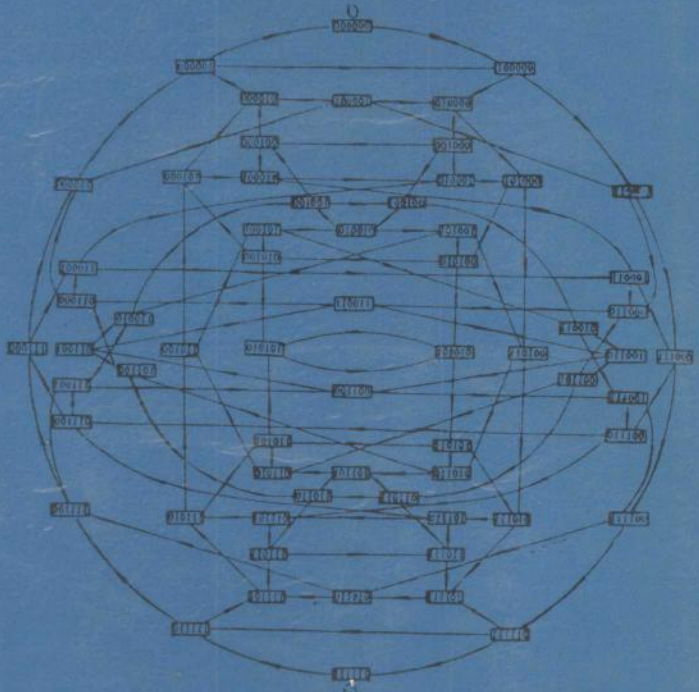
第三次全国密码学会议会议录

中国电子学会信息论学会



西安电子科技大学情报资料室出版

一九八八年十二月·西安

第三次全国密码学会议会议录

中国电子学会信息论学会

西安电子科技大学情报资料室出版
一九八八年十二月

前　言

本会议录收录了在第三次全国密码学会议上宣读的论文或论文摘要。这次会议是由中国电子学会信息论学会委托西安电子科技大学负责筹办的，会议于1988年12月13日至16日在西安举行。有关密码的前两次会议于1984年10月26日至29日和1986年8月1日至5日分别在西安和浙江建德举行。由于条件关系，前两次会议的论文都未能编辑成会议录出版。根据广大与会者的要求，本次会议我们提前编印了会议录以便于交流。我们希望今后能坚持在会前印好会议录。

七十年代以来，密码学和数据安全保密问题在国际上受到了越来越广泛的重视，国内自八十年代初也开始重视起来，密码学的研究队伍在不断扩大，许多充满朝气的年青力量加入了这个队伍。就这次会议收到的论文来看，无论从深度和广度都远远超过了前两次会议。特别是论文的作者大都是年青的同志们，这是十分可喜的现象，我国密码学研究寄希望于他们。

本会议录从收到的论文中选收63篇，包括的内容有以下几个部分：

1. 单钥密码体制，包括各种移位寄存器序列理论、有限自动机密码、布尔函数 、Walsh 谱论等共计　31篇
2. 双钥或公钥密码体制　8篇
3. 数据加密标准　7篇
4. 模拟语音加密　6篇
5. 计算机安全　6篇
6. 硬件芯片实现　4篇
7. 其他　1篇

这次会议收到的论文虽较前两次丰富而深刻，但距密码学的发展还有差距，有些重要理论课题，如零知识证明、概率密码、数据网的安全协议等课题涉及很少或无人问津，又如图象信号加密、密钥分配管理、EFT、EMS和OA中的安全保密等实际问题也很少涉及，当然并非国内根本无人搞这方面的研究。这反映了本次会议在组织宣传方面的工作也做得不够。

本会议录所以能在会前付印，首先应当感谢各位作者的大力合作。在会议录编辑出版的整个过程中，西安电子科技大学情报资料室的同志们付出了辛勤的劳动，没有他们的紧密配合，本会议录是不可能在会前印好的。

