密码学进展—CHINACRYPT'92

第二届中国密码学学术会议论文集

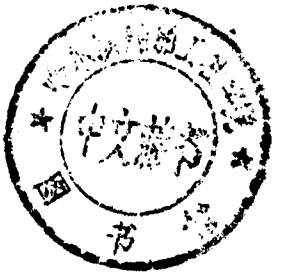
陶仁骥　李　祥　裴定一　编

科学出版社

密码学进展——CHINACRYPT'92

第二届中国密码学学术会议论文集

陶仁骥　李　祥　裴定一 编



科学出版社

1992

（京）新登字092号

内 容 简 介

本书为第二届中国密码学学术会议论文集，收录论文41篇及1篇特邀报告，内容涉及密码学的各个领域，其中包括与密码学有关的数学和计算机科学，以及密码学在信息安全方面的应用.

本书的主要内容有：密码分析，密码体制，认证码，自动机密码，序列与线性复杂度，与密码学有关的数论、统计、组合和逻辑问题，复杂性理论，计算机安全和密钥管理等.

本书可供从事密码、信息安全、数学、计算机通信专业的科技人员和高校的师生阅读、参考.

密码学进展——CHINACRYPT'92

第二届中国密码学学术会议论文集

陶仁骐 李 祥 裴定一 编

责任编辑 那莉莉 李淑兰

科学出版社出版

北京东黄城根北街16号

邮政编码：100707

北京市华星机算机公司激光照排

天津市静一胶印厂印刷

新华书店北京发行所发行 各地新华书店经售

*

1992年8月第 一 版 开本：787×1092 1/16

1992年8月第一次印刷 印张：17$\frac{1}{2}$

印数：1—1 500 字数：406 000

ISBN 7-03-003069-9/TP·224

定价：18.00元

前　言

第二届中国密码学学术会议（1992年8月20～24日在贵阳举行）上提出的研究工作涉及密码学的各个领域，其中包括与密码学有关的数学和计算机科学，以及在信息安全方面的应用等内容，诸如密码分析，密码体制，认证码，自动机密码，序列与线性复杂度，线性阵列，与密码学有关的代数、统计、组合、数论和逻辑问题，复杂性理论，计算机安全和协议，密钥管理，等等.

本书包括的41篇论文和1篇特邀报告是关于上述工作的具体论述.本届会议共收到文章88篇，每篇文章均由程序委员会指定的1～3名专家进行评审. 除个别迟到的以外，所有文章或摘要都在截稿后及时送至程序委员会的各位成员. 3月6日在北京召开的程序委员会会议，对文章作出了最后的选择. 出席这次会议的有王萼芳、卢开澄、刘木兰、朱洪、杨义先、杨君辉、陶仁骥、黄民强、黄祖良和裴定一. 未能与会的委员中，王育民、李祥、肖国镇以书面或其他方式表达了他们的意见.

我们衷心感谢所有投稿的作者对会议的关心和支持!衷心感谢程序委员会的同事和所有的文章评审者的辛勤工作!最后，还要感谢中国科学院软件研究所的许多同事，特别是自动机理论组的全体同事在本届会议筹备和召开的整个过程中所给予的帮助!

第二届中国密码学学术会议程序委员会

目　录

密码分析

密钥体制中的密钥熵漏与密钥信息冗余（特邀报告） …… 曾肯成 (1)
相关免疫缺陷与复合序列的攻击 …… 曾肯成　吕述望　杨君辉 (2)
偏差分析的一个例子——分析一个分组密码方案
…… 杨君辉　戴宗铎　曾肯成 (11)
Pieprzyk 公开钥密码及 Yang 破译方法分析 …… 谢冬青 (16)
W 签名方案与 ECPS2中的签名都是不可信赖的 …… 隆永红 (20)

单钥密码体制

热流密码体制线性模型加解密理论及试验方法分析
…… 楚泽甫　宋惠元　赵新田　李东红 (24)
一种实用的模拟密话系统…… 高玉良　钟持瑞 (34)

双钥密码体制

具有直接身份鉴别和可靠数字签名功能的基于 ID 的公开密钥密码体制
…… 陆浪如　赵仁杰 (40)
一种抗幂剩余函数周期性的新型公钥密码体制…… 乔秦宝　陈新明 (49)
复合加密的公钥密码系统…… 于秀源　薛昭雄 (55)
关于 Niederreiter 代数码公钥密码体制的安全性及参数优化
…… 李元兴　王新梅　成　坚 (59)

认证码

认证码及其构造的一些研究…… 裴定一 (66)
用平面曲线构造认证码…… 王也菁 (74)

自动机密码

拟线性有限自动机的可逆性…… 陈世华　陶仁骐 (77)
基于身份的密码体制和数字签名的有限自动机公开钥密码实现
…… 陶仁骐　陈世华 (87)
FA 公开密钥密码体制的软件实现
…… 张焕国　戴大为　覃中平　吴　逵　崔保秋　韩　海 (105)
有限自动机公开钥密码体制和数字签名的软件实现 …… 李健宝　高　翔 (110)

