# 计算机数据保护

## ——序列密码的分析与设计

Rainer A. Rueppel 著
郑秀林　杨荣焱　译
蔡吉人　校

一九八八年六月出版

## 内 容 简 介

序列密码是目前世界上广泛用于计算机和通信信息保护的一类密码。本书系统地介绍了序列密码的分析与设计方法，综合了有关最新结果。

本书可供从事计算机数据保护和保密通信工作的有关人员参考，也可供大专院校数学、计算机、通信、信息论、电子学等有关专业的师生阅读。

Rainer A. Rueppel

Analysis and Design of Stream Ciphers

Spring Verlag Berlin, Heidelberg 1986

计算机数据保护——序列密码的分析与设计

Rainer A. Rueppel 著

郑秀林　杨米森　译

蔡吉人　校

开本 787×1092 毫米　1/32

1988 年 6 月印刷

# 译 者 的 话

当今，世界已进入信息时代。数据保密与验证问题已成为现代计算机和通信系统不可分离的方面。在数据的存贮和传送过程中，为了防止数据被非法窃取、伪造和篡改，采用密码方法已成为最有效、最普遍的手段之一。

密码方法一般分为分组密码和序列密码，对它们的分析方法不尽相同。在过去的一些密码学著作或刊物中，对分组密码讨论较多。*Rainer A. Rueppel* 是一位颇有才华的年轻密码学家。他着重研究了序列密码，在所取得的一系列结果的基础上，综合了序列密码研究方面的最新成果，写成了"*Analysis and Design of Stream Ciphers*"这本书。书中对于序列密码分析和设计中的两个重要指标，即序列的线性复杂度和组合函数的相关免疫性作了比较完整的阐述。在此基础上，设计了许多序列密码算法，并对它们的保密性作了系统的分析。内容新颖，自成系统，是计算机和通信信息保护方面的一部佳作。

由于众所周知的原因，目前国内这方面的书籍寥若晨星，求之难得。我们特译出此书，奉献给诸位读者。为从事计算机数据保护和保密通信工作的有关人员，以及大专院校数学、计算机、通信、信息论、电子学等有关专业的师生提供一本有价值的参考书，为我国密码科学的繁荣尽绵薄之力。

我们将书名略有改动。此外，删去了原书中"致谢"部分

和"图表目录",并对所发现的一些错误(大部分是排印错误)作了更正,但未一一注明,只在少数几处加了译者注。

本书共九章。前五章由杨荣兹同志译,后四章由郑秀林同志译。蔡吉人研究员审校了全部译稿。张颂军及打字室的其他同志在计算机录入方面给予了很大帮助。在此一并谨表谢忱。

由于我们的水平有限,时间仓促,译文中难免有缺点甚至错误,恳请读者批评指正。

译 者

1988年6月于北京

# 序　言

自 W. 迪菲和 M. E. 赫尔曼一鸣惊人的论文《密码学的新方向》问世以来，至今已整整十年了。这篇论文不仅开辟了公开密钥密码学的新领域，而且也引起了秘密密钥密码学方面的兴趣，这一领域曾经几乎是保密机构及数学业余爱好者的专门领域。自一九七六年以来，已出现了一些优秀的有关密码学学科方面的书。这些书基本上都是论述公开密钥体制及分组密码(即在加密变换中不具有记忆的秘密密钥密码)的，而对序列密码(即在加密变换中具有记忆的秘密密钥密码)却很少关注。然而，序列密码(诸如转轮机器所实现的)在过去的密码实践中曾起过主导作用。据我断定，序列密码今后仍然是商业、军事、外交上主要使用的保密体制。

我自己在序列密码方面的研究兴趣，引起了我的在苏黎世的瑞士联邦理工学院的博士生 Rainer A. Rueppel 的共鸣。当 Rainer 于 1984 年下半年完成他的学位论文时，出现了这样一个问题，即在什么场合下发表自己关于序列密码的许多新结果。因为他的工作完全是基础性的，涉猎的范围很广，而且在论文中叙述得也非常清晰，因此，他不愿意将自己的工作按通常的习惯零碎地发表。鉴于这种情况，我要求 Rainer 为这套丛书写一本书，将论文予以扩展，对序列密码进行充分的阐述。简单地说，我要求他在这本书中注重论述序列密码，对其重要价值予以充分关注，弥补这方面的不足。

