 为攻击者查询的次数.如果 $\mathrm{Adv}^{\mathrm{PRP}}_{\mathcal{E}_{\mathrm{FPE}}}(\cdot,\cdot)$ 是一个可忽略的量,则称该保留格式加密方案$\mathcal{E}_{\mathrm{FPE}}$ 为伪随机置换,也即达到了 PRP 安全.

2 基本方法

2002 年,Black 和 Rogaway[4]提出了 3 种 FPE 构建方法:Prefix,Cycle-Walking 和 Generalized-Feistel.这 3 种方法不仅在一定程度上解决了整数集上的 FPE 问题,而且成为构造 FPE 模型的基本方法.其中,Cycle-Walking 和 Generalized-Feistel 在后续的加密模型中得到了广泛使用,包括 FFSEM[8],FFX[10],BPS[13]等.

2.1 Prefix

Prefix 方法很简单,首先在内存中建立一个随机的置换表,然后基于该置换表对数据进行加、解密.这意味着加、解密速度非常快,但在较大消息空间上建立置换表将会耗费更多的时间,因此只适用于较小的有限集 $X=\{0,1,\ldots,n-1\}$,$n<10^6$.

将 Prefix 方法记为密码$\mathcal{P}_{\mathrm{FPE}}$,其密钥空间为 K,消息空间为整数集 $X=\{0,1,\ldots,n-1\}$,$n<10^6$.为了建立置换表,采用基础分组密码 E,其对称密钥为 $k\in K$,计算如下元组:$I=(E_k(0),E_k(1),\ldots,E_k(n-1))$.由于 I 中每个分量 $E_k(i)$,$i\in X$ 是长度为分组长度的不同二进制符号串,可以按照数值关系对其进行排序,由此得到 $E_k(i)$对应的排序值 r_i.进一步对 I 进行操作,将分量 $E_k(i)$换成其对应的排序值并得到元组 $J=(r_0,r_1,\ldots,r_n)$,这样就建立了消息空间 X 上的一个置换表:给定任意明文 $x\in X$,返回元组 J 中相同序号的分量 r_x,就得到了对应的密文.

举例:消息空间为 $X=\{0,1,2,3,4\}$,为了建立置换表,选择基础分组密码 E 为 AES,计算 $E_k(0)=166$;$E_k(1)=6$;$E_k(2)=130$;$E_k(3)=201$;$E_k(4)=78$(这里,AES 的加密结果原本为分组长度的二进制字符串,但为方便起见,将其用十

© 中国科学院软件研究所 http://www.jos.org.cn

进制数来表示),得到元组 I=(166,6,130,201,78),将每个分量替换为其对应的排序值得到元组 J=(3,0,2,4,1).从而,假设要加密明文 0,返回元组 J 中序号为 0 的分量为其密文,即 3.

实际应用中,通常会有对密钥进行更新的要求,然而对于 Prefix 方法来说,密钥的更新意味着重新建立置换表,需要消耗较高的代价(重新加密整个消息空间并进行排序和替换).因此,有必要在特定环境里对密码应用调整因子[12],可以使其不需要密钥更新而更改加密函数.文献[7]已经提出一些构建可调整密码的方法:为分组密码 E 引入调整因子 t 来加密明文 x,可执行操作 $y=E_k((E_k(x)+t) \text{ MOD } n)$.可见,引入调整因子后的加、解密过程执行了两次加密,但是对 Prefix 方法而言,加、解密是在内存中查表的操作,因此不会影响效率.

Prefix 方法不会降低基础分组密码的安全性,即当 E 是 PRP 安全的时候,Prefix 也能达到相同的安全性.

2.2 Cycle-Walking

Cycle-Walking 方法为确保密文为消息空间内的元素提供了一种通用的解决思路,其加密的原理是利用基础分组密码(AES 或 3DES 等)对中间输出值不断地进行处理,直至其在可接受的输出范围内.

设 $Cycle_k(x)$表示使用 Cycle-Walking 方法对明文 x 加密,密钥为 k,加密过程为:要加密明文 $x\in\{0,\ldots,n-1\}$,选用分组密码 E(如 AES),设 $y=E_k(x)$,如果 $y\in\{0,\ldots,n-1\}$,则返回 y;否则,循环执行 $y=E_k(y)$,直到有{0,…,n−1}范围内的 y 产生为止.Cycle-Walking 可以将不在期望范围内的密文加密到此范围内,但是需要多次调用 E.

举例:设 X={0,1,…,10^6−1},首先确定所采用的基础分组密码,由于 $10^6<2^{64}$,选用 64 位的 DES 来处理,可以保证其输出范围始终包含 X.假设现在要加密明文 x=314159,计算得到 $c_1=E_k$(314159)=1040401(这里,E 采用 DES 算法,为方便起见,将其 E 的加密输出用十进制数表示),因为 $c_1\notin X$,迭代计算 $c_2=E_k$(1040401)=1729.因为 $c_2\in X$,所以 $Cycle_k$(314159)=1729.

Cycle-Walking 不会降低传统分组密码的安全性[4],但在效率方面,一次加密可能需要多次调用基础分组密码,当明文的二进制位数远小于分组长度时,会因为迭代次数增加而导致性能降低.因此,Cycle-Walking 方法适合大小接近分组长度的整数集.比如,如果采用 DES 算法,适合的范围是 54~64 二进制位的整数集.

2.3 Generalized-Feistel

Generalized-Feistel 方法要比 Prefix 和 Cycle-Walking 复杂,可以适用于更加广泛的加密范围.

由于 Cycle-Walking 方法对于接近分组密码大小的整数集完成保留格式加密时具有较高的性能,因此 Generalized-Feistel 方法的核心思想是基于 Feistel 网络来构建符合整数集大小的分组密码,并结合 Cycle-Walking 方法使最终密文输出在合理范围内.Generalized-Feistel 方法由两部分组成:① 由 Feistel 网络构造的分组密码 E,假设消息空间元素个数为 n,则 E 的分组长度要略大于 $\log_2 n$;② Cycle-Walking 方法,确保数据被加密到合理范围内.

Feistel 网络是目前主流的分组密码设计模式之一,基于 Feistel 网络,可以通过自定义的分组大小、密钥长度、轮次数、子密钥生成、轮函数等来构造一个分组密码.它将输入的分组分为左半部分 L 和右半部分 R,进行指定轮数的迭代运算,每次迭代执行轮函数 f 产生新的 L 和 R(记为 L'和 R')作为输出,如果没有完成指定轮数的迭代,输出结果将作为新一轮的输入参与下一轮迭代.

为了构建 Generalized-Feistel 密码$\mathcal{GF}_{n,f,r}(x)$,使用一个基本的伪随机函数 f 和 r 轮 Feistel 运算来加密整数集 X={0,1,…,n−1}内的值 x,首先定义一个基于 Feistel 网络的对称密码 $E_{n,f,r}(x)$:

1) 寻找最小的 w 使得 $2^{2w}\geqslant n$,2w 就是需要构建的分组密码的分组长度;
2) 定义 $f'(x)=trunc(f(x),w)$表示截取 $f(x)$的低 w 位数据;
3) 定义轮运算 $Round(R,L)=L$ XOR $f'(R)$;
4) 计算 $E_{n,f,r}(\cdot)$:① 寻找 R,L,使得 $x=L\times 2^w+R$;② 重复 r 次:{$T=Round(R,L),L=R,R=T$};③ 输出 $L\times 2^w+R$.

$\mathcal{GF}_{n,f,r}(x)$将使用所构建的分组密码 $E_{n,f,r}(x)$进行如下计算:① $y=E_{n,f,r}(x)$;② while ($y\geqslant n$){$y=E_{n,f,r}(y)$}.

很显然,$\mathcal{GF}_{n,f,r}(x)$使用了 Cycle-Walking 方法,确保数据加密到合理的范围内.由于其构建的分组密码的分组长度与待加密消息空间大小的二进制位数接近,因此具备较好的性能.

© 中国科学院软件研究所 http://www.jos.org.cn

安全性方面,Black 和 Rogaway[4]证明了,当攻击者拥有明文密文对少于 $2^{w/2}$ 时,$\mathcal{GF}_{n,f,r}(x)$是足够安全的.

2.4 结 论

Prefix,Cycle-Walking 和 Generalized-Feistel 方法能够解决特定范围整数集上的 FPE 问题,见表 2.

Table 2 Applicable scopes of basic methods

表 2 基本方法的适用范围

方法	处理的整数集大小(二进制位数)	处理的整数集大小(十进制整数)
Prefix	1~20	1~6
Cycle-Walking	50~63	16~19
Generalized-Feistel	40~240	12~80

表 2 描述了 3 种基本方法的适用范围,对于 6~12 位整数集的 FPE 问题,2009 年,Morris 和 Rogaway[17]提出了基于 Thorp Shuffle 的解决方案.