序列、线性复杂度和线性阵列

有限自动机输出序列的周期和线性复杂度 …… 马永彪 (115)
延迟采样序列的构造与性质 …… 孙登峰 (122)
环 $\mathbb{Z}/(p^e)$ 上线性序列簇 $G(f(x))$ 的结构及和序列的若干性质
…… 戚文峰 周锦君 (132)
序列综合问题的新方法及其应用 …… 陆佩忠 周锦君 宋国文 (140)
周期序列的布尔多项式表示及序列复杂度快速计算
…… 林须端 胡正名 蔡长年 (148)
De Bruijn 序列的 k 次齐次复杂度 …… 朱士信 (155)
产生一类 M 序列 …… 郭宝安 蔡长年 (160)
线性递归 m- 阵列窗口的计算 …… 刘木兰 李俊全 (165)

数论、代数与统计

关于安全质数和强质数的分布 …… 王 炜 (172)
DES 变换表研究之二——置换群方法引论 …… 徐茂智 (181)
置换方程的计算机解法 …… 王 杰 (186)
F_2 上矩阵秩的概率分布及其渐近性质 …… 生加明 黄民强 (195)
一类受污染数据的统计处理 …… 赵仁杰 郭启容 廉玉忠 (199)
GF(q) 上可逆矩阵的数量特征刻划与无重复枚举 …… 章中平 张焕国 (204)
拉丁阵的随机生成及其合痕类代表元的生成算法 …… 高 翔 (209)

布尔函数和逻辑设计

布尔函数的 Walsh 变换的推广及布尔函数的非线性逼近
…… 周锦君 陈卫红 (216)
布尔函数的独立性及其密码应用 …… 章照止 (222)
用于密码的 N 元 H- 布尔函数的计数 …… 杨义先 (227)
有限域上快速乘法器的新设计 …… 朱剑英 胡正名 (231)

复杂性理论

关于伪随机数发生器的保密性与不可区分性 …… 李 祥 (235)
关于线性复杂度与 Kolmogorov 复杂度之间关系的一个注记 …… 鲍 丰 (242)

计算机安全和密钥管理

为 DECnet-VAX 增设网络安全保密设施的研究 …… 宋自林 倪桂强 (247)
一个多用户验证安全协议的设想 …… 鲍振东 游之墨 赵一鸣 徐瑞芬 (253)
具有身份鉴别功能的密钥分配体制 …… 谯谨旭 (262)
关于密钥分享的2次密钥方案 …… 曹珍富 (267)

密钥体制中的密钥熵漏与密钥信息冗余

曾 肯 成

(中国科学院研究生院信息安全国家重点实验室，北京 100039)

[摘要]为了提出一些一般的观点和方法来评价某些保密密钥密码体制的质量，我们提出了两个方面的问题：

1)密钥体制中各种不同形式的密钥熵漏及其对整个体制安全性的影响.

2)密钥体制中某些子密钥的冗余性.

本报告围绕上述两个方面的问题进行并通过一些具体的例子来说明对密码体制进行这两方面评价的必要性.

On the Key Information Leakage and Redundancy in Cryptosystems

Zeng Kencheng

(*State Key Laboratory of Information Security, Graduate School of Academia Sinica, Beijing* 100039, *PRC*)

Abstract

In a strive to find general principles useful for the assessment of various secret key cryptosystems, the concepts of key information leakage and key information redundancy in cryptosystems are introduced and illustrated by concrete examples.

相关免疫缺陷与复合序列的攻击

曾肯成　吕述望

（中国科学院研究生院信息安全国家重点实验室，北京 100039）

杨君辉

（中国科学院计算中心，北京 100080）

［**摘要**］ 本文利用复合序列的 I/O 相关免疫缺陷，对 Jennings-序列给出统计分析的攻击方法.

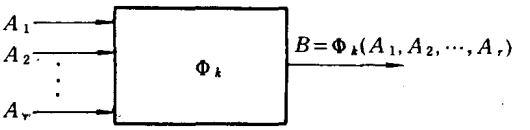
一、密码序列的相关免疫缺陷(CID)

为了获得高强度的密码序列，最常用的办法之一是从一组较简单的原始序列，比如说，从一组具有中等级数的 m-序列

$$A_1,\ A_2,\ \cdots,\ A_r$$

出发，经过某种非线性变形，得到一个新的序列

$$B = \Phi_k(A_1, A_2, \cdots, A_r)$$



这里的变形 Φ_k 由参数 k 来确定，后者也就是相应密码序列中的密钥. 如此产生的序列 B 究竟应满足怎样一些要求方可作为密码序列使用，恐怕没有也不可能提出任何简单明确的判定准则. 在讨论序列密码的公开文献中，对序列 B 只提到了如下一些最初步的要求：

1)有足够大的周期 $P(B)$；

2)有足够高的线性复杂度 $L(B)$；

3)有能够通过某些检验标准的统计均衡性.