本会议录的编辑出版工作中还有很多不足之处，希与会者和读者们提出意见，为今后的会议录的出版工作做得更好而共同努力！

第三次全国密码学会议筹备组

一九八八年十二月

目　录

一、单钥密码体制

二、双钥密码体制

三、数据加密标准

四、模拟语言加密

五、计算机安全

六、硬件芯片实现

七、其　它

介绍“叵测度”设想

Introduction to the Conception of Impredictability

高鸿勋
Gao Hongxun
（南开大学数学系）

【摘要】 通过关于已知序列的几种复杂性度量的回顾，观查到这些度量的内在缺陷。因而给出了一个新设想，叵测性度量的活动或动态分析。此后推导了叵测性度量的数学模型。这种活动分析与有关序列集合的某些分类密切相关。并在M序列集合的某些情况下考虑了集合分类的一些理论问题。

关键词：复杂度，叵测度，本原M序列，剪接法。

Abstract, By a general review of several measurements of complexity for a given sequence, it appears an intrinsic deficiency in these measurements. Thus, a new conception, the activity or dynamic analysis of measurement of impredictability is given. Then, a mathematical model of the measurements of impredictability is derived. This activity analysis is tightly related to some of the classifications of the regarded sets of sequences. Some theoretical considerations under the sets of M-sequences, for several cases, are considered.

Key words, Complexity, Impredictability, Primitive M-sequence, Cut-join Method.

一、关于复杂性的度量

设s是一个在有限集上的有限长或周期序列，如何估量s“复杂”程度，时常是一个很重要的实际问题。但从人们的不同观点出发往往得到不同的结果，比较一致的表示形式是：

$$s\text{的复杂度} \quad C(s)=C_A(s)$$

其中A是估算s的方法或手段。

下面首先回顾两种重要的有代表性的复杂度类型，它们是分别按照s的计算量与信息量规定的，所谓线性复杂度与信息熵复杂度。

［1，2，3］等给出了线性复杂度的定义与算法。它等于能产生s的等效线性移存器的最小级数；也是s的联接（极小）多项式的次数；还可以证明，它又等于s所生成的循环方阵的秩。这样，线性复杂度由于其线性性的规律简单的优点而受到重视。

［4，5，6］等研究了n级2元M序列线性复杂度的分布情况，证明了：

$2^{n-1}+n \leqslant C(s) \leqslant 2^n-1$，但$C(s) \neq 2^{n-1}+n+1$，对本原M序列来说$C(s)=2^n-1$，当$n \geqslant 3$时有s存在使$C(s)=2^{n-1}+n$。

但有些看来并不复杂的序列，如本原M序列，(1 0 … … 0)等都有很高的线性复杂度。若将本原M序列去掉一个0化为m序列，则其线性复杂度就骤降为n。

关于信息熵复杂度，在[7]中 Lempel 等人给出将序列s“穷举史”的分史个数定义为s的信息熵复杂度，可以证明其中较复杂的“双层”定义可以等价的更简捷地定义为“空前分史”的个数。对于n级2元M序列来说当n足够大时，M序列的复杂度均接近其上限。这点与线性复杂度的结论相符，但它也有与线性复杂度相类似的缺陷。

正如[6]中指出，C(s)高的序列其可预测性不一定低，但C(s)低的序列其可预测性一定高。从而限制并降低了用线性复杂度与信息熵复杂度检验序列“复杂”程度的有效性。对于其它由计算与信息量的大小所规定的复杂度来说，由于其内部缺陷，一般来说也有类似的问题。

