第14卷 第2期 电子科学学刊 Vol. 14, No. 2
1992 年 3 月 JOURNAL OF ELECTRONICS Mar., 1992

密码分析的连分数策略*

李 大 兴
(山东大学数学系,济南 250100; 中国科技大学研究生院信息安全国家重点实验室,北京 100039)
李 大 为
(杭州电子工业学院计算机系,杭州 310037)

提要 本文基于连分数的 Legendre 定理提出了用于密码分析的连分数策略。

关键词 密码分析;连分数策略;算法

一、引 言

本文基于连分数的 Legendre 定理[1],提出了一种密码分析的有效方法——连分数策略. 该策略由渐近条件和求解算法组成. 任何满足渐近条件的密码体制都可由本文给出的通用求解算法来破译. 利用这一策略, 不但可使已有的对某些密码体制的攻击方法大为简化,而且还可用来破译一些未曾破译的密码体制. 这一策略原理简明且适用性较广,从而易于掌握和灵活运用. 它不但是一种密码分析手段, 同时也可提醒密码设计者应回避那些满足或接近满足渐近条件的密码体制.

二、连分数策略的渐近条件和求解算法

众所周知,对正整数 $a_0, a_1, \cdots, a_{N-1}$, 分数 $a_0+1/(a_1+1/(a_2+1/(\cdots/(a_{N-2}+1/a_{N-1})\cdots)))$ 称为(有限)连分数,并简记为 $[a_0, a_1, \cdots, a_{N-1}]$. 对 $0\leqslant n\leqslant N-1$, 令 $[a_0,a_1,\cdots,a_n]=p_n/q_n$, 称 p_n/q_n 为 $[a_0,a_1,\cdots,a_{N-1}]$ 的第 n 个渐近值, 且 p_n 和 q_n ($0\leqslant n\leqslant N-1$) 可由如下递归式确定

$$\left.\begin{aligned} p_{-1}\triangleq 1,\ p_0=a_0,\ p_n=a_np_{n-1}+p_{n-2}\\ q_{-1}\triangleq 0,\ q_0=1,\ q_n=a_nq_{n-1}+q_{n-2}\end{aligned}\right\} \tag{1}$$

反之, 对有理数 $p/q((p,q)=1)$, 依关系: $\alpha_0=p/q$, $a_0=\lfloor\alpha_0\rfloor$, 令 $\alpha_n=1/(\alpha_{n-1}-a_{n-1})$, $a_n=\lfloor\alpha_n\rfloor$ 直到 $\alpha_{N-1}=a_{N-1}$ 为止,便可确定 p/q 的连分数表示

$$p/q=[a_0,a_1,\cdots,a_{N-1}]$$

1991.03.19 收到, 1991.10.19 定稿.
* 国家自然科学基金资助项目

将(1)式中的 $\{q_n|n=-1,0,1,\cdots,N-1\}$ 与 Fibonacci 数列 $\{F_n|n=0,1,\cdots,N\}$ 做比较，不难归纳地验证 $q_n \geq F_{n+1} > \phi^n$ (其中 $\phi=(1+\sqrt{5})/2$)，从而 $q=q_{N-1}>\phi^{N-1}$，或 $N \leq \lfloor \log_\phi q \rfloor + 1$，因此我们有

性质 1 求 p/q 的连分数表示可以在 $O(|q|)$ 次迭代步内完成，其中

$$|q| = \lfloor \log_2 q \rfloor + 1$$

为 q 的位长。

对实数 α，令 $[a_0, a_1, \cdots, a_{N-1}]$ 为 p/q 的连分数表示，$q_n(0 \leq n \leq N-1)$ 如(1)式所确定的，命 $\alpha - p/q = \varepsilon\theta/q^2$，$\varepsilon = \pm 1$，$0<\theta<1$. 由 Legendre 判别条件知 p/q 为 α 的渐近值的充要条件为 $\theta \leq q_{N-1}/(q_{N-1}+q_{N-2})$.

因 $[a_0,a_1,\cdots,a_{N-2},1]=[a_0,a_1,\cdots,a_{N-2}+1]$，所以在 p/q 的连分数表示中不妨设 $a_{N-1} \geq 2$，从而总有 $q_{N-1}/(q_{N-2}+q_{N-1})>2/3$，结合 Legendre 判别条件可得

性质 2 若存在一有理数 p/q 适合 $|\alpha - p/q| < (2/3)\cdot(1/q^2)$，则 p/q 必为 α 的某一渐近值。

依性质 2，我们可以给出如下适用于密码分析的连分数策略。

设 $\alpha = P/Q$ 是由已知信息（密文和公钥体制中的公钥或传统体制中的一些明密对等)构成的有理分数，称之为**公开分数**. 设 $\alpha' = P'/Q'$ 是由秘密信息(明文或密钥等)构成的有理分数，其中 $(P',Q')=1$，称 α' 为**秘密分数**. 若 α 与 α' 满足如下渐近条件：