这样一些要求，实际上只反映了序列 B 的一般结构，完全没有涉及这个序列对抗密码学攻击的能力. 从保证序列 B 的抗攻击强度这一角度来看，我们认为至少应明确提出如下一项补充要求：

4)输入序列 $A_i(1\leqslant i\leqslant r)$ 与输出序列 B 之间不应存在可被攻击一方利用的明显相关特征.

密码序列中的输入输出 (I/O) 相关性有各种不同的表现形式，其中一种形式表现为输入序列 $A_i=\{a_i(t)\}$ 的信号与输出序列 $B=\{b(t)\}$ 的信号之间的某种符合优势：如果在密码体制的工作过程中有

$$\text{Prob}(a_i(t)=b(t))=\frac{1}{2}+\varepsilon \qquad \varepsilon>0$$

我们就说输入序列 A_i 与输出序列 B 之间有符合优势 $\varepsilon>0$，或者说，序列 A_i 在序列 B 中的相关走漏率为 ε. 当这种相关走漏大到一定程度时，就有可能被攻击一方利用，根据 B 来推断 $A_i(1\leqslant i\leqslant r)$ 及相应密钥 K. 在这种情况下我们就说该序列密码体制带有相关免疫缺陷(CID——Correlation Immunity Deficiency). CID 带病的序列不宜用作密码序列.

公开文献中见到的一些较简单的密码序列或者直接带有 CID，其中带病最为突出而又经常被人引用的例子有：

1)Geffe-序列[1,2,12]. 这个序列有简单的表达式

由于

$$\begin{aligned} B&=A_1A_2+\overline{A_2A_3}\\ B&=\overline{A}_2(A_1+A_3)+A_1\\ &=A_2(A_1+A_3)+A_3 \end{aligned}$$

所以我们有

$$\text{Prob}(a_1(t)=b(t))=\text{Prob}(a_3(t)=b(t))=\frac{3}{4}$$

因此，输入序列 A_1，A_3 在输出序列中的相关走漏率为 $\varepsilon=0.25$. 我们可以把序列 B 看成 m-序列 A_1(或 A_3)带有 25%错。如果 m-序列 $A_i(1\leqslant i\leqslant 3)$ 的极小多项式为已知，那么我们就可以用线性纠错的方法根据序列 B 的一个适当长的截段还原 A_1(和 A_3). 序列 A_1 和 A_3 还原之后，再确定序列 A_2 便只需利用密码分析的简单技巧即可.

当所用输入序列 $A_i(1\leqslant i\leqslant 3)$ 均为 40～50 级 m-序列，其极小多项式为已知时，利用中国科学院研究生院信息安全国家重点实验室的算法[3]，根据 Geffe-序列的一个长为 2000 比特左右的截段，在普通的 PC 机上来解算三个输入序列，所需工作时间仅以若干秒计.

2)Pless-序列[4]. 与 Geffe-序列相比较，这个序列要稍许复杂一些，但稍经变形后仍可发现明显的 CID. 黄民强[5]利用 Pless-序列的这种 CID 大大改进了 Rubin[6]对这个序列所作的攻击.

本文的目的是要利用 CID 分析的技巧，对近年来颇受重视的复合序列(multiplexed sequence)及此种序列的一个大大强化了的变种——钟控复合序列作一次相当彻底的密码学攻击. 本文下述钟控复合序列表明，戏剧性地提高一个序列的线性复杂度或为密码序列的线性复杂度设计某种确保下界并非难事，而下述的“复合序列的破译”表明，发现和消除密码序列的潜在 CID 则是一个必须郑重其事地加以对待的问题.

二、Jennings-复合序列

复合序列的讨论最初见于 S. M. Jennings 的博士论文[7]，其后又受到另一些作者的注意[8,9]，并受到了较好的评价[10~12]. 在 Jennings 所讨论的方案中，两个原始的输入序列 A 和 B 是级数分别为 m 和 n 的 m -序列，且 m 与 n 互素. 序列 A 的发生器 SR_1 称为主选序列

发生器，序列 B 的发生器 SR_2 称为被选序列发生器. 主选序列发生器 SR_1 带有 $h \leqslant [\log_2(n-1)]$ 个抽头，其位置可以是任意的. 为简单计，在下面的讨论中规定这些抽头取自 SR_1 的第 0 至 $h-1$ 位寄存单元. 这样，在任意时刻 t 自 SR_1 的这 h 个单元中输出一个整数

$$k_A(t) = a(t) + a(t+1)2 + \cdots + a(t+h-1)2^{h-1}$$

Jennings 方案中的密钥表现为整数集合

$$H = \{0,1,2,\cdots,2^h - 1\}$$

到整数集合

$$N = \{0,1,2,\cdots,n-1\}$$

的一个可以随机布置的单射 r：

$$k \in H, k \to r(k) \in N$$

密钥 r 给定之后，输出序列 $C=\{c(t)\}$ 的信号按如下方式产生：在任意时刻 t 输出被选序列发生器中第 $r(k_A(t))$ 位寄存单元的内容，即有

$$c(t) = b(t + r(k_A(t)))$$

可以证明下述定理.

