文章编号:1007-5321(2006)04-0119-04

A5/1 算法可抵抗相关攻击的改进方法

陈 伟[1], 杨义先[1], 钮心忻[2]
(1. 北京邮电大学 信息安全中心, 北京 100876; 2. 北京邮电大学 数字内容研究中心, 北京 100876)

摘要:在介绍 A5/1 算法的线性初始化弱点基础上,通过分析基于择多逻辑的互钟控机制的非平衡输出特点,详细论述了由此带来的相关攻击漏洞,指出了 A5/1 算法丢弃起始 100 bit 远不能保证算法安全性. 针对该漏洞,将互钟控移位延伸到 A5/1 算法初始化过程中,从而有效堵塞该漏洞,增强了 A5/1 算法的安全性.

关 键 词:A5/1 算法;初始化;相关分析
中图分类号:TP309 **文献标识码**:A

Improvement of A5/1 Algorithm Against Correlation Attack

CHEN Wei[1], YANG Yi-xian[1], NIU Xin-xin[2]
(1. Information Security Center, Beijing University of Posts and Telecommunications, Beijing 100876, China;
2. Research Center of Digital Contents, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: On the basis of introduction of A5/1 linear initiation weakness, by the characteristic of non-balance output in inter-clocking mechanism based on majority logic is analyzed, and the correlation attack weakness is discussed. Meanwhile, that thrown-out of original 100 bits output in A5/1 algorithm that can not approve its security is pointed out. For such weakness, inter-clocking mechanism is extended to the originate process of A5/1 to eliminate such weakness. It is concluded that A5/1 algorithm is safe under such attacks.

Key words: A5/1 algorithm; initialization; correlation analysis

无线通信的兴起引发了人们对相关协议和算法的研究热潮[1-2]. 作为欧洲数字蜂窝移动电话系统(GSM)中采用的流密码加密标准,A5/1 算法(简称 A5/1)自然成为中外密码专家的研究热点[3-4]. 除用特殊机器、耗费巨大资源直接攻击 A5/1 外,目前破解 GSM 通信较简便的方法是利用 GSM 系统漏洞,即通过 A5/2 算法漏洞发现共享的 A5/1 密钥[5].

A5/1 在初始化过程中存在密钥比特和帧号比特都是线性填充的弱点. 因此,文献[6]提出一种简便的相关攻击方法,其攻击的有效性与 A5/1 的线性移位寄存器长度和抽头数无关,而与其产生密钥流之前丢弃的输出比特数(A5/1 为 100)有关;只需 2^{16} 帧 A5/1 数据报(约 5 min GSM 通信量),破译它的成功率就可达到 70%,并且所需的存储量非常小. 文献[7]则运用相关分析技术和统计工具找到了区分 A5/1 输出密钥流与随机比特序列的方法.

1 A5/1 线性填充漏洞

A5/1 工作在流密码计数器模式下,输入为 86 bit 密钥,其中 64 bit 为会话密钥 $K=(k_0, k_1, \cdots, k_{63})$、22 bit 为帧序列号 $F_n=(F_0, F_1, \cdots, F_{21})$ 生成的 228 bit 密钥流控制双向信道,每个方向的 114 bit 密钥流与 114 bit 的明文/密文进行异或运算,产生 114 bit 密文/明文. 算法内部结构见图 1.

收稿日期:2005-06-09
基金项目:国家自然科学基金项目(60372094)
作者简介:陈 伟(1973—),男,博士生,E-mail:cyberella2000@163.com.