3 种基本方法为保留格式加密方案的构造提供了基本的思路:

1) Prefix 方法揭示了保留格式加密的本质,即消息空间内的置换.Prefix 方法采用建立预置换表的做法,为复杂问题域上的 FPE 问题提供了一种思路,即通过某种方式(比如排序)来建立消息空间内的置换,RtE 方法排序后加密的思想[9]与此异曲同工;
2) Cycle-Walking 方法是一种解决任意有限问题域的 FPE 问题的通用办法,目前所提出的 FPE 模型在解决任意有限域上的 FPE 问题时,通常基于该方法来确保数据被加密到正确的范围内;
3) Generalized-Feistel 方法允许用户通过定义 Feistel 网络的相关参数来构建合适分组长度的对称密码,成为保留格式加密模型普遍适用的构造方法,目前所提出的 FFSEM[8],FFX[10],BPS[13]等模型都基于其完成模型的构造.

3 加密模型研究

首先,从构造角度对已有 FPE 模型进行了分析,总结出 FPE 模型的构造特点;然后,介绍了 3 种典型的保留格式加密模型,分析了其适用的问题域及存在的主要问题;最后,针对 Feistel 网络在 FPE 构造中的广泛应用,描述了 Feistel 网络的类型及其在 FPE 模型中的应用情况,并分析了其特点.

3.1 构造特点

Black 和 Rogaway 在 2002 年提出的 3 种构建 FPE 方案的基本方法[4]成为后来 FPE 模型设计的主要参照,已提出的主要模型或方案及其采用的构建方法见表 3.

Table 3 Modes or schemes of format-preserving encryption

表 3 保留格式加密模型或方案

时间	FPE 模型或方案	所解决的问题域	FPE 基本方法			
			Prefix	Cycle-Walking	Generalized-Feistel	其他
2008	社会保险号的 FPE 方案[7]	社会保险号	√		√	
2008	FFSEM[8]	整数集		√	√	
2009	RtE[9]	任意正则语言				√
	基于 Feistel 网络的整数 FPE[9]	整数集			√	
2009	FFX[10]	$chars^n$			√	
2009	基于 Thorp Shuffle 的方案[17]	小型整数集			√	
2010	BPS[13]	$chars^n$			√	
	基于 Type-2 Feistel 网络的整数 FPE[18]	整数集			√	

注:*chars* 为一有限字符表

由表 3 可以看出:

1) Generalized-Feistel 方法得到了更多的关注.绝大多数 FPE 方案都用到了 Feistel 网络(RtE 方法强调先

© 中国科学院软件研究所 http://www.jos.org.cn

排序后加密,其加密过程使用了整数 FPE 方案,而目前已提出的整数 FPE 方案,包括文献[9]中所提出的两种整数 FPE 方案,都是基于 Feistel 网络来实现),这与 Feistel 网络相对完备的研究成果和 FPE 算法本身对保留格式的要求相关;

2) 复杂问题域上的 FPE 问题通常需要结合 Cycle-Walking 方法.FPE 问题的复杂性除了表现在保留格式上以外,也表现在待解决消息空间的复杂性上,这种复杂性使得很难发现对不同消息空间的通用解决办法.消息空间越复杂,FPE 问题也就越难解决.Cycle-Walking 方法为确保密文输出,在消息空间内提供了一种通用的解决思路,通常在解决复杂问题域上的 FPE 问题时,需要使用 Cycle-Walking 方法;

3) FPE 模型的设计思想逐渐向降低问题域复杂性的方向发展.在解决愈加复杂的消息空间上的 FPE 问题时,无论是 RtE 模型还是 FFX 模型,都不在复杂问题域上直接寻求 FPE 问题的解决办法,而是将 FPE 问题转化到等价的具有较低复杂度的整数集上.这种通过降低问题域的复杂性来解决复杂消息空间上的 FPE 问题的解决方法,将会成为保留格式加密领域的一种可采纳的有效研究手段.

3.2 主要模型

FFSEM,RtE 和 FFX 模型是 3 类具有代表意义的 FPE 模型,研究其构造特点和优缺点,对于研究分析 FPE 加密模型具有重要意义.

3.2.1 FFSEM 模型

FFSEM 是基于 Generalized-Feistel 方法的整数集上的典型 FPE 方案,它由两个基本部分组成:① 平衡 Feistel 网络,用来产生指定分组长度的分组密码;② Cycle-Walking,用 $2m$ 位的分组密码对大小为 $n(n<2^{2m})$的集合进行加解密的普遍方法.

- 算法 Gen:FFESM 初始化阶段主要定义:① FFSEM-PRF,即平衡 Feistel 网络中所使用的伪随机函数,FFSEM 使用截断基础分组密码 AES 输出的方式构造了 FFSEM-PRF;② 基础分组密码的密钥 k、消息空间的大小 n 和轮次数 r 等.详细的参数信息参阅文献[8];
- 算法 Enc:输入为明文 x,输出为满足格式要求的密文 y.

首先,将明文 x 编码为 $l=\lfloor \log_2 n \rfloor+1$ 位的二进制数(这里,$\lfloor x \rfloor$表示不超过 x 的最大整数),不足的二进制位用 0 来填充;然后执行 Cycle-Walking 过程,每次 Cycle-Walking 都将执行 r 轮 Feistel 轮运算,直到产生合适的密文;最后,对密文进行二进制解码得到对应数值的整数,其加密过程如图 1 所示.

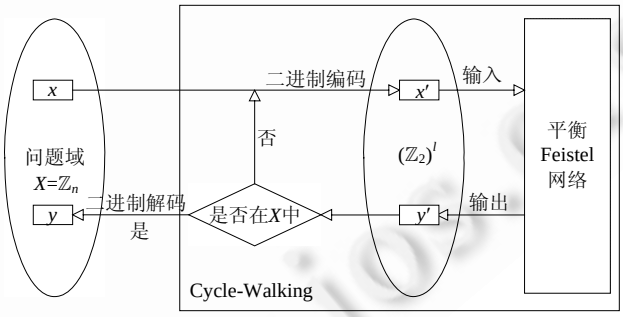

Fig.1 Feistel finite set encryption mode

图 1 FFSEM 模型

目前,该加密模型已被 Voltage 公司广泛应用,而且被 NIST 所采纳.然而,FFSEM 仅解决了整数集上的 FPE 问题,并不能成为一种普遍适用的 FPE 模型;而且 Cycle-Walking 过程需要多次调用基础分组密码,存在不确定的性能问题.

3.2.2 RtE 方法

RtE 是一种较通用的方法,其基本思想是,对消息空间内的元素先排序后加密,将解决复杂有限域上的 FPE

© 中国科学院软件研究所 http://www.jos.org.cn

问题转化到建立索引与元素的对应关系和设计整数 FPE 算法上.

为了实现对元素的排序,需要有效地描述消息空间.在编码理论研究领域,自 1977 年文献[20]发表以来,已经开始了对特定消息空间上高效 rank 算法的研究.其中,文献[9]针对可用正则语言描述的消息空间,介绍了高效的 rank 和 unrank 算法.

RtE 方法的工作原理如图 2 所示.

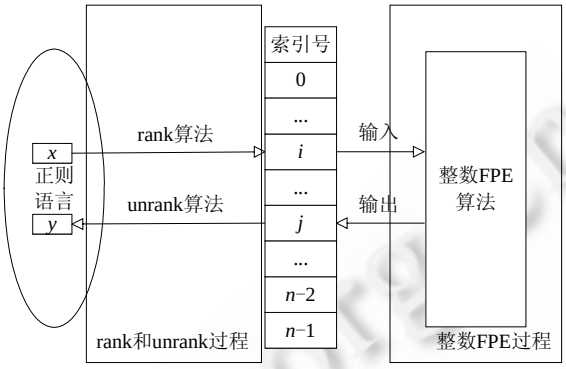

Fig.2 Rank-then-Encipher approach

图 2 RtE 方法

算法 Gen:RtE 方法的初始化阶段,主要确定:

1) rank 和 unrank 函数.对于消息空间 X,其上的排序函数可以描述为映射 $rank:X\to\mathbb{Z}_{|X|}\cup\{\perp\}$(对于有多个分片的消息空间,通常对每个分片分别进行 rank、整数 FPE 和 unrank,这里假设待加密消息空间只有一个分片),通过 rank 算法可以找到 $x\in X$ 在 X 中的索引 $i\in\mathbb{Z}_{|X|}$,如果 $x\notin X$,则 $rank(x)=\perp$.与此对应,反排序函数将一个整数映射为其在 X 中对应的元素,描述为 $unrank:\mathbb{N}\to X\cup\{\perp\}$.通过 unrank 算法,可以根据索引 $i\in\mathbb{Z}_{|X|}$找到其在 X 内的对应元素 $x\in X$,如果 $i\notin\mathbb{Z}_{|X|}$,则 $unrank(i)=\perp$.
2) 整数 FPE 方案 $E_{\mathbb{Z}_n}$ 及其所需的参数,文献[9]基于非平衡 Feistel 网络提出了两种整数 FPE 方案,而 RtE 方法并不限制所采用的整数 FPE 方案,比如 FFSEM[8]等.