二、回测性度量的初步设想

从另一观点出发是估量欲测序列s的不可(难)预测性(impredictability 或 unpredictability)简称回测性。

假如已知q元域上序列s的周期是T，而对其它特征则一无所知。如果将s的可能范围的势规定为s的回测度，则记为：

$$I_r(s)=q^T \quad 或 \quad U_r(s)=q^T$$

当$T \to \infty$时，$q^T \to \infty$，称s是完全回测的。假如又知s的某些特征，则按照这些特征将s的“范围”分类后，s的回测度就会降低，因此回测度除依赖于A、s之外还需依赖于对s特征的了解，这样回测性又是独立于复杂性之外的动态性的概念，而记为：

$$I_r(s)=I_r(s,t) \quad 或 \quad U_r(s)=U_r(s,t)$$

其中t表时间。

例如在n级2元M序列中任取其一，则其回测度为：

$$2^{2^{n-1}-n}$$

若又知s为本原M序列，则这个“额外信息”的获得使s的回测度下降为

$$\frac{\phi(2^n-1)}{n}$$

若将这个度量的下降现象（求$\log_2$后)与§1的相应下降现象相对照，则将

$$(2^{n-1}-n) \searrow \log_2 \frac{\phi(2^n-1)}{n} < (n-\log_2 n),$$

改换为

$$2^n-1 \searrow n$$

就相差不多较为合理了。这个例子也刻划了回测度的动态性质。

对于任选有关范围(回测度)内的M序列，我们可以提供三种算法：

①据升级演算法，通过归纳法的形式按照字典式排列法列出n级2元M序列，从而任求其一。

②在q元域上反求具有

$$((q-1)!)^{q^{n-1}} \text{ 个}$$

M置换剪接的部分集合的任一个τ。这批M剪接的排列可依其组成的初等组合性质，按字典式排列法排列之。然后再算出其中任一个为（τM_0）的M序列。

③应用2次M剪接的充要条件［10］可推导出一些结论、定理，据此将算号集：

$$R_n = \{1, 2, \cdots, 2^{n-1}-2\}$$

中的诸偶元子集到n级2元M序列集合Ω_n上的一个满射（多一映射），即可唯一地确定一个n级2元M序列——任选的M序列。

三、回测度的数学模型与分析

序列回测度的基本着眼点是对于预测序列范围的估量。而在不同的阶段这种“范围”也是不同的，所以随时间t的发展，所预测的可能的序列范围也在相应的变化。用数学的语言来说“预测范围”由一组（族）等式与不等式所构成的“约束条件”所界定。其中的系数或某些参数是t的函数。

设时间t所界定的预测范围是Ω_t，当$t=0$时记$\Omega_0=\Omega$

将初始范围Ω内的各序列按计算量（大体上）区分为：

$$\mathfrak{Q}_1, \cdots, \mathfrak{Q}_k \quad \text{等k个子集}$$

设$\mathfrak{Q}_i$内序列的计算量为Ω内序列平均计算量的P_i倍。则据§2的记号可得：

$$I_p(s, t) = \sum_{i=1}^{k} P_i |\mathfrak{Q}_i \cap \Omega_t|$$

其中将Ω内序列的平均计算量算作1。

或

$$U_p(s, t) = \frac{1}{|\Omega|} \sum_{i=1}^{k} P_i |\mathfrak{Q}_i \cap \Omega_t|,$$

其中将Ω中序列的总计算量算作1。

由于我们假定了t的有限与离散性，不仿设$t=0, 1, \cdots, T$。

则

$$I_p(s, T) = \sum_{i=1}^{k} P_i |\mathfrak{Q}_i \cap \Omega_t|$$

就是一个表示现有的侦破（能力或设备）的数量限额了。

由上述的数字模型结合［12，13，14］中的论述可将回测度的要点综述如下：

①预测序列的范围与估计——范围愈广，回测度愈高。

②预测范围的动态或活动分析（Dynamic or Activity analysis），而后者可能更贴切些，因为它涉及到双方人的活动与作为。显然②是最重要也是最复杂的一个方面。