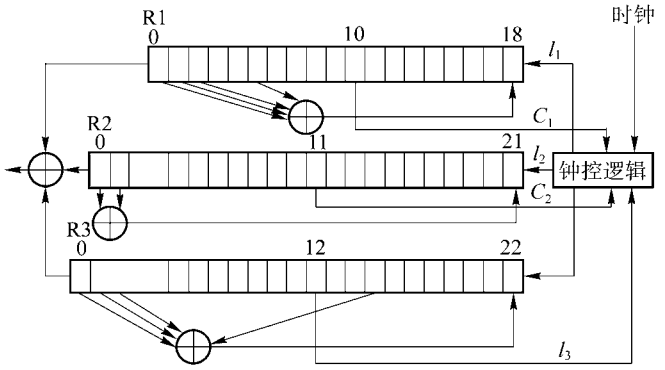


图1 A5/1的内部结构

A5/1内部由3个线性反馈移位寄存器(LFSR)R0、R1和R2组成,级数分别为19、22、23,抽头数分别为4、2、4个.将R0、R1和R2的异或作为位输出,每一帧密钥流的产生都分为"初始化"和"输出"2个过程.初始化过程为,先将3个LFSR全部清0;然后,在64个时钟周期内,将每个LFSR都移位64次,每次移位后将密钥比特 k_i(i为0～63)与每个LFSR输出异或运算后置入该LFSR的最低位;最后,在22个时钟周期内,将每个LFSR都移位22次,每次移位后将帧序列号比特 F_j(j为0～21)与每个LFSR输出异或运算后置入该LFSR的最低位.对于每个 n 级LFSR,初始化过程可以表达为

$$a_{t_1+n}=k_i+\sum_{k=1}^{n}c_k a_{t_1+n-k},\quad 0\leqslant t_1\leqslant 63$$

$$a_{t_2+t_1+n}=F_j+\sum_{k=1}^{n}c_k a_{t_2+t_1+n-k},\quad 0\leqslant t_2\leqslant 21$$

式中 $c_k\in F_2$.所以,初始化结束时每个LFSR的内部状态都是 k_i 和 F_j 的线性函数,与所有密钥比特和帧序列号比特都线性相关.以 u_t^0 表示R0在初始化完成后在 t 时刻的输出,则

$$u_t^0=\sum_{i=0}^{63}c_{it}^0k_i+\sum_{j=0}^{21}d_{jt}^0f_i \qquad (1)$$

式中 c_{it}^0($i=0,1,\cdots,63$)、d_{jt}^0($j=0,1,\cdots,21$)均为未知的0、1常数.令 $s_t^0=\sum_{i=0}^{63}c_{it}^0k_i$,称为 u_t^0 的密钥部分;$f_t^0=\sum_{j=0}^{21}d_{jt}^0f_i$,称为 u_t^0 的帧序列号部分,则式(1)可写成

$$u_t^0=s_t^0+f_t^0,\quad t\geqslant 0$$

表示R0在 t 时刻的输出是密钥 K 和帧序列号 F_n 的线性函数.对R1、R2也有类似等式成立.

2 基于择多逻辑的互钟控机制

A5/1的设计采用"3个LFSR互相控制时钟"的结构,有3个钟控输入(分别为每个LFSR的中间位)和3个钟控输出(分别控制每个LFSR的停/走).钟控机制采用择多逻辑,在每一轮中时钟驱动2个LFSR移位.择多逻辑真值表如表1所示.

表1 择多逻辑真值表

x_1	x_2	x_3	l_1	l_2	l_3
0	0	0	1	1	1
0	0	1	1	1	0
0	1	0	1	0	1
0	1	1	0	1	1
1	0	0	0	1	1
1	0	1	1	0	1
1	1	0	1	1	0
1	1	1	1	1	1

从表1可见择多逻辑的特点,每个时钟周期内,每个LFSR移位的平均概率都是3/4.设A5/1输出帧的第 z 比特为 z_z 时,R0、R1和R2正好分别移位 l_0、l_1 和 l_2 拍,即有等式 $z_z=u_{l_0}^0\quad u_{l_1}^1\quad u_{l_2}^2$ 成立.从择多逻辑函数可以导出如下概率递推公式

$$\begin{aligned}
&p(l_1,l_2,l_3,0)=1,\ 若\ l_1=l_2=l_3=0\\
&p(l_1,l_2,l_3,z)=0,\ 若\ l_1<0\ 或\ l_2<0\ 或\ l_3<0\\
&p(l_1,l_2,l_3,z)=0,\ 若\ l_1>z\ 或\ l_2>z\ 或\ l_3>z\\
&p(l_1,l_2,l_3,z)=0.25p(l_1-1,l_2-1,l_3-1,z-1)+\\
&\quad 0.25p(l_1-1,l_2-1,l_1,z-1)+\\
&\quad 0.25p(l_1-1,l_2,l_3-1,z-1)+\\
&\quad 0.25p(l_1,l_2-1,l_3-1,z-1)
\end{aligned} \qquad (2)$$

根据式(2)计算表明,$p_{l_0,l_1,l_2,z}$ 值随 l_0、l_1、l_2 和 z 的增大而下降.当 $z>100$ 时,其最大值已小于0.001,如表2所示.