算法 Enc:输入为 $E_{\mathbb{Z}_n}$ 方案所需的密钥 k、明文 x、调整因子 t 等,输出为满足格式要求的密文 y.

加密过程首先执行 rank 算法,获得明文 x 在消息空间 X 中的索引 i;然后,利用整数 FPE 方案 $E_{\mathbb{Z}_n}$ 对 i 加密得到密文索引 j;最后执行 unrank 算法,返回 X 中索引为 j 的元素 y 而得到密文.

RtE 方法的安全性等同于整数 FPE 方案的安全性,只要 $E_{\mathbb{Z}_n}$ 安全,它就是安全的.

RtE 方法的效率主要取决于消息空间内的排序算法,其时间复杂度约为 rank 算法和 unrank 算法时间的总和.Bellare 针对用正则语言描述的消息空间,描述了高效的排序算法[9].除此之外,对于无二义的上下文无关文法产生的语言,Bellare[9]指出,使用 Mäkinen 提出的 rank 算法[21]可以对其进行排序.其他各种组合结构也存在高效的 rank 算法,例如,如果想在域 $X_{n!}$(该问题域由 n 个元素的所有置换集合组成)上加密,可以使用 Lucas-Lehmer 编码提供的高效 rank 算法[22].其他的例子还有伸展树、B 树[23]、DYCK 语言[24]等.然而,并不是所有消息空间都存在基于排序的高效 FPE 算法,文献[9]将在实际问题中寻找一个不需要使用排序算法的高效 FPE 作为有趣的开放性问题保留下来.

3.2.3 FFX 模型

Bellare 在消息空间、Feistel 网络和运算等方面对 FFSEM[8]进行了扩展,提出了扩展机制 FFX 模型[10].该模型使用了非平衡 Feistel 网络,通过自定义运算可以解决 n 位字符串所构成的消息空间 $chars^n$ 的 FPE 问题.

FFX 模型的工作原理如图 3 所示.

© 中国科学院软件研究所 http://www.jos.org.cn

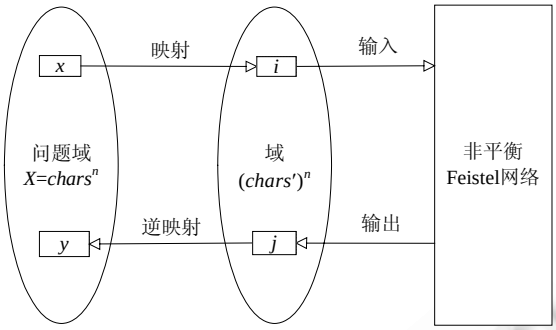

Fig.3 FFX mode

图 3 FFX 模型

算法 Gen:FFX 模型的初始化阶段,主要定义:

1) 字母表 $chars=\{char_0,char_1,char_2,char_3,\ldots,char_{radix-1}\}$及其基数 $radix=|chars|$;
2) 所使用的非平衡 Feistel 网络类型;
3) 消息空间中元素的长度 n;
4) 每轮中所用到的伪随机函数 f、所采用的运算类型、轮次数 r 和调整因子 t 等.

详细的参数信息参阅文献[10].

算法 Enc:输入为基础分组密码的密钥 k、调整因子 t 和字符串 x,输出为满足格式要求的字符串 y.字符串 x 和 y 都是由字母表 $chars=\{char_0,char_1,char_2,char_3,\ldots,char_{radix-1}\}$中字符组成的长度为 n 的字符串.

首先,将字符串 x 中的字符编码替换为数字:建立字母表 $chars=\{char_0,char_1,char_2,char_3,\ldots,char_{radix-1}\}$与 $chars'=\{0,1,2,3,\ldots,radix-1\}$的一一映射,将每个字符 $char_i$ 编码为对应的第 i 个数字.需要注意的是,$chars'$中每个数字前面的 0 与其他字符一样计入长度,例如,$chars=\{a,b,c,\ldots,z\}$,$x=acz$,将 x 编码后得到 $x=010326$.

然后,执行 r 轮指定非平衡 Feistel 网络的运算:首先,将输入(字符串 x 的编码)分割为左右两部分 L 和 R,$|L|\neq|R|$;然后,执行伪随机函数 f,对 L 和 $f_k(R)$执行选择的类型的运算得到 L':① $c_i=(a_i+b_i)$ MOD $radix$,当 $radix=2$ 时,该运算就是异或运算;② $\sum c_i radix^{n-i}=\left(\sum a_i radix^{n-i}+\sum b_i radix^{n-i}\right)$ MOD $radix^n$;最后,连接 L'与 R 得到输出 $L'||R$,并将其作为下一轮非平衡 Feistel 网络运算的输入.

可见,FFX 模型通过将非数字字母表与数字集合建立双射,将每个字符映射为对应的数字参与加密运算,实现对消息空间 $chars^n$ 的保留格式加密.与 FFSEM 相比,FFX 适用的范围更广,而且在处理信用卡号、社会保险号等 FPE 问题时避免了 Cycle-Walking,具有较高的效率.

3.3 Feistel网络及其应用

自 2002 年 Generalized-Feistel 方法首次使用 Feistel 网络来构造 FPE 算法以来,其后的 FPE 方案大多都采用了这种结构,这与 Feistel 网络相对完备的研究成果和 FPE 算法本身对保留格式的要求相关.

Feistel 网络是目前主流的分组密码设计模式之一,基于 Feistel 网络,可以通过定义分组大小、密钥长度、轮次数、子密钥生成、轮函数等来构造一个分组密码.在 FPE 模型的构造过程中,Feistel 网络被广泛应用于构造合适分组长度的分组密码,其中,FFSEM 使用了平衡的 Feistel 网络,FFX 模型中使用了两种非平衡 Feistel 网络.本节总结了常见类型的 Feistel 网络及其在不同 FPE 模型中的应用.

3.3.1 Feistel 网络的类型

常见的 Feistel 网络类型主要包括传统的 Feistel 网络、非平衡 Feistel 网络、交互式 Feistel 网络及以上 3 种的数值形式.

传统的 Feistel 网络[5],即平衡 Feistel 网络,每一轮运算的基本原理如图 4 所示:输入为长度为 $2n$ 的字符串,

© 中国科学院软件研究所 http://www.jos.org.cn

首先将其等分为长度相等的两部分 L 和 R,这里,$|L|=|R|=n$,然后对 R 执行轮函数 $f_k(R)$后并与 L 异或得到 $L'=f_k(R)$ XOR L,最后将新的字符串 $R||L'$作为下一轮迭代的输入.文献[25]证明了迭代$(6/\varepsilon-1)$轮次的传统 Feistel 网络在选择密文攻击模型下能够抵御 $2^{n(1-\varepsilon)}$次恶意查询.

Schneier 和 Kelsey 对传统 Feistel 网络进行改进,引入扩张或收缩轮的伪随机函数,提出了非平衡 Feistel 网络[11].非平衡 Feistel 网络能够支持对输入字符串不平衡的划分,即$|L|\neq|R|$.由于 L 与 R 不等长,这就需要对轮函数 f 进行调整,使得在每一轮迭代中$|f_k(R)|=|L|$恒成立.文献[25]证明,进行足够多轮次的迭代运算之后,该类型 Feistel 网络能够抵御 2^n 次选择密文查询.

Anderson 和 Lucks 分别在文献[26]和文献[27]中提出了交互式 Feistel 网络.交互式 Feistel 网络也支持对输入字符串的不平衡划分,其关键在于能够根据当前迭代轮次的奇偶性质来判定使用哪一种伪随机函数的类型(压缩还是扩张)及作用对象(L 或 R):① 在奇数轮次,伪随机函数 f 作用于 R,并将作用结果与 L 进行异或操作,得到下一轮迭代的左部分 L',下一轮迭代的右部分直接由 R 产生;② 在偶数轮次,伪随机函数 g 作用于 L,并将作用结果与 R 进行异或操作,得到下一轮迭代的右部分 R',下一轮迭代的左部分直接由 L 产生.

平衡、非平衡和交互式 Feistel 网络每一轮运算的过程如图 4 所示.