渐近条件 $|\alpha - \alpha'| < (2/3)\cdot(1/Q)^2$ 则 P'/Q' 必等于 P/Q 的某一渐近值 p_i/q_i，因 $(P',Q')=(p_i,q_i)=1$，所以 $P'=p_i$，$Q'=q_i$，从而 P' 和 Q' 可以通过计算 p_i 和 q_i 来获得，如下的算法便是这一思想的具体化。

求解算法 (简记 SOL-算法)

输入 P，Q

(1)(初始化) $\alpha_0 := P/Q$，$a_0 := \lfloor \alpha_0 \rfloor$，$p_{-1} := 1$，$p_0 := a_0$，$q_{-1} := 0$，$q_0 := 1$，$i := 0$.

(2)(判别条件)若 $B(i)$ 为真且 $\alpha_i - a_i \neq 0$ 则执行(3)；否则执行(4).

(3) $i := i+1$，$\alpha_i := 1/(\alpha_{i-1}-a_{i-1})$，$a_i := \lfloor \alpha_i \rfloor$，$p_i := a_i p_{i-1} + p_{i-2}$，$q_i := a_i q_{i-1} + q_{i-2}$，回到(2).

(4) 输出 p_i，q_i，停止.

说明 1 上述算法中的谓词 $B(i)$ 为真，当且仅当 $q_i \neq Q'$ 或 $p_i \neq P$ 为真，因秘密信息 P' 或 Q' 并不知道，所以对 $B(i)$ 并不能直接判别，但常常可用其他条件代替，这是由于一旦 $q_i = Q'$ 或 $p_i = P'$，则我们已经得到了密码的关键秘密，从而导致密码的崩溃，因此判别 $B(i)$ 是否为真就可以通过将 q_i 和 p_i 视为 Q' 和 P' 并验证是否密码已崩溃来确定，参见本文后面的应用实例.

说明 2 由于 SOL-算法中起关键作用的是中间变量 $a_1, a_2, \cdots$等，它们都是正整数，分数 $\alpha_0, \alpha_1, \cdots$等是为计算 $a_1, a_2, \cdots$等服务的，在实际计算时没有必要求出 $\alpha_0, \alpha_1, \cdots$等分数的小数表示，只须用分子和分母表示出来就足够了(相当于用一个整变量的二元数组表示一个分数)，这样在计算 $a_0, a_1, a_2 \cdots$ 时只相当于计算一下 $\alpha_0, \alpha_1, \alpha_2, \cdots$ 的整部(用分母除分子计算到小数点为止)，其计算时间和作一次整数乘法是同一

量级的(只须注意到除法和乘法的时间是同一量级的[2]).

令 $M(k)$ 记两个不超过 k 位整数相乘的时间，D 记判别 $B(i)$ 的计算时间，则由性质 1 和整数算术的计算性质[2]，可知 SOL-算法的计算时间为

$O_B(k(D+M(k)))$，其中 $k=\max\{|P|,|Q|\}$

三、应　　用

1. 破译 Okamoto 体制

Okamoto 体制[3]以其加解密迅速而引起人们的兴趣．许多学者曾对其进行过各种形式的攻击，我们在文献[4]中利用 Euclid 算法彻底破译了这一体制．本文指出连分数策略亦可用来破译 Okamoto 体制．

Okamoto 体制的

公钥　$n=p^2q$，$u=a+bpq$

密钥　(p,q,a,b)，其中 p,q 为两个大素数，$p<q$，$0<a<\sqrt{pq}/2$，$a\in Z_{pq}^*$，$0<b<p$.

令公开分数 $\alpha=u/n$，秘密分数 $\alpha'=b/p$，$(b,p)=1$．不难验证

$$|\alpha-\alpha'|=a/(p^2q)<1/(2p^2)$$

这表明渐近条件成立．SOL-算法中的谓词 $B(i)$ 可由 "$q_i=1$ 或 $q_i\nmid n$" 代替，这是因为 $q_i=p$ 时我们就获得了 n 的一个非平凡因子．注意到 $q_i\nmid n$ 当且仅当

$$n-q_i\cdot\lfloor n/q_i\rfloor\neq 0$$

所以判别 $B(i)$ 的时间 $D=O_B(M(|n|))$，从而 SOL-算法的计算时间为 $O_B(|n|M(|n|))$，这与文献[4]的结果一致．

2. 破译丢番图 (Diophantine) 公钥体制

杨义先等人[5]曾基于二次丢番图方程的难解性，提出了二类丢番图公钥体制．我们在文献[6]中曾利用文献[4]的攻击方法破译了第一类丢番图体制并部分地破译了第二类丢番图体制．