表2 $z>100$ 时的最大概率简表

l_0	l_1	l_2	z	$p_{l_0,l_1,l_2,z}$
76	76	76	101	0.000 974 338
79	79	79	105	0.000 920 123
82	82	82	109	0.000 870 757
85	85	85	113	0.000 825 652
88	88	88	117	0.000 784 311

3 相关攻击

A5/1的每帧密钥流产生前,都要先丢弃钟控移位输出的前100 bit,以便充分降低初始化线性填充的不利影响,但攻击时不妨将之也计算在内.相关

攻击的主要思路是,利用 A5/1 初始化中的线性填充弱点,可以将密钥信息和帧号信息的影响分开.因为每个密钥流比特都是原始密钥信息和帧号信息的函数,如果去除了帧号信息,就能以高概率识别出原始密钥信息.攻击步骤为,先计算出 R0、R1、R2 输出与 A5/1 输出的相关概率,然后计算将 R0、R1、R2 输出与 A5/1 输出的一些相关比特异或值,去除密钥流中的帧号信息.在只剩下原始密钥信息后,大量帧的某个或某些位的集合将会表现出一定的 0-1 不平衡性,从而可以进行密钥信息的概率推导.

从 $z_z = u^0_{l_0} \quad u^1_{l_1} \quad u^2_{l_2}$ 可以得到

$$s^0_{l_0} \quad s^1_{l_1} \quad s^2_{l_2} = \hat{f}^0_{l_0} \quad \hat{f}^1_{l_1} \quad \hat{f}^2_{l_2} \quad z_z \qquad (3)$$

此等式左边为密钥信息的异或,右边是 A5/1 输出帧的第 z 比特和 R0、R1、R2 输出中的第 l_0、l_1、l_2 位进行异或,滤掉帧号信息影响的结果.

由于互钟控机制的存在,每个寄存器的"停/走"受其他寄存器状态的影响.A5/1 输出某一帧的第 z 比特不一定由 R0、R1 和 R2 输出序列中对应帧的第 l_0、l_1 和 l_2 位异或计算得到.当正好第 z 比特由第 l_0、l_1 和 l_2 位异或计算得到时,式(2)成立的概率为 1;否则,式(2)成立的概率为 1/2.设情况 1 发生的概率为 $p_{l_0,l_1,l_2,z}$,则情况 2 发生的概率为 $1 - p_{l_0,l_1,l_2,z}$.

对于同一次会话,n 个不同帧的第 z 比特构成 n 长 0-1 序列,式(3)成立的概率为

$$p(s^0_{l_1} \quad s^1_{l_1} \quad s^2_{l_2} = \hat{f}^0_{l_0} \quad \hat{f}^0_{l_1} \quad \hat{f}^1_{l_2} \quad z_z) =$$

$$p_{l_0,l_1,l_2,z} \times 1 + (1 - p_{l_0,l_1,l_2,z}) \times \frac{1}{2} = \frac{1 + p_{l_0,l_1,l_2,z}}{2} \qquad (4)$$

式(4)是 A5/1 输出帧的第 z 比特和 R0、R1、R2 输出的第 l_0、l_1、l_2 位之间的相关概率.

对同一次会话产生的密钥流序列,取 n 个不同帧的第 z 比特构成一条 n 长 0-1 序列,它和 R0、R1、R2 输出中对应帧的第 l_0、l_1、l_2 位构成的 3 条 n 长序列进行异或运算,得到 $\hat{f}^0_{l_0} \quad \hat{f}^1_{l_1} \quad \hat{f}^2_{l_2} \quad z_z$,即第 z 比特过滤掉帧号信息影响后的 n 长"畸变 A5/1 序列".从式(4)知它是个不平衡 0-1 序列,具有一定的分布不平衡偏差率.有关攻击细节见文献[6-7].