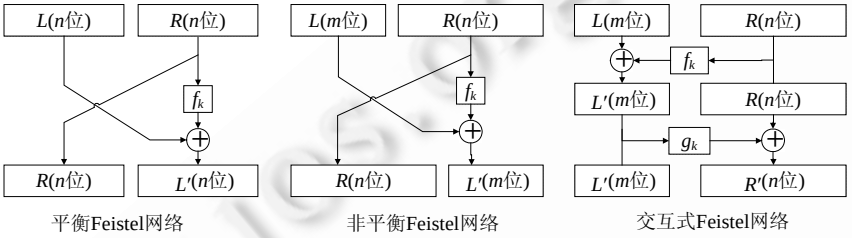

Fig 4 Balanced, unbalanced and alternating Feistel networks

图 4 平衡、非平衡和交互式 Feistel 网络

数值的 Feistel 网络是对以上各类 Feistel 网络的输入进行优化,使其支持$\mathbb{Z}_n$ 范围内的加密.假设输入为整数 $x\in\mathbb{Z}_n$,一个可行的划分办法就是选定整数 a,b,使得 $n=ab$,取 $L=\lfloor x/a\rfloor$(其中,$\lfloor\cdot\rfloor$为向下取整运算),$R=x$ MOD a,则有 $L\in\mathbb{Z}_a,R\in\mathbb{Z}_b$.于是,再根据 Feistel 网络的具体类型,进行相应的迭代运算.

在 Feistel 网络中,每一轮迭代的输出都保持了长度(对输入为字符串的 Feistel 网络)或范围(对输入为整数的 Feistel 网络)不变,这一特点使其非常适合于处理 FPE 问题.不同类型的 Feistel 网络在 FPE 方案中的应用现状可见表 4.

Table 4 FPE modes and their Feistel networks

表 4 FPE 模型及其采用的 Feistel 网络

FPE 方案	Feistel 网络
Generalized-Feistel[4]	交互式 Feistel 网络的数值形式
FFSEM[8]	传统 Feistel 网络(平衡 Feistel 网络)
FFX[10]	非平衡 Feistel 网络,交互式 Feistel 网络
基于 Feistel 网络的整数 FPE[9]	非平衡 Feistel 网络的数值形式,交互式 Feistel 网络的数值形式
BPS[13]	交互式 Feistel 网络

通过表 4 可以看出,目前,典型的 FPE 方案所采用的 Feistel 网络类型都将输入划分为等长或不等长的 L 和 R 两部分,即所采用的都是 2-分割的 Feistel 网络类型.最近,文献[18]基于 l-分割的 Type-2 Feistel 网络类型提出了一种新的整数 FPE 方案.

Type-1,Type-2 和 Type-3 Feistel 网络是 Feistel 网络的 3 种扩展形式[28],将输入长度为 lm 的比特串等分为 l 个分组 $X_1,X_2,\ldots,X_l$,然后分别进行各自的轮运算(图 5 描述了 l=4 时的工作原理):

1) Type-1 Feistel 网络执行 $X_2'=f_k(X_1)\oplus X_2$,并连接 $X_2'\|X_3\|X_4\|X_1$ 作为下一轮迭代的输入;

© 中国科学院软件研究所 http://www.jos.org.cn

2) Type-2 Feistel 网络执行 $X_2' = f_k(X_1) \oplus X_2$ 和 $X_4' = g_k(X_3) \oplus X_4$,并连接 $X_2' \| X_3 \| X_4' \| X_1$ 作为下一轮迭代的输入;

3) Type-3 Feistel 网络执行 $X_2' = f_k(X_1) \oplus X_2, X_3' = g_k(X_2) \oplus X_3$ 和 $X_4' = h_k(X_3) \oplus X_4$,并连接 $X_2' \| X_3' \| X_4' \| X_1$ 作为下一轮迭代的输入.

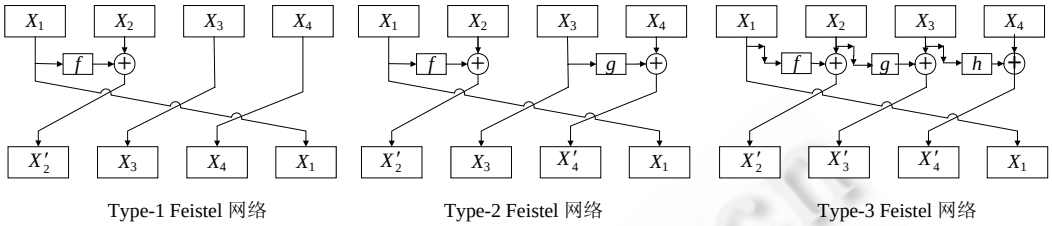

Fig.5 Type-1, Type-2, Type-3 Feistel networks

图 5 Type-1,Type-2,Type-3 Feistel 网络

3.3.2 基于 Feistel 网络的 FPE 模型和方案的特点

基于 Feistel 网络的 FPE 模型和方案,包括 FFSEM[8],FFX[10],BPS[13]等,普遍具有如下特点:

1) 通常与 Cycle-Walking 方法结合使用.在表 4 中,只有 FFSEM[8]在内部构造中使用了 Cycle-Walking 方法.事实上,基于 Feistel 网络的加密模型在解决任意有限域上的 FPE 问题时,与 Cycle-Walking 方法相结合是一种通用的做法;

2) 具有可证明的安全性.Luby 和 Rackoff 证明了,当分组长度为 $2m$ 且攻击者拥有明文密文对少于 $2^{m/2}$ 时,Feistel 网络是安全的[5].如果用更多轮次运算替代 Generalized-Feistel 方法中的 3 轮运算,只有当攻击者至少拥有 2^m 的明文密文对时才可能破译密码[6],使得 Feistel 结构能够在更大范围内得以应用;

3) 现有 FPE 方案的 Feistel 网络中所用到的伪随机函数的构造方法主要有两类:① 采用直接截断基础分组密码输出的办法,构造伪随机函数,比如 FFSEM 中 FFSEM-PRF 的构造方法;② 以基础分组密码为核心算法,采用 CBC-MAC,HMAC 等消息认证码方案,将产生的消息认证码作为 PRF 的输出,比如 FFX 中伪随机函数的构造方法[10].这两类构造方法都具有可证明的安全性.

4 安全性研究

在设计密码学方案时,主要考虑在可能面临的攻击模型下所要达到的安全目标,通常使用安全目标与攻击模型相结合的方式来定义密码学方案的安全性.

对于公钥密码体制而言,安全目标主要有:语义安全(semantic security)[29]、不可区分性(indistinguishability,简称 IND)[29,30]、不可展性(non-malleability,简称 NM)[31]和明文可意识性(plaintext awareness,简称 PA)[32].由于明文可意识性是在随机预言机模型(random oracle model)[33]下定义的,而随机预言机模型又是一种理想化的模型,因此,对于实际系统中安全性的讨论较少提到明文可意识性.Goldwasser 和 Micali[29,30,34,35]证明了,安全目标中的语义安全等价于不可区分性安全.

公钥密码体制的安全目标通常也适用于对称密码体制.除此之外,因为密码学方案一般建立在底层基础模块上,所以方案的安全性通常也可以规约到基础模块的安全性.由于对称密码体制的基础模块是分组密码和伪随机函数,因此,对称密码体制的另一个重要的安全目标是伪随机性.

保留格式加密是一种对称密码,因此,对称密码体制的安全目标对其通常都适用.在 2009 年,Bellare[9]对 FPE 问题进行了深入研究,结合保留格式加密的特性,定义了相关的多种安全目标.

4.1 安全目标

对于保留格式加密而言,标准的安全目标是 PRP 安全.也就是说,FPE 本质是在特定消息空间内的伪随机置

© 中国科学院软件研究所 http://www.jos.org.cn

换.较低的安全目标包括 SPI(single point indistinguishability),MP(message privacy)和 MR(message recovery),这些安全目标更容易实现,而且可以满足典型应用的安全性要求.

4.1.1 定 义

PRP 安全:PRP 的标准安全目标在第 1.4 节已经给出定义,要求攻击者不能区分是保留格式加密方案还是消息空间内置换集合中的某个随机置换.Bellare 进一步在文献[9]中,基于游戏模型定义了适用于 FPE 一般定义的 PRP 安全目标.

PRP 安全的定义基于游戏 $\mathrm{PRP}_{\mathcal{E}_{\mathrm{FPE}}}$,该游戏通过攻击者与挑战者之间的博弈,量化攻击者区分保留格式加密或随机置换的能力.

首先从密钥空间随机选取一个密钥 $k \xleftarrow{\$} K$,从{0,1}中随机选择一个值 $b \xleftarrow{\$} \{0,1\}$,从消息空间 X 内的置换集合中随机选取一个置换 $P \xleftarrow{\$} Perm(X)$.