③计算量的估计。当Ω中序列的计算量较均匀的情况下上述k个子集的区分是不必要的；但一般来说它们是不均匀的，前者仅为特例。另外，计算量的估计是在给定的情况下取最佳

计算方案所需的计算量。

四、叵测度与序列集合的分类

在叵测性度量的论述中我们曾两次遇到序列集合Ω的分类问题，可见叵测度与序列集合Ω的分类是密切相关的。一种分类是由约束条件组所对应的序列范围经适当组合而成的，如前所述约束条件组是一个很复杂的问题，它们可能组成Ω的形形色色的分类；另一种是按序列的计算量所形成的分类。对于后者我们指出在具体化了的情况下所遇到的一些理论问题。

规定两个n级二元M序列α，β间的M距离，记为$M(\alpha,\beta)$，它表示α，β间的最小剪接次数。已经证明[8，9]当α，β均为本原M序列时，其M距离恒等于2^{n-3}（$n\geq 4$）。因而定义

$$m_\alpha=\min_{(\gamma\text{本原})} M(\alpha,\gamma)$$

为M序列α的极距。M距离与极距具有某些典型的基本性质，其中关于极距的上界，若据[10，11]中的标号表示与推理可以推出：

当$n\geq 4$时，该上界$\leq 2^{n-2}-2$。

进一步的类似推理还可以得出：

该上界$\leq 3\times 2^{n-4}-1$。

自然，这个上界仍有改进的余地。

以上所提到的那些性质对于第二种M序列集合的分类是直接相关联的。

参 考 文 献

[1] E.J.Groth, IEEE Trans. IT—17, 1971, 283-296.

[2] E.J.Key, IEEE Trans. IT—22, 1976, 732—736.

[3] J.L.Massey, IEEE Trans. IT—15(1969), 122—127.

[4] A.H.Chan, R.A.Games, E. l. Key, J.Combin. Theory, A33, 233—246(1982).

[5] R.A.Games, J.Combin. Theory, A34, 248—251(1983).

[6] Tuvietzion, A. Lempel, IEEE Trans. IT—30(1984), 705—709.

[7] A. Lempel and J. ziv, IEEE Trans, IT—22(1976), 75—81.

[8] 王洪涛：M序列剪接的几点性质，1987。

[9] 康庆德：组合数学讲义。

[10] 高鸿勋：求全部n级M序列及其反馈函数的一个方法与证明，《应用数学学报》，2：4，316—324，1979。

[11] ——：非奇函数是M序列反函数的一个充要条件，《应用数学学报》，7：1，1984。

[12] 高鸿勋：关于复杂性与叵测性的度量，1987。

[13] ——：叵测度的数学模型与序列集合的分类，1988。

[14] ——：关于叵测度的几个注。

流密码的安全性新度量指标

New Measure Index on the Security of Stream Ciphers

丁存生　肖国镇
Ding Cunsheng Xiao Guozhen
（西安电子科技大学应用数学系）

单炜娟
Shan Weijuan
（华北电力学院）

〔摘要〕 本文给出了一类流密码的破译方法，对流密码的安全性给出了新度量指标。为了度量序列线性复杂 度的稳定性，本文引入了“重量复杂度或球面复杂度”和“u—球体复杂度”的概念，并给出了NLFSR序列和非线性滤波ML—序列的重量复杂度上界。最后对各种流密码的安全性从线性逼近的角度进行了分析。

关键词：线性复杂度稳定性，重量复杂度，球面复杂度，u—球复杂度。

Abstract: A deciphering method for a class of stream ciphers, new Points of View on stream ciphers and a measure index on the securty of stream ciphers are presented. In order to measure the stability of linear complexity of a sequence, 'weight complexity or sphere surface Complexity' and 'u-sphere complexity' are also introduced. Upper bounds on the weight complexity of NLFSR sequence and of nonlinear filtered ML-Sequence are given. The security of many kinds of stream ciphers is analysed from the approximation Point.

Key Words: Stability of linear complexity, Weight complexity, Sphere surface complexity, u-Sphere complexity.