定理 1(Jennings)　当 m,n 互素时，Jennings-复合序列 C 的周期为

$$P(C) = (2^m - 1)(2^n - 1)$$

线性复杂度为

$$L(C) = n(1 + C_m^1 + C_m^2 + \cdots + C_m^h)$$

此外还可以证明，除非

$$\tau \equiv 0 \quad [\mathrm{mod}(2^m - 1)]$$

这个序列的自相关函数 $R(\tau)$ 没有其他突出副峰. 进一步的分析表明，Jennings-复合序列还有其他一些在文献[7]中没有讨论过的优点，兹不详述.

三、钟控复合序列

就提高线性复杂度这一点来说，Jennings 的方案并没有充分利用微电子学器件所提供的潜力. 实际上，只要对主选序列发生器 SR_1 的工作时钟稍加控制，就可保证输出序列 C 具有高得多的线性复杂度，其中一种作法如下：

1)自被选序列发生器 SR_2 的第 f 位输出一个序列 B_f(它是被选序列 $B=b_0$ 的一个 f 步平移)，并按图 1 所示方式产生的微分序列 B_f 来控制 SR_1 的工作时钟；

2)自主选序列发生器 SR_1 的某两位分别输出两个序列，将其中一个序列作一步延迟后与另一个序列作模二合成. 我们用这个合成序列来作为主选序列 A，并将它的信号输到一个 h 位移存器，以产生 Jennings 方案中的整数

$$k_A(t) = a(t) + a(t+1) \cdot 2 + \cdots + a(t+h-1) \cdot 2^{h-1}$$

延迟单元 R 的设置是为了保证新生主控序列 A 有较理想的统计均衡性. 可以证明下述定理.

定理 2　当 SR_1 和 SR_2 均为 n 级 m- 序列发生器时，输出序列的周期是

$$p = (2^n - 1)^2$$

线性复杂度是

$$L(C) = (2^n - 1)(1 + C_n^1 + C_n^2 + \cdots + C_n^h)$$

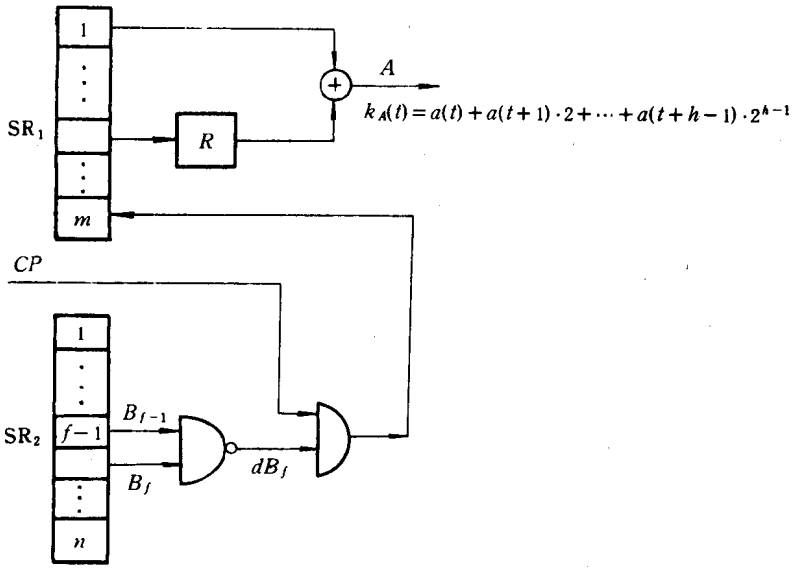


图 1

我们把如此产生的序列 C 称为钟控复合序列，以区别于没有钟控措施的 Jennings-复合序列. 从攻击上来看这两种复合序列可用同一方法来处理. 这就说明提高线性复杂度并不意味着增强抗攻击能力.

四、复合序列的攻击

复合序列的攻击指的是根据输出序列 C 的一个不太长的截段 $\bar{C}$ 来推断产生该截段时两个移存器 SR_1,SR_2 的初态 $S_0(A)$,$S_0(B)$及作为密钥使用的映射 r. 在这里我们假定 SR_1 和 SR_2 的联结多项式均为已知. 如果不从揭露序列 C 的弱点入手，完全靠穷搜来确定这三个量，那么在 Jennings-复合序列的情况下所需搜索量为

$$Q=(2^m-1)(2^n-1)A_n^{2^h}$$

而在钟控复合序列的情况下有

$$Q=(2^n-1)^2A_n^{2^h}$$

下面我们讨论钟控复合序列的攻击问题，并分以下两个步骤来进行.

首先我们讨论初态 $S_0(B)$和象集 $T=r(H)$的确定.