$\mathrm{PRP}_{\mathcal{E}_{\mathrm{FPE}}}$ 拥有加密预言机 Oracle(Enc),用于响应攻击者 A 的加密查询,如果 b=1,加密预言机 Oracle(Enc)以保留格式加密方案 E 响应;如果 b=0,加密预言机 Oracle(Enc)以随机选取的方式置换 P 响应.

攻击者 A 的目标是:在查询加密预言机 Oracle(Enc)一定次数后,给出一个判定——Oracle(Enc)使用的是 E 还是 P,即输出一个判定值 b'.如果 $b'=b$,说明攻击者 A 判定成功,相反,则失败.A 在该游戏中具有的优势为

$$\mathrm{Adv}^{\mathrm{PRP}}_{\mathcal{E}_{\mathrm{FPE}}}(A) = 2 \cdot Pr[\mathrm{PRP}^{A}_{\mathcal{E}_{\mathrm{FPE}}} \Rightarrow \mathrm{true}] - 1.$$

SPI 安全:SPI 要求攻击者即使能够访问真正的加密预言,也不能区分是对选定的消息还是对某个随机点的加密.

SPI 安全的定义基于游戏 $\mathrm{SPI}_{\mathcal{E}_{\mathrm{FPE}}}$.类似于 $\mathrm{PRP}_{\mathcal{E}_{\mathrm{FPE}}}$,$\mathrm{SPI}_{\mathcal{E}_{\mathrm{FPE}}}$ 首先随机选取密钥 k 和待猜测的布尔值 b.攻击者 A 首先对某个待猜测的明文单点执行一次 Oracle(Test)查询,该预言机根据 b 值返回 FPE 密码下该单点的对应密文或消息空间中的某一个随机点.

攻击者 A 的目标是:在对真正的加密预言机 Oracle(Enc)(该加密预言机只以 FPE 密码 E 响应)进行一定次数适应性的查询后(在此查询过程中,禁止重复查询,禁止对待猜测单点进行查询),给出一个判定——首次执行的 Oracle(Test)查询得到的结果是选定单点所对应的密文还是待加密消息空间中的某个随机点.即输出一个判定值 b'.如果 $b'=b$,说明攻击者 A 判定成功,相反,则失败.A 在该游戏中具有的优势为

$$\mathrm{Adv}^{\mathrm{SPI}}_{\mathcal{E}_{\mathrm{FPE}}}(A) = 2 \cdot Pr[\mathrm{SPI}^{A}_{\mathcal{E}_{\mathrm{FPE}}} \Rightarrow \mathrm{true}] - 1.$$

MR 安全:在对抗消息恢复攻击方面,MR 要求即使提供给攻击者加密预言机和明文的概率分布以及格式、调整因子等信息,攻击者也不能从密文得到明文.

通常情况下,如果加密是完全随机的,就要求已知的目标密文 y^*和能够访问加密预言机 Oracle(Enc)对于恢复明文是无用的.但是这对确定性加密的要求太高,因为攻击者可以通过 Oracle(Enc)查询消息 $x_1,\ldots,x_q$,得到其对应密文 $y_1,\ldots,y_q$.如果存在某个 $y_i=y^*$,那么攻击者就可以知道待猜测的目标明文 $x^*=x_i$.MR 安全就是规范此类攻击为针对密码方案最好的攻击方式(这里,最好以成功概率来衡量).即对于 FPE 方案 $\mathcal{E}_{\mathrm{FPE}}$,如果不存在成功概率明显大于上述攻击的攻击方式,则认为 $\mathcal{E}_{\mathrm{FPE}}$ 达到了 MR 安全.

MR 安全的定义基于游戏 $\mathrm{MR}_{\mathcal{E}_{\mathrm{FPE}}}$.$\mathrm{MR}_{\mathcal{E}_{\mathrm{FPE}}}$ 提供了 3 个供攻击者查询的预言机:Oracle(Enc),Oracle(Eq)和 Oracle(Test).其中,Oracle(Enc)用于以 FPE 密码的加密方案 E 响应攻击者的加密查询,Oracle(Eq)允许攻击者查询一个消息空间内的元素是否是目标密文所对应的原始明文 x^*,Oracle(Test)用于赋予攻击者获取目标密文 y^*的能力.

在游戏开始前,从攻击范围中随机选取一个目标明文 x^*,然后执行保留格式加密算法得到其对应密文 y^*,最后将加密所用到的格式信息和调整因子(ω,t)发布给攻击者.

$\mathrm{MR}_{\mathcal{E}_{\mathrm{FPE}}}$ 中定义了两类攻击者:攻击者 A 执行一次 Oracle(Test)查询和 q 次 Oracle(Enc)查询来获得有利的信息以便做出判定;攻击者 S 放弃 Oracle(Test)查询,而使用 q 次的 Oracle(Eq)查询来替换 Oracle(Enc)查询.无论是 A 还是 S,其目标都是通过多次查询各自预言机后,输出一个对目标明文的猜测值 x,如果 $x=x^*$,说明攻击者判定成

© 中国科学院软件研究所 http://www.jos.org.cn

功,否则失败.事实上,MR 安全目标度量了攻击者在能够获取目标密文、并能 q 次访问加密预言机的情况下,与最一般的攻击方式——从待加密消息空间中选取 q 个候选值依次与目标密文判定相比,成功的概率差.即,攻击者 A 在该游戏 $\mathrm{MR}_{\varepsilon_{\mathrm{FPE}}}$ 中具有的优势为 $\mathrm{Adv}^{\mathrm{MR}}_{\varepsilon_{\mathrm{FPE}}}(A)=Pr[\mathrm{MR}^{A}_{\varepsilon_{\mathrm{FPE}}}\Rightarrow \mathrm{true}]-pa$, 其中, $pa=\max_S Pr[\mathrm{MR}^{S}_{\varepsilon_{\mathrm{FPE}}}\Rightarrow \mathrm{true}]$.

MP 安全:MP 试图量化拥有加密预言机的攻击者,从目标密文中计算其对应的目标明文被某函数作用后所得到的结果的能力.

MP 安全的定义基于游戏 $\mathrm{MP}_{\varepsilon_{\mathrm{FPE}}}$.类似于 $\mathrm{MR}_{\varepsilon_{\mathrm{FPE}}}$,$\mathrm{MP}_{\varepsilon_{\mathrm{FPE}}}$ 也提供了 3 个预言机:Oracle(Enc),Oracle(Eq)和 Oracle(Test),以及两类攻击者:A 和 S,其中,A 可以执行 1 次 Oracle(Test)和 q 次 Oracle(Enc),S 可以执行 q 次 Oracle(Eq).不同之处在于,在 $\mathrm{MP}_{\varepsilon_{\mathrm{FPE}}}$ 中,攻击者不仅需要猜测明文信息,还需要计算出该明文在某个函数 f 作用后的输出值,并将该输出值与目标明文在 f 下的作用结果进行比较,判断攻击是否成功.

事实上,MP 可以看作是 MR 的一般形式:MR 仅仅是 MP 中 f 为恒等函数的情况.因此,攻击者 A 在游戏 $\mathrm{MR}_{\varepsilon_{\mathrm{FPE}}}$ 中具有的优势为 $\mathrm{Adv}^{\mathrm{MP}}_{\varepsilon_{\mathrm{FPE}}}(A)=Pr[\mathrm{MP}^{A}_{\varepsilon_{\mathrm{FPE}}}\Rightarrow \mathrm{true}]-pa$, 其中, $pa=\max_S Pr[\mathrm{MP}^{S}_{\varepsilon_{\mathrm{FPE}}}\Rightarrow \mathrm{true}]$.

4.1.2 关 系

4 个安全目标之间的关系如图 6 所示.

Fig.6 Relation of FPE security goals

图 6 FPE 安全目标的关系

图 6 中,实箭头表示密切关联(就是能推导出),虚箭头表示有损.通过图 6 可以看出,PRP 是保留格式加密的标准安全目标,SPI 是 PRP 的一个变种,与其具有类似的安全性,虽然安全性有损,但可以达到比 PRP 更好的安全边界.为了适应更早的确定性加密概念,MP 把语义安全带给 FPE,它要求密文不能泄露真实信息,安全性要比 PRP 和 SPI 安全性略低.最基本也是最常用的安全性要求就是对抗适应性或者非适应性攻击下的 MR,MR 从整体上使对手无力从消息的密文得到信息.

SPI 概念来源于文献[36],PRP 安全就意味着 SPI 安全,但在优势边界上有额外 q/M 的损失,q 是攻击者执行查询的次数,$M=\max_{\omega\in\Omega}|X_\omega|$是消息空间最大分片的元素数目,但并不影响 SPI 作为一种安全工具来使用.文献[36,37]中的混合参数表明,SPI 安全同样也意味着 PRP 安全.