引　　言

流密码主要有以下两种 类型及它们的变型。

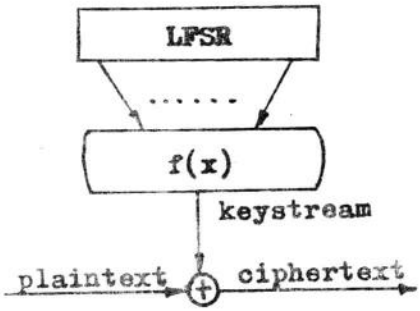


图 1　流密码系统 I

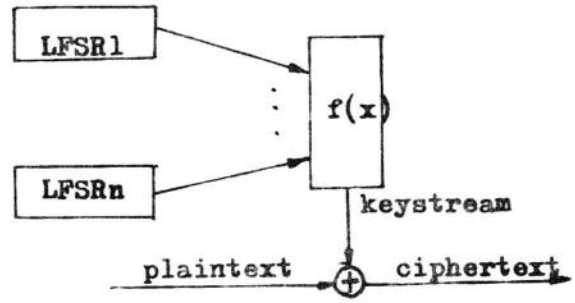


图 2　流密码系统 II

众所周知，以上两个类型的流密码的安全性主要在于种子密钥的安全性。Sieg-enthaler在文献〔1〕中对体制 II 提出了一种攻击方法，即逐步确定单个LFSR i 种子密钥的方法。为了对付这种攻击方法，他引入了组合函数 $f(x)$ 相关免疫的概念，并证明了 $f(x)$ 的相关免疫阶数和它的次数之间有一个限制关系。肖和Massey给出了相关免疫函数的一个谱特征〔2〕。

从本文作者已知的有关流密码的文献来看，上述两种流密码的安全性度量指标主要有三个：① $f(x)$ 平衡的程度；② $f(x)$ 的相关免疫程度；③流密钥序列的线性复杂度。我们的观点之一是，上述三个指标并不能保证体制 I 和 II 的安全性；另一个观点是，带有相关免疫性的函数未必是"好"的滤波函数。

本文的主要思想是，虽然一个序列的线性复杂度可以很高，我们可以用一个复杂度很低的序列来逼近它且使二者的复合率很高。所以线性复杂度高的序列可能会很"坏"。我们的观点是，只有线性复杂度较高且线性复杂度稳定性较好的序列才是好的流密钥序列。

内　容　目　录

(1) Walsh变换的基本性质和两种Walsh变换之间的关系。

(2) 布尔函数的最佳仿射逼近。

(3) 一类流密码的破译。

(4) 带有一定相关免疫性的函数未必是"好"的非线性组合函数——新的安全性指标。

(5) 序列的新复杂度——重量复杂度或球面复杂度和球体复杂度。

① 重量复杂度的基本性质。　② N LFSR序列重量复杂度的上界。

③ 非线性滤波ML序列的重量复杂度上界。

(6) 带有记忆的FSM流密码的安全性分析。

(1) Walsh变换的基本性质和两种Walsh变换之间的关系

假定$f(x):GF(2)^n \to GF(2)$是一个布尔函数。两种Walsh变换定义为[4]：

$$S_f(w)=2^{-n}\sum_{X\in GF(2)^n} f(x)(-1)^{wx}$$

$$f(x)=\sum_{w\in GF(2)^n} S_f(w)(-1)^{wx}$$

$$S_f(w)=2^{-n}\sum_{X\in GF(2)^n}(-1)^{wx\oplus f(x)}$$

$$f(x)=1-2\sum_{W\in GF(2)^n} S_{(f)}(w)(-1)^{wx}$$

这里$WX=W_1X_1\oplus W_2X_2\oplus\cdots\oplus W_nX_n$。

基本性质：

1) $\sum_W S_{(f)}(W)=2^n S_{(f)}(0)$

2) $\sum_W S_f(W)=f(0)$

定理1：（两种Walsh变换之间的关系）

$$S_{(f)}(\omega)=\begin{cases}-2S_f(W), & W\neq 0\\ 1-2S_f(W), & W=0\end{cases}$$