为了压缩搜索工作量，我们采取各个击破的攻击策略，先设法确定 SR_1 的初态$S_0(B)$及 H 在 r 下的象集 T，后者是 N 的一个 2^h 元子集. 对于任意两个二元序列

$$X=\{x(t)\} \qquad Y=\{y(t)\}$$

我们把

$$\rho(X,Y)=\mathrm{Prob}(x(t)=y(t))$$

称为这两个序列之间的相关数.

定理 3（第一 CID 定理） 如命 B_1 表示自 SR_2 的第 l 个寄存单元输出的序列，则

$$\rho(B_1,C)=\begin{cases}\dfrac{1}{2}+\dfrac{1}{2^{k+1}} & \text{如果 } l\in T\\ \dfrac{1}{2} & \text{如果 } l\notin T\end{cases}$$

[证明] 先设 $l\in T$,并设 $l=r(k),k\in H$. 我们区分 $k_A(t)=k$ 和 $k_A(t)\neq k$ 两种情况. 在条件 $k_A(t)=k$ 下,恒有

$$C(t)=b(t+r(k_A(t)))=b(t+r(k))=b(t+l)$$

而在条件 $k_A(t)\neq k$ 下,我们有 $r(k_A(t))\neq l$,因而有

$$\text{Prob}(C(t)=b(t+l))=\text{Prob}(b(t+r(k_A(t)))=b(t+l))=\frac{1}{2}$$

但由序列 A 的均衡性有

$$\text{Prob } k_A(t)=k)=\frac{1}{2^k}\qquad \text{Prob}(k_A(t)\neq k)=1-\frac{1}{2^k}$$

故知应有

$$\rho(B_l,C)=\frac{1}{2^k}1+\left(1-\frac{1}{2^k}\right)\frac{1}{2}=\frac{1}{2}+\frac{1}{2^{k+1}}$$

现设 $l\notin T$. 这时对任意 t 有 $r(k_A(t))\neq l$,因而有

$$\rho(B_l,C)=\text{Prob}(b(t+r_A(k_A(t)))=b(t+l))=\frac{1}{2}$$

定理证毕.

上面的证明中隐含了一些硬性假定,因此不能认为是严格的*,但实验结果与定理中所作的估计高度符合,故应认为这个定理的确揭示了钟控复合序列的某种 CID. 它表明序列 C 与 SR_2 的各位输出序列之间有两种不同的相关水平. 使得相关数为$\frac{1}{2}$的那些位称为弱相关位,而其余各位称为强相关位,后者共有 2^k 个.

两种相关水平的划分为根据序列 C 的一个适当长为 L 的截段 $\overline{C}$ 来确定 $S_0(B)$和 T 提供了一个有效的办法:任意选取 $S\in V_n(F_2)$作为 SR_2 的初态,自这个移存器的 n 个位输出 n 个长为 L 的截段

$$\overline{B'}_0,\overline{B'}_1,\cdots,\overline{B'}_{n-1}$$

并计算信号符合率

$$\rho(\overline{B'}_0,\overline{C}),\rho(\overline{B'}_1,\overline{C}),\cdots,\rho(\overline{B'}_{n-1},\overline{C})$$

如果根据这 n 个符合率能分辨出 2^k 个强相关位,那么所取的 S 就有可能是 $S_0(B)$,而相应强相关位的标号就组成集合 T,否则应舍弃该 S.

可是必须注意,用这里所描述的办法只能将 $S_0(B)$限制在一个很小的范围之内,而不一定能将它唯一确定. 同样,对于集合 T 也只能确定其中各个数的相对位置,而不能确定它们的绝对位置. 为了更具体地说明这一点,用 $l_{\min},l_{\max}$分别表示集合 T 中最小和最大的两个数,我们有下述定理.

定理 4 以 n 维向量 S 作为 SR_2 的初态时能按上述方式分辨出 2^k 个强相关位的充要条件是 SR_2 能由状态 SR_0 出发经 $d\leqslant l_{\min}$次推移进入状态 S,或由状态 S 出发经 $d\leqslant n-l_{\max}$

---

* 以 B 的周期为步长对 C 采样,可以给出严格证明,此处略.

-1 次推移进入状态 SR_0.

[证明] 把 SR_2 由状态 S 出发进行工作时的各位输出记作 $B_l'(0\leqslant l\leqslant n-1)$. 如果 SR_2 能由 $S_0(B)$出发经 $d\leqslant l_{\min}$次推移进入 S,那么由 $B'_{l-d}=B_l$,有

$$\rho(B'_{1-d},C)=\rho(B_l,C)$$

这样,设

$$T=\{l_0,l_1,\cdots,l_{2^k-1}\}$$

则当 SR_2 自 S 出发进行工作时,通过上述相关检验将会发现 SR_2 中的第

$$0\leqslant l_0-d,l_1-d,\cdots,l_{2^k-1}-d\leqslant l_{\max}-d$$

位是强相关位. 同样,如果 SR_2 由 S 出发经 $d\leqslant n-l_{\max}-1$ 次推移进入 $S_0(B)$,那么相关检验将会显示它的第

$$l_{\min}+d\leqslant l_0+d,l_1+d,\cdots,l_{2^k-1}+d\leqslant n-1$$

位为强相关位.