第一类丢番图体制的

公钥　$m=p^2q$，$n=up^2+v$

密钥　(p,q,u,v)，其中 $p>(1/4)q^{3/2}$，$(q,up^2+v)=1$，$v<(1/2)p^{1/2}q^{1/4}$

文献[6]已指出，若 $(m,n)=(u,q)=1$ 不成立，则该体制很容易破译，所以只须考虑 $(m,n)=(u,q)=1$ 的情景．令公开分数 $\alpha=n/m$，秘密分数 $\alpha'=u/q$，则不难验证渐近条件满足．判别谓词 $B(i)$ 可由 "$q_i=1$ 或 $q_i\nmid m$" 代替．类似于第三·1 节的讨论，SOL-算法的计算时间与文献[6]的 SOL-算法是同一量级的．

用类似的方法可以讨论第二类丢番图公钥体制，并可得到比文献[6]稍强一些的结论．

3. 破译短解密指数的 RSA 体制

RSA 体制是当今最为成功的公钥密码体制，已在国内外众多的信息安全系统中得到应用．在使用 RSA 体制建立信息系统的安全机制时，常会遇到象 Smart 卡与计算机、

终端与主机这样的计算能力不平衡的用户之间的通信保密问题. 有些学者曾建议为计算能力弱的用户设计短解密指数的 RSA 体制，以使其解密时间大为缩短. 最近，M. J. Wiener[7] 利用连分数的性质给出了短解密指数的 RSA 体制的攻击方法. 下面我们指出，连分数策略可使 Wiener 的攻击方法无论在原理上，还是在算法上都大为简化.

众所周知，RSA 体制的公钥由 $n=pq$ 和 e 组成，其中 p,q 为两个不同的大素数，e 为加密指数，满足 $(e,\varphi(n))=1$，通常 p 和 q 的位长取得一样或相近. p,q 保密，解密指数 d 由 $ed\equiv 1(\mathrm{mod}\varphi(n))$ 确定. 短解密指数的 RSA 体制其设计如下：选某一较短的 $d,(d,\varphi(n))=1$，由 $ed\equiv 1(\mathrm{mod}\varphi(n))$ 确定加密指数 e，通常 e 取在 $0\sim\varphi(n)-1$ 之间.

由加密指数和解密指数的关系可知存在整数 k 使 $ed=k(p-1)(q-1)+1$，由此立得 $(k,d)=1$ 且 $e/n=(k/d)(1-\delta)$，其中 $\delta=(p+q-1-1/k)/n$，亦即 $|e/n-k/d|=(k/d)\delta$. 令 $\alpha=e/n$ 为公开分数，$\alpha'=k/d$ 为秘密分数，则渐近条件成立，当且仅当 $(k/d)\delta<(2/3)\cdot(1/d^2)$ 成立，即当且仅当 $kd<(2/3)\cdot(1/\delta)$ 成立，若 $e<\varphi(n)$，则 $k<d$，从而当 $d<\sqrt{(2/3)(1/\delta)}(\approx O(n^{1/4}))$ 时渐近条件成立.

注意到如下等式组

$$\begin{cases}(p+q)/2=(1/2)[-(ed-1)/k+(n+1)]\\(p-q)/2=\pm\sqrt{[(p+q)/2]^2-n^2}\end{cases}$$

可见当 $p_i=k$，$q_i=d$ 时，p 和 q 可立即得知，因此判别谓词 $B(i)$ 可由如下子程序替代

(1) $x_i:=(1/2)[-\lfloor(eq_i-1)/p_i)\rfloor+(n+1)]$

$y:=\lfloor\sqrt{x^2-n^2}\rfloor$，$z:=x+y$

(2) 若 $z=1$，n 或 $z\nmid n$，则视 $B(i)$ 为真；否则为假.

由于除法和开平方运算与乘法的时间是同一量级的，所以上一子程序的计算时间

$$D=O_B(M(|n|))$$

从而由 SOL-算法求短解密指数 d 的计算时间为 $O_B(|n|M(|n|))$.

RSA 体制更为紧凑的形式是加解密指数 e 和 d 满足关系

$$ed\equiv 1(\mathrm{mod}LCM[p-1,q-1])$$

即存在整数 K 使 $ed=K\cdot LCM[p-1,q-1]+1$. 令 $G=(p-1,q-1)$，则

$$ed=(K/G)(p-1)(q-1)+1$$

再令 $k=K/(K,G)$，$g=G/(K,G)$，则 $k/g=K/G$ 且 $(k,g)=1$，和

$$(k,d)\leqslant(K,d)=1$$

由此得 $(k,dg)=1$. 不难看出 $e/n=(k/(dg))(1-\delta)$，其中

$$\delta=(p+q-1-g/k)/n$$

令公开分数 $\alpha=e/n$，秘密分数 $\alpha'=k/(dg)$，则渐近条件成立，当且仅当

$$k\delta/(dg)<(2/3)\cdot(1/(dg))^2$$

亦即 $kdg<(2/3)\cdot(1/\delta)$. 与文献[7]的结果一致，依此可推出文献[7]中的全部结论.