Bellare[9]证明了 SPI 可以推导出 MP 安全,但是 MP 安全不能推导出 SPI 安全.而且,Bellare[9]证明了 MP 可以推导出 MR 安全,而相反不成立.事实上,MP 的目标与 MR 的目标类似,不同之处就是,在后者中,攻击者 A 从一开始就确定其所选用的函数为恒等函数,因此,MR 安全是 MP 安全的一种特殊形式.

4.2 攻击模型

攻击者根据利用的条件可以采用不同的攻击模型,常见的攻击模型包括唯密文攻击(ciphertext only attack)、选择明文攻击(chosen plaintext attack)和选择密文攻击(chosen ciphertext attack)等.进一步地,根据询问预言机方式的不同,可分为非适应性攻击模型和适应性攻击模型.在非适应性攻击模型中,攻击者在开始询问之前就选好询问的全部内容,询问过程中内容不再发生变化;而在适应性攻击模型下,攻击者可以随时根据每次询问的结果来调整询问的内容.显然,适应性攻击强度比非适应性攻击要高.

在设计加密方案时主要关注的攻击模型有:非适应性选择明文攻击(non-adaptive chosen plaintext attack,简称 CPA1)、适应性选择明文攻击(adaptive chosen plaintext attack,简称 CPA2)、非适应性选择密文攻击(non-adaptive chosen ciphertext attack,简称 CCA1)、适应性选择密文攻击(adaptive chosen ciphertext attack,简称 CCA2).

一般使用安全目标与攻击模型相结合的方式来定义加密方案的安全性,这与保留格式加密的 4 种安全目标和上述的攻击模型相结合,就得到了保留格式加密主要的安全性,比如 PRP-CPA1,PRP-CPA2,PRP-CCA1,PRP-CCA2,MR-CPA1,MR-CPA2,MR-CCA1,MR-CCA2 等.很显然,PRP-CCA2 具有最高的安全性,而 MR-CPA1

© 中国科学院软件研究所 http://www.jos.org.cn

是保留格式加密最基本的安全性.

如果攻击者只允许对加密预言机进行查询,而不允许查询解密预言机,就定义了选择明文安全的攻击模型.在这种情况下,如果要求攻击者在第 1 次查询加密预言机之前就必须准备好所有要查询的问题,则定义了非自适应性选择明文攻击的游戏模型,即 PRP-CPA1;相反地,如果允许攻击者通过分析前面的查询结果再给出下一次要查询的问题,那么就定义了自适应性选择明文攻击的游戏模型,即 PRP-CPA2.

上一节关于安全目标的描述都基于 CPA 攻击模型.如果要达到 CCA 安全标准,上述的游戏模型中可以通过加入一个解密预言机来完成,比如在 $\mathrm{PRP}_{\mathcal{E}_{\mathrm{FPE}}}$ 游戏中引入解密预言机 Oracle(Dec),用于响应攻击者的解密查询.如果 b=1,解密预言机 Oracle(Dec)以保留格式加密方案对输入的密文进行解密并响应;如果 b=0,解密预言机 Oracle(Dec)以随机选取的方式置换 P 对输入的密文进行解密并响应.在这种情况下,如果攻击者既允许对加密预言机进行查询,又允许对解密预言机进行查询,同时要求攻击者在第 1 次查询加密预言机之前就必须准备好所有要查询的问题,于是定义了非自适应性选择密文攻击的游戏模型,即 PRP-CCA1;相反地,如果允许攻击者通过分析前面的查询结果再给出下一次要查询的问题,那么就定义了自适应性选择密文攻击的游戏模型,即 PRP-CCA2.

5 应用领域及待解决的问题

自 2008 年美国 Voltage 公司公布其安全产品所使用的 FPE 技术的白皮书[7]以来,FPE 得到了更多研究学者的关注,发展成为极具实用价值的密码学体系:在加密模型方面,提出的典型模型(如 FFSEM[8],RtE[9]与 FFX 模型[10])为特定领域的 FPE 问题提供了较好的解决办法;在应用研究方面,目前已提出了适用于支付卡行业安全的信用卡号[7]等 FPE 方案、适用于数据库加密的日期型 FPE 方法[19]、满足格式兼容要求的 JPEG2000 加密[38]等实用的解决方案,开拓了 FPE 在不同领域的应用研究.

5.1 应用领域

由于 FPE 保持密文与明文具有相同格式的特性,因此适合于格式敏感的数据加密领域,主要包括:

1) 增强数据库应用系统的安全性

在数据库中加密数据一直是一个难题,因为加密数据库中的信息就意味着扩充数据并改变格式.然而,多数大型商用应用系统都是基于数据库的应用系统,如金融、社保、电子政务、电子商务等,如果数据库中存储的大量用户敏感信息(如银行卡号、社保卡号、用户名和密码等)被窃取,将造成致命的破坏.引入 FPE 技术,将极大地提高数据库的安全性.无论新部署还是已有的数据库应用系统,都可以引入 FPE 技术增强系统安全性,因为其具有以下优势:① 不改动现有软件系统的代码;② 不改动现有数据库结构.

2) 数据遮蔽

数据遮蔽源于解决数据从生产环境向测试环境(或者开发环境)导入时可能产生的数据内容安全问题,它通过克隆原始数据进行掩码转换,输出一个与原数据格式、关联等一模一样的数据,用以进行功能测试、性能测试和模拟测试等.

数据遮蔽内嵌丰富的数据修改规则,通过各种复杂算法,可以自动批量、快速地完成对敏感数据的修改,同时保证克隆出来的数据库的数据量完全等同于生产库的数据量;敏感数据(如身份证号、电话号码、信用卡号码等)又作了伪装,看起来是真实数据,实际上是已经进行了修改的假数据,从而消除了敏感数据的泄露隐患.如果充分且合理利用的话,数据遮蔽必成为保障数据安全的重要技术.

3) 支付卡行业安全

支付卡行业数据安全标准(payment card industry data security standard,简称 PCI DSS)是一套被广为接受的政策和程序,目的是为了保证信用卡、借记卡和现金卡交易的安全,保护持卡人的个人信息,防止被他人利用.PCI DSS 所规定和阐述的六大目标之一就是:持卡人信息无论存在哪里,都必须受到保护.应确保存放的社会保险号和身份证号等重要数据不被破解.当持卡人的数据通过公共网络进行传送时,这些数据必须经过有效的加

© 中国科学院软件研究所 http://www.jos.org.cn

密.数据加密技术在各种形式的信用卡交易中都是非常重要的.

金融交易过程很复杂、约束也很多,传统的分组密码加大了改变这些系统的复杂度,代价昂贵.而 FPE 因为实现简单,保留敏感信息的格式,避免了从根本上去设计和审查整个系统的繁杂.

4) 格式兼容的加密领域

除了对上述敏感信息的数据加密以外,FPE 还非常适合于遵循既有协议格式、格式兼容的数据加密的应用领域,这将是保留格式加密在未来的一个重要的应用领域.

比如,2010 年,Stütz[38]讨论了 FPE 在多媒体加密领域的应用,指出 FPE 方法可以应用到 JPEG 2000 的加密中.JPEG 2000 的加密有两类解决方案:在图像压缩过程中进行加密;对压缩后的码流进行加密.FPE 技术应用于后者.JPEG 2000 标准将范围 0xff90 至 0xffff 之间的值分配作为限定码流的标记码,通常存在于包头中.因此,JPEG 2000 压缩流的包体中不包括超过 0xff8f 的序列,且不以 0xff 结尾.为了保证加密后的图像仍然能被标准解码器识别并解码,必须满足不产生多余的标记字段,即加密后的码流也不包括超过 0xff8f 的序列,且不以 0xff 结尾这样的条件.Stütz[38]通过有限自动机描述了 JPEG 2000 压缩流的包体部分,从而将这个加密问题转化为了正则语言上的 FPE 问题.

5.2 待解决的问题

FPE 问题的复杂性在于待解决问题消息空间的复杂性,已有的模型为特定领域的 FPE 问题提供了较好的解决办法.然而,目前的研究工作还存在很多亟待解决的问题,主要包括:

1) 性能问题

无论是应用 FFSEM[8],还是 FFX 模型[10]来解决任意有限域上的 FPE 问题,理论上都需要使用 Cycle-Walking,因而存在不确定的性能问题.

RtE 方法[9]较通用,对于存在高效排序算法的可描述的消息空间,可以通过“排序后加密”的思想达到保留格式加密的目的.然而,并不是所有消息空间都存在高效的排序算法,而且不同空间的排序算法其效率也各不相同,因此无法定量考核 RtE 方法的性能.

现有的 FPE 方案多数基于 Feistel 网络来构造指定分组长度的对称密码,通常使用基础分组密码来构造伪随机函数,执行一次保留格式加密要执行数次基础分组密码,相比基础分组密码性能偏低.