反之,如果 SR_2 由 $S_0(B)(S)$出发经 $l_{\min}<d\leqslant 2^k-1(n-l_{\max}-1<d\leqslant 2^k-1)$次推移进入 $S(S_0(B))$,则由 $l_{\min}-d<0(l_{\max}+d>n-1)$可知,以 S 作为 SR_2 的初态时,相关检验所能显示的强相关位不到 2^k 个.

证毕.

推论 1 利用相关检验的办法可根据序列 C 的一个适当长的截段确定出

$$e=n-l_{\max}+l_{\min}\leqslant n-2^k-1$$

个 n 维向量作为 SR_2 的可能初态,且右端的等号成立当且仅当 T 由 2^k 个蝉联整数组成.

推论 2 通过相关检验能将 $S_0(B)$唯一确定的充要条件是 0 与 $n-1$ 都在 T 中.

再来讨论 $S_0(A)$和 r 的确定.

SR_2 的初态 $S_0(B)$和映射 r 的象集一经确定之后,下一步是设法确定 SR_1 的初态 $S_0(A)$和映射 r 本身. 在下面的分析中,我们总是让 SR_2 从状态 $S_0(B)$开始工作,并考察另一种相关数

$$\rho_{S,k}(B_l,C)=\mathrm{Prob}(C(t)=b(t+l)\,|\,k_A(t)=k,S_0(A)=S)$$

即在条件

a)SR_1 自状态 S 开始工作

b)$k_A(t)=k$

的制约下,SR_2 的第 l 位输出序列 B_l 与复合序列 C 之间的信号符合率.

定理 5(第二 CID 定理) 视 S 与 $S_0(A)$,k 与 l 之间的不同关系,相关数 $\rho_{S,k}(B_l,C)$显示出以下三种不同水平:

$$\rho_{S,k}(B_l,C)=\begin{cases}1 & \text{如果 } S=S_0(A),r(k)=l\\ \dfrac{1}{2} & \text{如果 } S=S_0(A),r(k)\neq l\\ \dfrac{1}{2}+\dfrac{1}{2^{k+1}} & \text{如果 } S\neq S_0(A)\end{cases}$$

[证明] 先设 $S=S_0(A)$. 注意在条件 $k_A(t)=k$ 的约束下,序列 C 的信号 $C(t)$都是从 SR_2 的第 $r(k)$位输出的,即当 $k_A(t)=k$ 时有

$$C(t)=B(t+r(k))$$

因此，当 $k_A(t)=k$ 时，视 $l=r(k)$ 与否，序列 C 与 B_1 相应信号分别出自 SR_2 同一寄存单元或两个固定的不同单元. 这就意味着当 $S=S_0(A)$ 时我们有

$$\rho_{S,k}(B_l,C)=\begin{cases}1 & \text{如果 } l=r(k)\\ \dfrac{1}{2} & \text{如果 } l\neq r(k)\end{cases}$$

现设 $S\neq S_0(A)$. 将 SR_1 自状态 $S_0(A)$ 和 S 出发进行工作所产生的两个主选序列分别记作 A 和 A'. 由于

$$\text{Prob}(a(t)=a'(t))=\frac{1}{2}$$

我们可以把约束条件 $k_A'(t)=k$ 和 $r(k_A(t))=l$ 看作是彼此独立的，而计算相关数 $\rho_{S,k}(B_l,C)$ 相当于在条件 $k_A'(t)=k$ 的约束下计算概率

$$\text{Prob}(b(t+r(k_A(t)))=b(t+l)$$

我们区分以下两种情况：

a) $r(k_A(t))=l$. 这时我们有

$$\text{Prob}(b(t+r(k_A(t)))=b(t+l))=1$$

b) $r(k_A(t))\neq l$. 这时我们有

$$\text{Prob}(b(t+r(k_A(t)))=b(t+l))=\frac{1}{2}$$

但上述两种情况出现的概率分别为 $\frac{1}{2^k}$ 和 $1-\frac{1}{2^k}$，所以当 $S\neq S_0(A)$ 时我们有

$$\rho_{S,k}(B_l,C)=\frac{1}{2^k}\times 1+\left(1-\frac{1}{2^k}\right)\frac{1}{2}=\frac{1}{2}+\frac{1}{2^k}$$

这里给出的证明较定理 3 的证明更不严格，好在通过相关检验的办法来确定 $S_0(A)$ 和 r 时并不需要计算这些相关数的具体数值，只要看它们是否等于 1 即可. 实际上，设 $S_0(B)$ 和 T 已经确定，那么

为了确定 $S_0(A)$，只须取定一个 $l\in T$，并对每个 $S\in V_n(F_2)$ 检查相关数

$$\rho_{S,k}(\overline{B}_l,\overline{C}),\rho_{S,l}(\overline{B}_l,\overline{C}),\cdots,\rho_{S,2^k-1}(\overline{B}_l,\overline{C})$$

中是否有一个为 1. 如然，则判定 $S=S_0(A)$，否则应舍弃该 S.