4　A5/1 算法的改进

A5/1 初始化中线性填充的目的是使 3 个移存器的内部状态与每一密钥比特和帧序列号比特都相关,但是线性运算的使用却犯了密码设计大忌,招致了相关攻击.基于择多逻辑的互钟控非线性前馈组合机制有助于弱化线性漏洞.因此,算法设计者不得不降低效率,丢弃钟控输出的前 100 bit,以此弥补线性漏洞.即使如此,通过 3 阶统计,还是可以剥离 A5/1 密钥流的伪随机性[7].

作为非线性前馈组合的互钟控机制有很多优良特性[8].通过该机制也可以得到线性填充的"每一比特相关"性.由于相关攻击的有效性仅与 A5/1 产生密钥流之前丢弃的输出比特数有关,如果将基于择多逻辑的互钟控机制从 A5/1 的密钥流产生过程延伸到初始化过程,等于增加丢弃的互钟控输出比特数,则基本丧失了相关攻击的有效性.表 3 为 z 较大时根据式(2)计算得出的部分数据.

表 3　关于 z 的部分最大概率简表

l_0	l_1	l_2	z	$p_{l_0,l_1,l_2,z}$
132	132	132	176	0.000 432 013
135	135	135	180	0.000 416 971
138	138	138	184	0.000 401 898
141	141	141	188	0.000 387 146
144	144	144	192	0.000 373 011

由上看出必须对 A5/1 进行改进,形成新的 A5/1 算法.

1) 初始化过程.①将 3 个 LFSR 全部清 0.②在64 个时钟周期内,将每个 LFSR 都互钟控移位 64 次,并在每次移位后将密钥比特 k_i($i = 0 \sim 63$)与每个 LFSR 输出异或后置入该 LFSR 最低位.③在 22 个时钟周期内,将每个 LFSR 都互钟控移位 22 次,并在每次移位后将帧序列号比特 F_i($j = 0 \sim 21$)与每个 LFSR 输出异或运算后置入该 LFSR 最低位.

2) 产生输出密钥流过程.将 3 个 LFSR 钟控移位 100 拍,丢弃输出.钟控输出 228 bit,作为双向通信中使用的密钥流.

改进后的 A5/1 相当于丢弃互钟控移位生成的 186 bit,从第 187 bit 开始输出.从表 3 可见,不平衡偏差率仅相当于原算法的 2/5.根据文献[7]中的 3 阶统计指标,当 $n = 90\,000$ 时,原算法的正态偏移值 $E'(X') = 3.365$,$D'(X') = 1.174$.如图 2 所示.而改进算法的 3 阶统计指标正态偏移值 $E'(X') = 0.988$,$D'(X') = 0.153$.

下面对改进算法和原算法进行 2 组各 200 条随

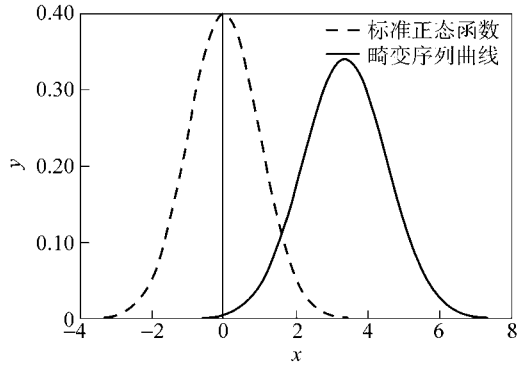
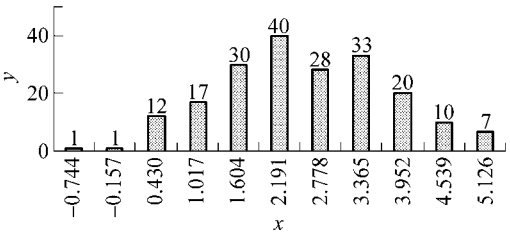
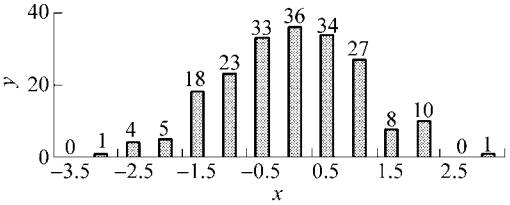


图2 n=90 000时畸变序列产生的3阶统计曲线与标准正态曲线的对比

机密钥下的对比实验,所得数据如图3和4所示.