(2) 布尔函数的最佳仿射逼近

布尔函数的最佳仿射逼近在逻辑设计中有广泛的应用（Moraga[4]），但在密码学中直到现在几乎没有多少应用（Rueppl[8]）。在这篇文章中我们将探讨它在流密码中的应用。

定义1：如果仿射函数$WX\oplus V$使得下述表达式达到它的极小值，则称$WX\oplus V$是$f(x)$的最佳仿射逼近。

$$\sum_{X\in GF(2)^n}(f(x)\oplus WX\oplus V),\quad W\in GF(2)^n,\quad V\in GF(2)$$

引理1：命$P(WX\oplus V)$表示$f(x)$和$WX\oplus V$的复合率，则有

$$P(WX)=\frac{1}{2}+\frac{1}{2}S_{(f)}(W),\quad W\in GF(2)^n$$

$$P(WX)=\begin{cases}\frac{1}{2}-\frac{1}{4}S_f(W), & W\neq 0\\ 1-S_f(W), & W=0\end{cases}$$

从引理1易知下述定理2成立

定理2：假定$a=\max\{|S_{(f)}(W)|:W\in GF(2)^n\}$且$|S_{(f)}(W)|=a$。则

(1) 如果$S_{(f)}(W)\geq 0$，则$f(x)$的最佳仿射逼近为WX且

$$P(WX)=\frac{1}{2}+\frac{1}{2}a$$

(2) 如果$S_{(f)}(W)<0$，则最佳仿射逼近是$1\oplus WX$且

$$P(1\oplus WX)=\frac{1}{2}\mp\frac{1}{2}a$$

同样，我们可以用第一种Walsh变换来描述函数的最佳仿射逼近。

注记：一个函数的最佳仿射逼近不是唯一的。在逻辑设计和密码学中我们所需要的是具有最小Hamming重量$W_H(W)$的仿射函数$WX\oplus V$。

(3) 一类流密码的破译

在引言中我们断言，上面提到的三个指标不能保证流密码体制Ⅰ和Ⅱ的安全性。现在我们来讨论在一定的条件下如何攻击它们。

引理2：(T. Herlestam)设S_1，S_2，…，S_t是t个周期序列，则有

$$L(S_1\oplus\cdots\oplus S_t)\leq L(S_1)+\cdots+L(S_t)$$

特别地，如果S_1，…，S_t是同一个序列的t个不同相应序列，则

$$L(S_i\oplus\cdots\oplus S_t)=L(S_i),\quad 1\leq i\leq t$$

引理3：命$S=S_1\ S_2\cdots S_n$是一个二元序列，则序列$\bar{S}=\bar{S}_1\ \bar{S}_2\cdots\bar{S}_n$的线性复杂度满足