$S_0(A)$ 确定后，对每个 $l\in T$，必有一个唯一的 $k\in H$，使 $\rho_{S_0(A),k}(\overline{B}_l,\overline{C})$ 的值为 1，这时就可判定 $r(k)=l$.

应该指出，用上面的方法来确定 $S_0(B)$ 和 $S_0(A)$ 时，都必须进行搜索，但这时所需的搜索量仅是

$$Q_0=2(2^n-1)<<Q=(2^n-1)^2A_n^{2^k}$$

当我们通过某种方式将两个原始序列相结合来产生一个新的序列时，为了保证新序列的抗攻击能力，当然希望新序列的解算问题中所遇到的搜索量相当于将两个原始序列的搜索量相乘. CID 分析使本来希望的搜索量相乘退化成搜索量相加，而映射 r 所带来的搜索则几乎被完全消除. 因此我们认为，如果没有进一步的混乱措施，Jennings 论文中所讨论的复合方案，即使再加上面提到的那种钟控措施，也未必是一个好的方案. 这个方案的特点是将多路选通器这样一种现成的电子学元件用到密码序列的设计中来，但这种用法所造成的结合是可分割的，因而是无效的.

下面来看一个攻击实例. 取 $n=13, h=3$，并取 SR_1, SR_2 的联结多项式为

$$f_1(x) = f_2(x) = x^{13} + x^4 + x^3 + x + 1$$

对于如此获得的序列 C，几个有关的参数是：

a)周期：$P(C) = (2^{13}-1)^2 = 67092481$；

b)线性复杂度：$L(C) = (2^{13}-1)(1+C_{13}^1+C_{13}^2+C_{13}^3) = 3096198$；

c)穷搜量：$Q = (2^{13}-1)^2 A_{13}^8 = 3.48 \times 10^{15}$

下面是序列 C 的一个 500 比特截段 $\overline{C}$：

0110110111000010111101111011100110111100011110000
1010010111011110100100100101111001010100110011111 0
1111000011010011100000010110011011110101111001000 1
0000001001010111111110001011010010101110110001011 0
0101111010010011101101101110011001010010100011101 1
0011110111000001100100010100100011100011110001100 1
1111111011110100111010000001001001011100001000110 1
0101101100011001110110110110100000011100111110000 0
1111111110110111001110100111111000101111010001100 00
0000110001111010010101100101101010101000101010111 1

首先，通过对 SR_2 的初态 S 的一次合理组织的搜索发现，当

$$S = 1101001010101$$

时，13 个相关数 $\rho(\overline{B}_l, \overline{C})(0 \leqslant l \leqslant 12)$ 如下：

0	1	2	3	4	5	6	7	8	9	10	11	12
.522	.450	.558	.478	.518	.534	.462	.558	.534	.524	.496	.498	.566

因此可以判定映射 r 的象集 T 为

$$0,2,4,5,7,8,9,12$$

特别我们看到 0 和 12 都在集合 T 内，故可断定这个向量就是 SR_2 的初态.

再取 $k=0 \in HS$ 对 SR_1 的初态 S 进行搜索，发现当

$$S=0111001011001$$

时，相关数 $\rho_{S,0}(\overline{B}_l, \overline{C})(l \in T)$ 如下表：

l	0	2	4	5	7	8	9	12
$\rho_{S,0}(\overline{B}_l, \overline{C})$	*	*	*	*	*	*	*	1

其中 * 表示小于 1 的数. 这就说明

$$S_0(A) = S = 0111001011001$$

取定 $S_0(A)$，对不同 $k \in H$ 计算 $\rho_{S,0(A),k}(\overline{B}_l, \overline{C}), (l \in T)$，得如下数据表.

由此断定产生 $\overline{C}$ 时所用密钥为

$$r = \begin{bmatrix} 0,1,2,3,4,5,6,7 \\ 12,5,8,4,2,9,7,0 \end{bmatrix}$$

k \ l	0	2	4	5	7	8	9	12
0	*	*	*	*	*	*	*	1
1	*	*	*	1	*	*	*	*
2	*	*	*	*	*	1	*	*
3	*	*	1	*	*	*	*	*
4	*	1	*	*	*	*	*	*
5	*	*	*	*	*	*	1	*
6	*	*	*	*	1	*	*	*
7	1	*	*	*	*	*	*	*

注意，在这个例子中，复合序列的不确定度是

$$H = \log_2 Q = 51.6(\text{比特})$$

我们看到，由于 CID 带病之故，平均每个输出信号使序列的熵损失约 0.1 比特. 由此可见，即使是带钟控的复合序列也是一只漏水相当严重的“洋铁壶”.