Rainr欣然接受了这一任务，写成此书，书中尽可能包括了许多新结果，使序列密码的研究自成一体。

本书是斯普林格(Springer)丛书中关于密码学的第一本书，其余的正在进行中。我希望这套丛书对于密码学方面的重要新论文能够起到抛砖引玉的作用。Rainer的书已为它们设立了一个高的标准。

James L. Massey

# 目　录

# 第一章　绪　论

当今，世界已进入信息时代，数据保密与验证问题已成为现代计算机和通信系统不可分离的一个方面。越来越多的数据用电子方法存贮和传送，因之数据也易于公开暴露，通信信息有可能被非法地泄漏和篡改，个人和团体的秘密完全依赖于严密的通信信息保护。

密码体制是运用变换来提供保密性的。在发送地点，明文消息在加密密钥控制下变换为密文，它对于没有掌握秘密密钥的任何对手来讲是难懂的．而对于合法的接收者，密文在秘密的脱密密钥控制下变换为原来的明文。密码体制一般分为分组密码和序列密码两大类。

分组密码通常是按固定规模将明文分组，对每组独立地进行运算。因此，分组密码是一种简单的代替密码，它必须具有大的字母表，以阻止穷举搜索的密码分析。

序列密码将明文分为字符，并使用一个随时间变化的函数加密每个字符，该函数的时间相关性是由序列密码的内部状态控制的。就是说，在每个字符被加密之后，装置按照某种规则改变状态。因此，当同一明文字符两次出现时，一般不会得到相同的密文字符。

由此可见，分组密码和序列密码可看作分别具有一个或多个内部状态的密码体制。这种分类法与纠错码分为分组码和卷积码相似。

在同步序列密码中，下一状态只依赖于前一状态，而与输入无关。从而，状态的改变不依赖于接收的字符序列。这种加密是无记忆的，但随时间变化。对应于一个特殊的输入，其输出只依赖于在它之前或在它之后的那些字符。

在所有密码中最出色的密码之一是一次一密乱码本(One－Time－Pad)。在这里，明文消息与一个同样长度的无重复的随机序列逐位相加。弗纳姆(G. Vernam)于1917年将这一原理用于电报通信〔Kahn 67〕。为了纪念他，也把一次一密乱码本称为弗纳姆密码。关于一次一密乱码本的一个值得注意的事实是它具有完全保密性。在假定唯密文攻击情形下，香农〔Shan 49〕证明了：即使是具有无限计算资源的密码分析者，也决不能从所有其它有意义的明文中识别出真正的明文。当然，该体制的缺点是需要无限的密钥量。

受一次一密乱码本这一富有吸引力的特点的启发，我们可构造使用伪随机序列加密明文的同步序列密码，从而取消了无限密钥量的必要条件。这种伪随机序列，是在一个秘密的密钥控制下，由一个称之为密钥序列发生器的确定性算法所产生的。在同步序列密码情形下，假定明文已知，则相当于密码分析者可得到密钥序列。对于保密的体制来讲，密码分析者要想预测密钥序列的任何部分，只相当于随意猜测，不论已观测了多少密钥序列字符，对后续字符的预测都无济于事。然而，由线性自动机产生的序列是容易预测的，这是密码编制上的弱点，因而必须结合非线性变换，以得到所期望的高复杂度的伪随机序列。

本书重点讨论同步序列密码的非线性体制，尤其是在如何生成高度不可预测的伪随机序列方面。

第二章讨论序列密码加密原理，并定义密钥序列发生器。

将生成的密钥序列的线性复杂度视为一个具有头等重要性的设计参数。非线性变换只限于密钥序列发生器的前馈部份，以保证它的可分析性；并导出了非线性前馈部分的若干性质。

第三章扼要地介绍了代数工具，以用于线性反馈移位寄存器非线性组合的分析。除了基本的定义之外，这一章也含有扩域中算术运算的独立的重要结果，以及未被普遍采用的然而在周期序列的非线性理论中使用方便的定义。

在第四章中，证明了线性复杂度作为有限序列不可预测性(或相当于随机性)测度的重要性质。根据相伴的线性复杂度的期望与方差，得到了长度为 n 的二元序列的特征。随着序列长度 n 的增加，线性复杂度的增长可用随机游动来描述。"典型的"随机序列显示出"典型的"复杂度轮廓。对于周期地重复一个有限随机序列这种有实际意义的情形，计算了它的期望线性复杂度。本章的目的在于：导出评价确定地生成的随机序列不可预测性的准则。