$$L(S)-1\leq L(\bar{S})\leq L(S)+1$$

其中$\bar{S}_i=1\oplus S_i$。

假定我们知道滤波函数，则我们可以通过应用最佳仿射逼近及已知的明文密文对来

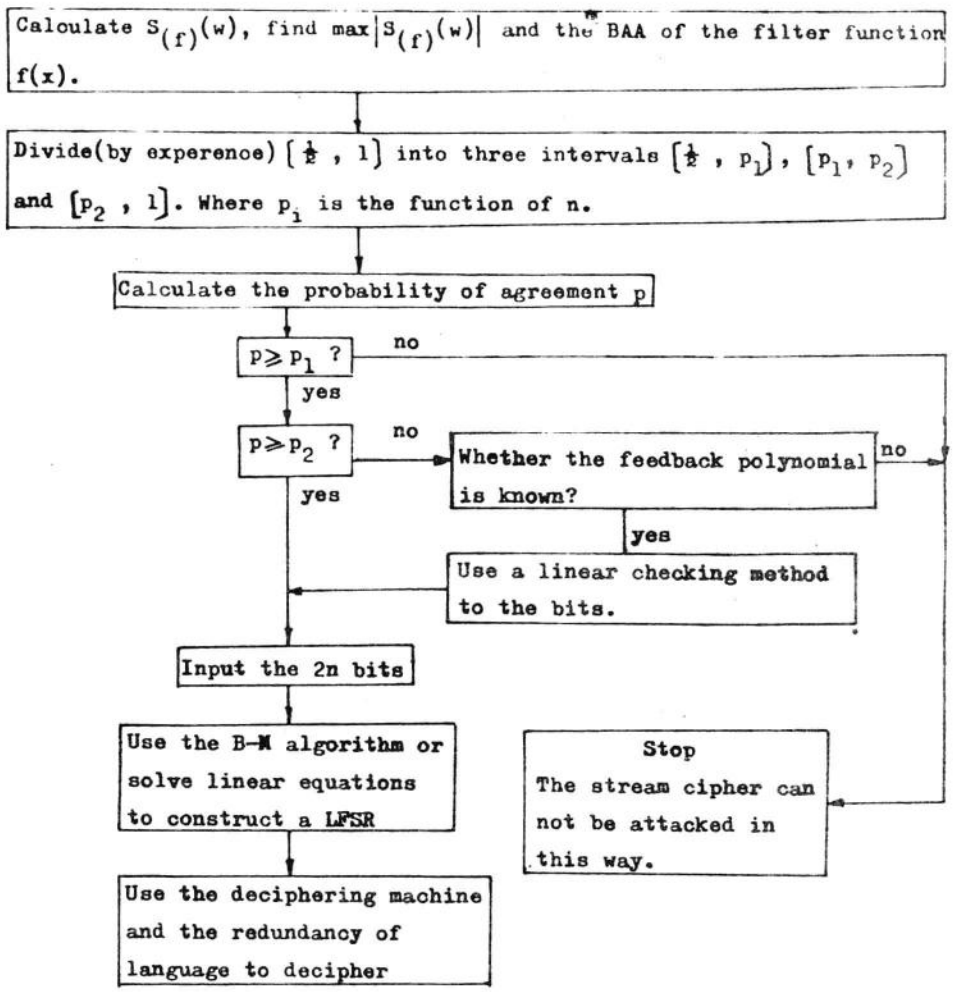


图 8　对流密码的BAA攻击的流程图

构造一个线性反馈移位寄存器使其与原流密钥序列的复合率最高，然后用该线性移位寄存器作为解密机，最后通过语言的多余度纠错。

(1) 流密码体制 I 的攻击

假定线性反馈移位寄存器的长度为 n（$n \geq L$）选择 $f(x_1, x_2, \cdots, x_L) = x_1 \oplus \cdots + X_{\frac{1}{2}L} \oplus X_{\frac{1}{2}L+1} \cdot X_{\frac{1}{2}L+2} \cdots X_L$ 作为例子，Rueppel [3] 曾证明，如果一个滤波函数形如上述 $f(x)$，则输出流密钥序列的线性复杂度满足

$$\text{线性复杂度} \geq \binom{L}{\frac{1}{2}L} \approx 2^{\frac{1}{2}}(L)^{-\frac{1}{2}} 2^{L-1} \qquad (L >> 1)$$

另一方面，$f(x)$ 是平衡的 且为 $\frac{1}{2}L - 1$ 阶相关免疫的。按照上述三个指标 $f(x)$ 是一个好滤波函数。通过计算我们得到

$$a = \max |S_{(f)}(w)| = 1 - 2^{1-\frac{1}{2}L}$$

$f(x)$ 的最佳仿射逼近为 $X_1 \oplus \cdots \oplus X_{\frac{1}{2}L}$，复合率为 $1 - 2^{-\frac{1}{2}L} > 99.8\%$ $(L \geq 12)$。