参 考 文 献

[1] P. R. Geffe, How to Protect Data with Ciphers That Are Really hard to Break, Electronics, 1973, 4, 99～101.

[2] 许祥秦，Geffe-序列的进一步研究，西北电讯工程学院硕士论文，1984.

[3] 中国科学院研究生院 DCS 中心，线性校验方法，中国科学院研究生院 DCS 中心论文汇编.

[4] V. S. Pless, Encryption Schemes for Computer Confidentiality, IEEE Trans. Comp., 1977, 26, 1133～1136.

[5] 黄民强，对 Pless-序列的攻击，国外通信保密现状研讨会报告，1985.

[6] F. Rubin, Decrypting a Stream Cipher Based on J-K Flip-Flops, IEEE Trans. Comp., 1979, 28(7), 483～487.

[7] Jennings, S. M., A Special Class of Binary Sequences, PhD Thesis, University of London, 1980.

[8] 熊荣华，Multiplex 序列的分析与综合，北京大学硕士论文，1984.

[9] 刘木兰、万哲先，广义复合序列的分析与综合，中国科学院研究生院学报，1985, 2, 85～94.

[10] H. Beker and F. Piper, Cipher System, Northwood Books, 1982.

[11] T. Beth(ed.), Cryptography, Springer-Verlag, 1983, 17.

[12] 齐忠涛，序列密码的若干动向，国外通信保密研讨会会议录，1985, 1～6.

Correlation Immunity Deficiency and an Attack to Jennings' Scheme

Zeng Kencheng　　Lu Shuwang

(*State Key Laboratory of Information Security, Graduate School of Academia Sinica, Beijing* 100039, *PRC*)

Yang Junhui

(*Computing Center, Academia Sinica, Beijing* 100080, *PRC*)

Abstract

Making use of certain special autocorrelation characteristic of the multiplexing sequences, a statistical attack to the Jennings' scheme is proposed and realized.

偏差分析的一个例子——分析一个分组密码方案

杨君辉

（中国科学院计算中心，北京 100080）

戴宗铎　曾肯成

（中国科学院研究生院信息安全国家重点实验室，北京 100039）

[摘要]　本文运用偏差分析(differential cryptanalysis)方法，阐明加法函数的输入输出偏差特性，攻击一个分组密码算法[3].

一、引　言

偏差分析是 E. Biham，A. Shamir 提出的一种可选择明文的密码分析方法，是分析攻击 DES 型分组密码算法的有力工具. 运用这一方法，已成功地攻击了 8 层的 DES，Feal-8，GDES 等算法[1,2]. 在这篇文章中，我们运用这一方法，分析攻击在第七届全国计算机网络学术会议上提出的一个分组密码方案[3]，给出阐明偏差分析方法的一个例子. 在该方案中，即使数据和密钥长度加倍，加密层数增至 16 层等，本文提供的偏差分析方法仍然有效.

先介绍这个方案(见图 1). 这是一个变形的 DES 型分组密码算法，数据和密钥的长度都是 32 比特. 图中“移位并交换”不仅左右两组 16 比特数据交换，而且每组数据本身的高权位的 8 比特和低权位的 8 比特也交换位置. 各层的子密钥 K_1,K_2,K_3,K_4 通过特定算法由 32 位主密钥产生. 该方案中的变换函数(图 1 中 9)是简单的加法：$F_K(x)=x+K_i(\mathrm{mod}\ 2^{16})$，这里 K 是各层的子密钥. 容易看出，对任何密钥 K，该算法将有 2^{16} 个不动点，即有 2^{16} 个明文经过加密后保持不变，密文等于明文. 显然，在分析中不必考虑方案中的初始置换和逆初始置换.

首先引入一些记号. 本文中出现的变量符号 a,a^*,a',A,A^*,A',K_i 等都是 16 比特的量，可以将它们看作是在区间$[0,2^{16}-1]$中的整数. 对这些量，本文有时用十六进制表示，如 10 表示整数 16，80 表示整数 128；有时用向量$(a_{15}a_{14}\cdots a_0)$表示数 $\sum a_i 2^i$. 我们用小写字母 a,a^*,b,b^* 等表示由子密钥 K_i 控制的各层变换函数 $F_{K_i}(x)=x+K_i$ 的输入变量 x 的取值，用大写字母 A,A^*,B,B^* 等表示相应的输出量，即 $A=F_{K_i}(a)$，$A^*=F_{K_i}(a^*)$，$B=F_{k_i}(b^*)$，$B^*=F_{K_i}(B_*)$等. 用 a' 表示一对输入量 a,a^* 的偏差：$a'=a\oplus a^*$(⊕是按位模 2 加)，用 A' 表示相应输出 A,A^* 的偏差：$A'=A\oplus A^*$ 等. 我们用下图表示一对输入量 a,a^*：