第五章的目的在于：根据生成的输出序列的线性复杂度来分析非线性组合提供一种理论基础，并将 GF(2) 上函数的代数标准型选作标准。对于最大长度的序列(m 序列)不同相位上的非线性运算，给出了一种矩阵方法，它提供了完整的可分析性。然而，对于有实际意义的情形，这种方法一般在计算上是不可行的。Key 导出了生成序列线性复杂度的一个上界，它依赖于所使用函数的非线性阶以及 m 序列递归式的阶；证明了所选择的某个函数，其相伴的输出序列的线性复杂度远远小于此上界的概率随着 m 序列递归式的素数阶增大而趋向于零。有一大类函数，其生成序列的线性复杂度下界是 $C_L^k$，这里L 表示 m 序列递归式的素数阶，k 表示函数的

非线性阶．对于具有不同的(但未必是不可约的)极小多项式的序列的非线性组合运算的情况，证明了所生成的输出序列的线性复杂度等于将非线性组合函数的序列自变量代之以相伴线性复杂度之后，所计算的该函数的实数值．这里，要求每个极小多项式的根都是单根，并都在某一扩域中，该扩域的次数与含有别的极小多项式的每个根的扩域的次数互素．

第六章讨论了 $LFSR_s$ 的计时速率高于系统的计时速率的作用，并说明了这种多倍计时将导致一个不同的 LFSR，它可模拟以物理方式实现的 LFSR．模拟的 LFSR 可由原始 LFSR 及相伴的速度因子完全确定；对由一个线性密码问题提出的使用多倍计时的随机序列发生器进行了详细分析．

在第七章中，对 0/1 背包及有关问题在密码学中的可用性进行了讨论．对于整数加法，如果将整数及它们的和表为二进制形式，则在 GF(2)上也可完全地计算．生成和的第 i 位的函数的非线性阶可用数量精确地表示，并已证明它随 i 按指数律迅速增长．通过将背包重量参数化，并将背包看作二元 N－ 数组的集合到整数集合上的函数，可得到背包的完全的 GF(2)描述．

在第八章中，为序列密码体制提出了一种流密钥发生器．在那里，一个背包作用于最大长度线性反馈移位寄存器的状态上．对由背包的个别和的二进位所定义的序列进行了分析，并证明了这种序列具有密码上的重要性质．本章还包括了所提出的背包序列密码的某些小规模模拟．

第九章讨论在非线性组合函数中允许具有记忆的作用．于是，组合器本身变成一个有限状态机(FSM)．证明了这种 FSM－ 组合器容易避免相关性攻击．而许多新近提出的密钥序列发生器都屈服于这种威协；还给出了 r 元序列的

实加法，以定义这种有记忆的抗相关结构；证明了递归式的阶互素的两个 m 序列在实数范围内相加时，所得到的和序列的线性复杂度一般等于或者很接近于输入序列的周期之积。

书中有许多用图说明的例子，以便于概念的理解和接受。它们是尽可能根据实践和现实情况而提出的。

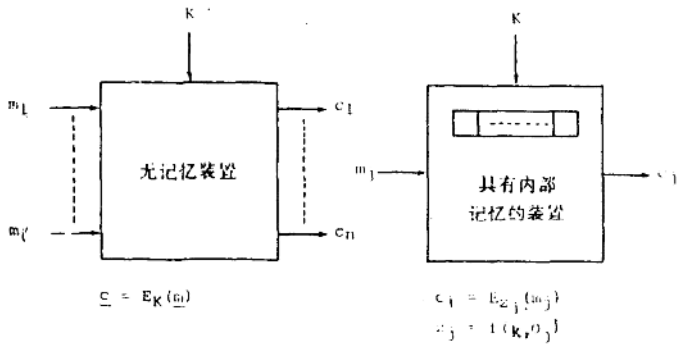
# 第二章 序列密码

出于现实的考虑(比如延迟或者处理方便),通常将半无限的消息序列细分为固定规模的实体顺次地加密。存在两种基本不同的但能够合理地完成加密的方法。当加密变换独立地作用于每一个这样的消息实体上时,这就是分组密码。为了阻止强力的密码分析,它必须具有足够大的字母表。"组"这个示意性的名字,标志着一个消息实体的规模。相反,序列密码则用一个随时间变化的函数加密每个消息实体,这个函数的时间相关性是由序列密码的内部状态控制的。根据这种加密原理,它的消息实体不需要大,并以"字符"来标志一个消息实体的规模。每个字符被加密之后,序列密码按照某种规则改变状态。因此,当出现两个相同的明文字符时,一般不会产生相同的密文字符。