因此,在保证安全性的基础上,优化 Feistel 网络,或者采用其他构建方式,设计高效的 FPE 算法,仍然是未来一个待解决的问题.

2) 完整性认证问题

目前的 FPE 加密模型都只能完成加密,而不能对消息提供完整性认证功能.消息认证通常需要通过附加认证码或进行签名来实现,这意味着要扩充密文空间,与 FPE 确保密文和明文在相同的消息空间是一种矛盾,该问题是对 FPE 的新挑战.

3) FPE 解决实际应用存在的关键问题

FPE 为保留数据的类型和长度提供了相对完美的解决方案,并具有较高的安全性.然而,每种应用问题都具有特定的格式要求或者某些约束,FPE 需要为此量身定制,而且由于加密破坏了明文数值,可能会违背一些应用的约束,从而增加了 FPE 应用的难度.

一个 FPE 的典型应用难题是 FPE 在数据库加密中的关键问题,由于加密破坏了明文数值,使其无法满足常规数据库的一些查询和聚集操作,若要对加密的字段进行模糊查询或者统计、平均、求和等数学运算很困难,该问题也是制约保留格式加密应用到数据库加密中的关键问题.如果对所有加密数据进行解密,然后再执行查询、聚集操作,由于解密操作开销巨大,这将对查询性能造成极大影响.因此,如果将 FPE 广泛应用到数据库加密中,需要找到一种机制,既能保证其安全性,又不会对查询、聚集操作等性能产生较大影响.

与该问题相关的两类研究领域为保序加密和秘密同态.保序加密[39−42]是一种对数值型数据保留顺序的加密方案,它允许在密文上直接进行比较等操作,而不需要对其进行解密.秘密同态[43−45]是指通过构造明文空间上的秘密同态函数,实现在不进行解密的情况下直接对密文进行数学运算等.

© 中国科学院软件研究所 http://www.jos.org.cn

如何在 FPE 中引入保序和同态的性质,使其支持常规数据库的一些查询和聚集操作,从而更好地应用到数据库加密领域,将是 FPE 未来研究的一个挑战.

6 结 论

本文关注近年来保留格式加密研究的进展,主要从基本方法、加密模型的构造、安全性等角度对已有的研究成果进行了总结分析,以期对保留格式加密在国内的研究起到一定的推动作用.

从构造角度,一方面,Generalized-Feistel 方法得到了更多的关注,FPE 模型通常都采用 Feistel 网络来构造满足要求的对称密码.常见的 Feistel 网络类型主要包括传统 Feistel 网络(平衡 Feistel 网络)、非平衡 Feistel 网络、交互式 Feistel 网络及以上类型的数值形式,它们都是 2-分割的 Feistel 网络类型,即将输入划分为左右两部分.最近,*l*-分割的 Type-2 Feistel 网络也被应用到构建 FPE 方案中;另一方面,FPE 模型的设计思想逐渐向降低问题域的复杂性方向发展,无论是 RtE 方法还是 FFX 模型,都将 FPE 问题转化到等价的具有较低复杂度的整数集上.这种通过降低问题域的复杂性来解决复杂消息空间上的 FPE 问题的解决方法,将会成为保留格式加密领域的一种可采纳的有效研究手段.

从安全性角度,保留格式加密主要有 4 种安全目标:PRP,SPI,MP 和 MR,本文介绍了这些安全目标并详细描述了与之相关的游戏模型.PRP 是保留格式加密的标准安全目标;SPI 是 PRP 的一个变种,与其具有类似的安全性,虽然安全性有损,但可以达到比 PRP 更好的安全边界;MP 把语义安全带给 FPE,它要求密文不能泄露真实信息,安全性要比 PRP 和 SPI 略低;最基本也是最常用的安全性要求就是对抗适应性或者非适应性攻击下的 MR,MR 从整体上使对手无力从密文得到明文.保留格式加密主要关心的攻击模型主要包括 4 种:CPA1,CPA2,CCA1,CCA2.使用安全目标与攻击模型相结合,可以定义保留格式加密的安全性.在所定义的安全性中, PRP-CCA2 具有最高的安全性,MR-CPA1 是保留格式加密最基本的安全性.

随着 2008 年 Voltage 公司 FPE 白皮书[7]的公布,FPE 问题逐渐成为密码学领域中的一个研究热点.当前已有一些研究成果:在加密模型方面,提出的典型模型(如 FFSEM[8],RtE[9]与 FFX 模型[10])为特定领域的 FPE 问题提供了较好的解决办法;在应用研究方面,已提出了适用于支付卡行业安全的信用卡号等 FPE 方案[7]、适用于数据库加密的日期型 FPE 方法[19]、满足格式兼容要求的 JPEG 2000 加密等实用的解决方案[38],开拓了 FPE 在不同领域的应用研究.然而,由于保留格式加密待解决的问题域的复杂性,目前的研究工作还存在很多亟待解决的问题:如性能问题、完整性认证问题以及 FPE 解决实际应用存在的关键问题等,都是需要进一步解决的问题.

References:

[1] Radhakrishnan R, Kharrazi M, Memon N. Data masking: A new approach for steganography? The Journal of VLSI Signal Processing, 2005,41(3):293−303. [doi: 10.1007/s11265-005-4153-1]

[2] Smith HE, Brightwell M. Using datatype-preserving encryption to enhance data warehouse security. In: Proc. of the 20th National Information Systems Security Conf. 1997. 141−149. http://csrc.nist.gov/nissc/1997/proceedings/141.pd

[3] National Bureau of Standards. FIPS PUB 74, Guidelines for Implementing and Using the NBS Data Encryption Standard, 1981.

[4] Black J, Rogaway P. Ciphers with arbitrary finite domains. In: Preneel B, ed. Proc. of the Topics in Cryptology—CT-RSA 2002. LNCS 2271, San Jose: Springer-Verlag, 2002. 114−130. [doi: 10.1007/3-540-45760-7_9]

[5] Luby M, Rackoff C. How to construct pseudorandom permutations from pseudorandom functions. SIAM Journal of Computing, 1988,17(2):373−386. [doi: 10.1137/0217022]

[6] Patarin J. Security of random Feistel schemes with 5 or more rounds. In: Franklin M, ed. Advances in Cryptology—CRYPTO 2004. LNCS 3152, Santa Barbara: Springer-Verlag, 2004. 106−122. http://www.iacr.org/archive/crypto2004/31520105/Version%20courte%20Format%20Springer.pdf [doi: 10.1007/978-3-540-28628-8_7]

[7] Spies T. Format preserving encryption. Unpublished Voltage White Paper. 2008. https://www.voltage.com/pdf/Voltage-Security-WhitePaper-Format-Preserving-Encryption.pdf

[8] Spies T. Feistel finite set encryption mode. 2008. http://csrc.nist.gov/groups/ST/toolkit/BCM/documents/proposedmodes/ffsem/ffsem-spec.pdf

© 中国科学院软件研究所 http://www.jos.org.cn

[9] Bellare M, Ristenpart T, Rogaway P, Stegers T. Format-Preserving encryption. In: Jacobsn MJ, eds. Proc. of the Selected Areas in Cryptography (SAC 2009). LNCS 5867, Calgary: Springer-Verlag, 2009. 295−312. [doi: 10.1007/978-3-642-05445-7_19]

[10] Bellare M, Rogaway P, Spies T. The FFX mode of operation for format-preserving encryption. 2010. http://csrc.nist.gov/groups/ST/toolkit/BCM/documents/proposedmodes/ffx/ffx-spec.pdf

[11] Schneier B, Kelsey J. Unbalanced Feistel networks and block cipher design. In: Gollmann D, ed. Proc. of the Fast Software Encryption'96. LNCS 1039, Cambridge: Springer-Verlag, 1996. 121−144. http://www.schneier.com/paper-unbalanced-feistel.pdf [doi: 10.1007/3-540-60865-6_49]

[12] Liskov M, Rivest RL, Wagner D. Tweakable block ciphers. In: Advances in Cryptology—CRYPTO 2002. LNCS 2442, Santa Barbara: Springer-Verlag, 2002. 31−46. [doi: 10.1007/s00145-010-9073-y]

[13] Eric B, Thomas P, Jacques S. BPS: A format-preserving encryption proposal. An NIST Submitted Proposal, 2010. http://brutus.ncsl.nist.gov/groups/ST/toolkit/BCM/documents/proposedmodes/bps/bps-spec.pdf

[14] National Institute of Standards and Technology. SP800-67: Recommendation for the triple data encryption algorithm (TDEA) block cipher. 2004. http://purl.access.gpo.gov/GPO/LPS69978

[15] National Institute of Standards and Technology. FIPS 197: Advanced Encryption Standard. 2001. http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf

[16] National Institute of Standards and Technology. FIPS 180-2: Secure Hash Standard. 2002. http://csrc.nist.gov/publications/fips/fips180-2/fips180-2.pdf

[17] Morris B, Rogaway P, Stegers T. How to encipher messages on a small domain—Deterministic encryption and the thorp shuffle. In: Advances in Cryptology—CRYPTO 2009. Santa Barbara: Springer-Verlag, 2009. http://www.cs.ucdavis.edu/~rogaway/papers/thorp.pdf [doi: 10.1007/978-3-642-03356-8_17]

[18] Jia CF, Liu ZL, Li JW, Dong ZQ, You XY. A new integer FPE scheme based on Feistel network. In: Zhu X, ed. Proc. of the Int'l Conf. on Services Science Management and Engineering 2010. Tianjin: IEEE Press, 2010. 305−308.