由引理 2 我们知道，如果用 $x_1 \oplus \cdots \oplus x_{\frac{1}{2}L}$ 来代替 $f(x)$，则输出流密钥序列的线性复杂度是 n。所以，如果已知 $2n$ 比特连续明密文对，假定 $f(x)$ 的非线性项对这些比特没有贡献（如果有的话，我们可以用线性校验方法来消除它们）。我们可以构造一个 LFSR 按上述方法解密。

一般地，破译过程可综合为图 3 中的流程图。

(2) 流密码体制 II 的攻击

假定 LFSR i 的长度是 n_i，同样，由引理 2 和 3 我们知道，如果已知 $2(n_1 + \cdots + n_{\frac{1}{2}L+1})$ 比特连续明密文对，我们可以用同样的方法来攻击它。

(4) 带有一定相关免疫性的函数未必是好的滤波函数——新的安全性指标

引理 4：（能量守恒定理）设 $f(x): GF(2)^n \to GF(2)$ 是一个布尔函数，则

$$\sum_{W=0}^{2^n-1} S_{(f)}(W)^2 = 1$$

定理 3：设 $f(x): GF(2)^n \to GF(2)$ 是一个布尔函数，则至少存在一个仿射函数 $WX \oplus V$ 使得复合率

$$P(WX\oplus V)\geq \frac{1}{2}+2^{-\frac{1}{2}n-1}$$

推论：如果$f(x)$是m阶相关免疫的，则至少存在一个仿射函数$WX\oplus V$使得

$$P(WX\oplus V)\geq\frac{1}{2}+\frac{1}{2}\left(2^n-\sum_{i=0}^{m}\binom{n}{i}\right)^{-\frac{1}{2}}$$

定理4：命$P(WX,1\oplus WX)=\max\{P(WX),P(1\oplus WX)\}$，则

$$\sum_{W\in GF(2)^n}(2P(WX,1\oplus WX)-1)^2=1$$

上述定理3和4可由引理4和1推出。

肖和Massey〔2〕证明了，$f(x)$是m阶相关免疫的当且仅当$S_f(W)=0$，$1\leq W_H(W)\leq m$。由定理1我们知道f是m阶相关免疫的当且仅当$S_{(f)}(W)=0$，$1\leq W_H(W)\leq m$。从而$f(x)$是m阶相关免疫的当且仅当$P(WX\oplus V)=\frac{1}{2}$对所有满足$1\leq W_H(W)\leq m$的W成立。但定理4告诉我们，$f(x)$与所有仿射函数的复合率是守恒的，这说明带有一定相关免疫性的函数未必是好的滤波函数。Siegenthaler在〔9〕中给出了所有本质不同的2个，3个和4个变元的相关免疫函数。因为$\max_W|S_{(f)}(W)|\geq\frac{1}{2}$对〔9〕中的所有函数都成立。所以，所有4个变元以下的相关免疫函数在流密码中都不能作为滤波函数。我们认为用相关免疫函数作为滤波函数的流密码没有用满足如下两条件的函数作为滤波函数的流密码安全，1）$S_{(f)}(0)=0$；2）对所有$W\neq 0$，$|S_{(f)}(W)|$是几乎相等的。

由以上的讨论得知，下述两个指标作为流密码的安全性度量是合理的：

a）$$PV(f)=\max_W|S_{(f)}(W)|$$

b）$$VS(f)=\sum_{W=0}^{2^n-1}(S_{(f)}(W)^2-2^{-n})^2$$

$$=\sum_{W=0}^{2^n-1}S_{(f)}(W)^4+2^{1-n}$$

这里我们视$\{S_{(f)}(W)^2, W\in GF(2)^n\}$为一个随机变量的概率分布。

两个指标的关系为

$$(PV(f)^2-2^{-n})^2\leq VS(f)\leq(2^{\frac{1}{2}n}PV(f)^2-2^{-\frac{1}{2}n})^2$$