如果运用记忆作为判断的准则,可以看出分组密码与序列密码之间的明显区别(见图 2.1)。

分组密码确定一个无记忆装置,它在密钥 K 的控制下,将消息组 $\underline{m}=(m_1,m_2,\cdots,m_L)$ 变换为密文组 $\underline{c}=(c_1,c_2,\cdots,c_n)$。通常,其消息报文字母表与密文字母表相同。而序列密码确定一个内部记忆装置,即用一个依赖于秘密密钥 K 以及序列密码在时刻 j 的内部状态 $\sigma_j$ 的函数,将消息序列的第 j 个数字 $m_j$ 变换成密文序列的第 j 个数字 $c_j$。

按照这种方法,将密码体制分为分组密码与序列密码,这




(a) 分组密码　　(b) 序列密码

图 2.1　两种基本的加密原理

同纠错码分成分组码与卷积码相似。对一个固定的密钥 K，分组密码将相同的消息组变换成相同的密文组。这就允许主动的窃听者插入、删除或者重放预先录制的密文组(除非使用了控制消息组顺序的某种附加的通信规程)。此外，分组加密还允许被动的窃听者通过搜索密文进行匹配。当把固定的消息实体(比如薪水记录等)作为明文分组处理时，就是一个严重的威胁。相反地，由于序列密码是在一个密钥的随时间变化的函数作用下加密每个字符，故它可以防止密文的删除、插入或者重放，并防止密文搜索。由于利用记忆提供了附加的变化因素，因此，我们可以说序列密码比分组密码有更好的内在保密性。但是应当指出，一个分组密码总可以通过结合

记忆增强为序列密码的。此时，分组密码只作为非线性函数使用。就目前最出众的分组密码——数据加密标准(DES)而言，在变换中引入记忆以提高其保密性方面，有许多不同的建议(参看〔Denn 83〕的一个简要评述)。序列密码可进一步分为同步体制与自同步体制。在同步序列密码中，下一状态只依赖于前一状态，而不依赖于输入。从而状态的后继与接受的字符序列无关，故加密变换是无记忆的，但随时间变化。然而装置本身不是无记忆的，它依靠内部记忆以产生必要的状态序列。所以，在同步序列密码中，把加密变换与控制加密变换的随时间变化参数的产生过程分开是很自然的事情(见图2.2)。

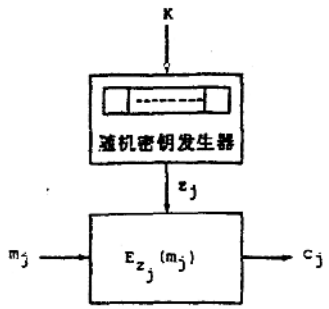



图 2.2 一种分解的同步序列密码

用于控制加密的序列 $\tilde{z}=z_0, z_1, \cdots$，称为密钥序列或者流密钥。根据现行的密钥 K 与内部状态来产生密钥序列的这

种确定性自动机，称为**流密钥发生器**．当发方与收方的密钥K以及装置的内部状态相同时，流密钥也必然相同，因而脱密是容易实现的，这时我们说流密钥发生器在发方和收方相互同步．一旦流密钥发生器失步，则不能完成脱密，这时必须给装置重新建立同步．

相反，在自同步序列密码中，脱密变换具有有限的记忆．由于受过去字符的影响，如有一个密文字符的错误或者丢失，对已脱密的明文只引起固定数目的错误，在此之后仍可产生正确的明文．

由于序列密码具有内部记忆并运用了非线性变换，故序列密码一般是难于分析的．在编码理论中也可以见到类似的情况，即研究分组码的刊物占绝大多数．Widman教授于1984年在苏黎世大学数字通信讨论班上指出：对于序列密码的重要性在公开文献中没有得到适当反映，在关于密码体制的刊物中，涉及序列密码的只占很小的百分比．他认为，序列密码的应用将具有广阔的前景．

## 2.1 理论保密与实际保密

在所有密码中最值得重视的密码之一是一次一密乱码本（亦称弗纳姆密码〔Kahn 67〕）．在这种体制中，密文是由明文消息与具有同样长度的不重复的随机序列作逐位模2加得到．由于在GF(2)上加法与减法相同，故脱密可通过再一次加密来完成．图2.3具体说明了一次一密乱码本．

一次一密乱码本的基本原理是：运用真正随机的密钥序列消除密文与明文之间在统计上的差异．产生这种随机序列（即每个比特是等可能的、且与前面比特无关的0或1的序