[19] Liu ZL, Jia CF, Li JW, Cheng XC. Format-Preserving encryption for datetime. In: Chen W, ed. Proc. of the Intelligent Computing and Intelligent Systems 2010, Vol.2. Xiamen: IEEE Press, 2010. 201−205. [doi: 10.1109/ICICISYS.2010.5658769]

[20] Cover T. Enumerative source encoding. IEEE Trans. on Information Theory, 1973,19(1):73−77. [doi: 10.1109/TIT.1973.1054929]

[21] Mäkinen E. Ranking and unranking left Szilard languages. Int'l Journal of Computer Mathematics, 1998,68(1-2):29−38. [doi: 10.1080/00207169808804677]

[22] Knuth DE. The Art of Computer Programming, Vol. 2: Seminumerical Algorithms. 3rd ed., Addison-Wesley, 1997.

[23] Kelsen P. Ranking and unranking trees using regular reductions. In: Puech C, ed. Proc. of the 13th Annual Symp. on Theoretical Aspects of Computer Science. LNCS 1046, Grenoble: Springer-Verlag, 1996. 581−592. http://www.springerlink.com/index/p2x337k42l09w032.pdf [doi: 10.1007/3-540-60922-9_47]

[24] Liebehenschel J. Ranking and unranking of a generalized DYCK language and the application to the generation of random trees. In: Séminaire Lotharingien de Combinatoire. 2000. 43−62.

[25] Hoang VT, Rogaway P. On generalized Feistel networks. In: Rabin T, ed. Advances in Cryptology—CRYPTO 2010. LNCS 6223, Santa Barbara: Springer-Verlag, 2010. 613−630. [doi: 10.1007/978-3-642-14623-7_33]

[26] Anderson R, Biham E. Two practical and provably secure block ciphers: BEAR and LION. In: Gollmann D, ed. Proc. of the Fast Software Encryption'96. LNCS 1039, Cambridge: Springer-Verlag, 1996. 113−120. http://www.springerlink.com/index/P3L22063H7817J35.pdf [doi: 10.1007/3-540-60865-6_48]

[27] Lucks S. Faster Luby-Rackoff ciphers. In: Gollmann D, ed. Proc. of the Fast Software Encryption '96. LNCS 1039, Cambridge: Springer-Verlag, 1996. 189−203. http://weisskugel.informatik.uni-mannheim.de/people/lucks/papers/pdf/LR-fast.pdf.gz [doi: 10.1007/3-540-60865-6_53]

[28] Zheng YL, Matsumoto T, Imai H. On the construction of block ciphers provably secure and not relying on any unproved hypotheses. In: Brassard G, ed. Proc. of the 9th Annual Int'l Cryptology Conf. LNCS 435, Berlin: Springer-Verlag, 1990. 461−480. ftp://www.hacktic.nl/pub/mirrors/Advances%20in%20Cryptology/HTML/PDF/C89/461.PDF [doi: 10.1007/0-387-34805-0_42]

[29] Goldwasser S, Micali S. Probabilistic encryption. Journal of Computer and System Sciences, 1984,28(2):270−299. [doi: 10.1016/0022-0000(84)90070-9]

[30] Micali S, Rackoff C, Sloan R. The notion of security for probabilistic cryptosystems. SIAM Journal on Computing, 1988,17(2):412−426. [doi: 10.1137/0217025]

[31] Dolev D, Dwork C, Naor M. Nonmalleable cryptography. SIAM Journal on Computing, 2000,30(2):391−437. [doi: 10.1137/S0097539795291562]

© 中国科学院软件研究所 http://www.jos.org.cn

[32] Bellare M, Rogaway P. Optimal asymmetric encryption—How to encrypt with RSA. In: De Santis A, ed. Advances in Cryptology-EUROCRYPT'94. LNCS 950, Perugia: Springer-Verlag, 1994. 92−111.

[33] Bellare M, Rogaway P. Random oracles are practical: A paradigm for designing efficient protocols. In: Proc. of the 1st ACM Conf. on Computer and Communications Security. Fairfax: ACM Press, 1993. 62−73. http://seclab.cs.ucdavis.edu/papers/Rogaway/ro.pdf [doi: 10.1145/168588.168596]

[34] Goldreich O. A uniform complexity treatment of encryption and zero-knowledge. Journal of Cryptology, 1993,6(1):21−53. [doi: 10.1007/BF02620230]

[35] Goldreich O. Foundations of Cryptography, Vol II: Basic Applications. Cambridge: Cambridge University Press, 2004.

[36] Goldreich O, Goldwasser S, Micali S. How to construct random functions. Journal of the ACM, 1986,33(4):792−807. [doi: 10.1145/6490.6503]

[37] Desai A, Miner S. Concrete security characterizations of PRFs and PRPs: Reductions and applications. In: Okamoto T, ed. Advanced in Cryptology- ASIACRYPT 2000. LNCS 1976, Kyoto: Springer-Verlag, 2000. 503−516. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.119.1491&rep=rep1&type=pdf [doi: 10.1007/3-540-44448-3_39]

[38] Stütz T, Uhl A. Efficient format-compliant encryption of regular languages: Block-Based Cycle-Walking. In: De Decker B, ed. Proc. of the 11th IFIP TC 6/TC 11 Int'l Conf. LNCS 6109, Linz: Springer-Verlag, 2010. 81−92. http://wavelab.at/papers/Stuetz10a.pdf [doi: 10.1007/978-3-642-13241-4_9]

[39] Ozsoyoglu, G, Singer DA, Chung SS. Anti-Tamper databases: Querying encrypted databases. In: Proc. of the 17th Annual IFIP WG, Vol.11. 2003. http://vorlon.case.edu/~chung/Publication/TechnicalReportIFIP03RevisedPaper.pdf

[40] Agrawal R, Kiernan J, Srikant R, Xu YR. Order preserving encryption for numeric data. In: Proc. of the 2004 ACM SIGMOD Int'l Conf. on Management of Data. New York: ACM Press, 2004. 563−574. http://rsrikant.com/papers/sigmod04.pdf [doi: 10.1145/1007568.1007632]

[41] Boldyreva A, Chenette N, Lee Y, O'Neill A. Order-Preserving symmetric encryption. In: Joux A, ed. Advances in Cryptology-EUROCRYPT 2009. LNCS 5479, Perugia: Springer-Verlag, 2009. 224−241. http://www.iacr.org/archive/eurocrypt2009/54790225/54790225.pdf [doi: 10.1007/978-3-642-01001-9_13]

[42] Lee S, Park TJ, Lee D, Nam T, Kim S. Chaotic order preserving encryption for efficient and secure queries on databases. IEICE Trans. on Information and Systems, 2009,E92-D(11):2207−2217. [doi: 10.1587/transinf.E92.D.2207]

[43] Rivest RL, Adleman L, Dertouzos ML. On data banks and privacy homomorphisms. Foundations of Secure Computation, 1978, 169−178.

[44] Gentry C. Fully homomorphic encryption using ideal lattices. In: Proc. of the 41st Annual ACM Symp. on Theory of Computing (STOC 2009). New York: ACM Press, 2009. 169−178. http://boxen.math.washington.edu/home/wstein/www/home/watkins/CG.pdf [doi: 10.1145/1536414.1536440]

[45] Van Dijk M, Gentry C, Halevi S, Vaikuntanathan V. Fully homomorphic encryption over the integers. In: Gilbert H, ed. Advances in Cryptology-EUROCRYPT 2010. LNCS 6110, Perugia: Springer-Verlag, 2010. 24−43. [doi: 10.1007/978-3-642-13190-5_2]

刘哲理(1978－),男,山东潍坊人,博士,主要研究领域为密码学及应用,智能卡操作系统.

李经纬(1987－),男,博士生,主要研究领域为密码学.

贾春福(1967－),男,博士,教授,博士生导师,主要研究领域为信息安全与可信计算,恶意代码发现与分析.

© 中国科学院软件研究所 http://www.jos.org.cn