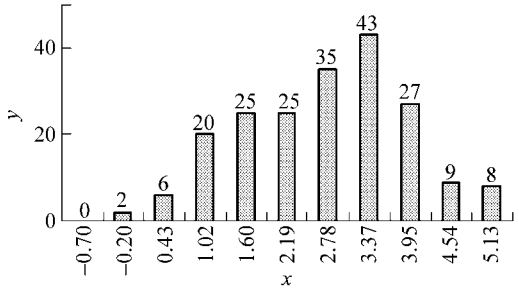


(a) A5/1算法计算值分布曲线图

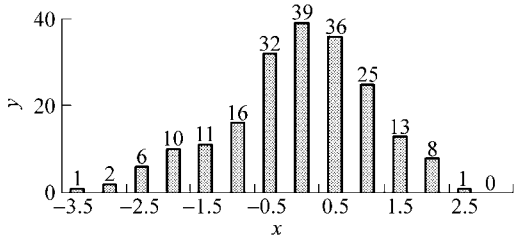


(b) 改进A5/1算法计算值分布曲线图

图3 第1组A5/1及其改进型的输出序列3阶统计值分布图对比



(a) A5/1算法计算值分布曲线图



(b) 改进A5/1算法计算值分布曲线图

图4 第2组A5/1及其改进型的输出序列3阶统计值分布图对比

从实验数据上可以看出,改进后A5/1的伪随机性与原算法相比有很大的改进.

5 结 论

A5/1初始化中的线性填充容易引发相关攻击.本文针对这一漏洞进行的改进计算代价很小,却有效地堵塞了对A5/1的相关攻击.A5/1是一个非常值得研究的经典的序列密码算法,因而对其还要做很多后续的研究工作.

参考文献:

[1] 姚惠明,隋爱芬,杨义先.3GPP网络AKA协议中若干算法的设计[J].北京邮电大学学报,2002,25(3):98-102.
Yao Huiming, Sui Aifen, Yang Yixian. Design of some algorithms in AKA protocol of 3GPP network[J]. Journal of Beijing University of Posts and Telecommunications, 2002, 25(3): 98-102.

[2] 沈苏彬.开放电信业务中的若干技术问题探讨[J].北京邮电大学学报,2004,27(增刊):16-24.
Shen Subin. Research of some technique issue in opening telecommunication business[J]. Journal of Beijing University of Posts and Telecommunications, 2004, 27(Sup): 16-24.

[3] Schneier B. Applied cryptography second edition: protocols, algorithms, and source code in C[M]. New York: John Wiley & Sons, 1996.

[4] 王育民,刘建伟.通信网的安全——理论与技术[M].西安:西安电子科技大学出版社,1999:100-101.

[5] Barkan E, Biham E, Keller N. Instant ciphertext-only cryptanalysis of GSM[R]. Haifa: Technion-Computer Science Department Technical Report CS-2003-05, 1-18.

[6] Ekdahl P, Johansson T. Another attack on A5/1[J]. IEEE Transactions on Information Theory, 2003, 49(1): 284-289.

[7] 陈伟,胡云,杨义先,等.基于相关攻击的A5/1算法识别[J].电子与信息学报,2006,28(5):827-831.
Chen Wei, Hu Yun, Yang Yixian, et al. Identify of A5/1 algorithm based on correlation attack[J]. Journal of Electronic Information, 2006, 28(5): 827-831.

[8] Kholosha A. Clock-controlled shift registers for keystream generation[D]. Netherlands: Technische Universiteit Eindhoven, 2002.