4. 破译短秘密指数的 Dickson 公钥体制

一些学者基于 Dickson 多项式的性质提出了几种公钥体制[8,9]，这里统称为 Dickson 公钥体制．当模 n 为两个位长相同或相近的大素数之积时是最为实用的形式．类似于第三·3 节的讨论，可以看出连分数策略可以攻击 $d = O(n^{1/2})$ 的短解密指数的 Dickson 公钥体制．

5. 预测 D-序列

作为伪随机序列家族一员的 D-序列已得到很多研究，并在相关特性和游程分布等方面取得了一批成果[10-12]．本节利用连分数策略证明 D-序列是很容易被预测的，无论其是否为极长 D-序列．这说明 D-序列作为流密码的密钥流是不安全的．

依文献[12]的记法，$GF(2)$ 上的 D-序列 $\{p/q\}_2$ 用 $a_1a_2\cdots a_{q-1}$ 表示，其中

$$a_i \in \{0,1\},\ (i = 1,\cdots,q-1)$$

由递归式 $(i = 1,\cdots,q-1)$

$$m_0 = p,\ a_i = \lfloor 2m_{i-1}/q \rfloor,\ m_i = 2m_{i-1} - qa_i$$

来确定，亦即 a_i 为 $2m_{i-1}$ 除 q 之商，m_i 为余数，从而 $0 \leqslant m_i < q$．由此递归式可以看出，若对某个 $i_0 \in \{1,\cdots,q-1\}$，m_{i_0} 和 q 已知，则序列 $\{p/q\}_2$ 中 a_{i_0+1} 及其以后的各项都可以求出来．不但如此，a_{i_0+1} 之前各项也可逐项求出来．这是因为 $2m_{i_0-1}$ 为偶数，而 q 为奇数，所以 $a_{i_0} = 0$，当且仅当 m_{i_0} 为偶数，再依 $m_{i_0-1} = (1/2)(m_{i_0} + qa_{i_0})$ 便可求出 m_{i_0-1}．依此类推，可求出 a_{i_0+1} 以前的各项．

下面我们来看一看如何利用 SOL-算法从 D-序列的一段不太长的连续若干个比特流中求出某个 m_{i_0} 和 q，从而等于完全掌握了整个 D-序列．

设 $\varphi = b_1b_2\cdots b_k$ 为长为 k 的比特流，$[\varphi]$ 表示将 φ 视为二进制整数的表示时所代表的整数，即 $[\varphi] = \sum_{i=1}^{k} b_i \cdot 2^{k-i}$．例如，$[101]=5$，$[011]=3$．

考查 D-序列，设 $1 \leqslant i \leqslant g-1$，$k \geqslant 1$，则有

$$\begin{aligned} m_{i+k} &= 2m_{i+k-1} - q \cdot [a_{i+k}] = 2(2m_{i+k-2} - q \cdot a_{i+k-1}) - q \cdot [a_{i+k}] \\ &= 2^2 m_{i+k-2} - q \cdot [a_{i+k-1}a_{i+k}] \\ &= \cdots = 2^k \cdot m_i - q \cdot [a_{i+1}a_{i+2}\cdots a_{i+k}] \end{aligned}$$

从而

$$[a_{i+1}a_{i+2}\cdots a_{i+k}]/2^k - m_i/q = -m_{i+k}/2^k \cdot q$$

若 D-序列 $a_1\cdots a_{q-1}$ 中的一段 $a_{i_0+1}a_{i_0+2}\cdots a_{i_0+k}$ 为已知，等价地说 $[a_{i_0+1}a_{i_0+2}\cdots a_{i_0+k}]$ 为已知，令 $\alpha = [a_{i_0+1}a_{i_0+2}\cdots a_{i_0+k}]/2^k$ 为公开分数，$\alpha' = m_{i_0}/q$ 为秘密分数，则有 $|\alpha - \alpha'| = m_{i_0+k}/(2^k \cdot q)$，从而渐近条件成立当且仅当 $m_{i_0+k}/(2^k \cdot q) < (2/3) \cdot (1/q)^2$，或写成 $2^{k+1} > 3 \cdot q \cdot m_{i_0+k}$．注意到 $1 \leqslant m_{i_0+k} < q$，所以当 $2^{k+1} > 3q^2$ 时渐近条件一定成立，亦即 $k \geqslant \lfloor \log_2(3q^2) \rfloor$（该值不超过 $2\lfloor q \rfloor$）时，m_{i_0}/q 为

$$[a_{i_0+1}a_{i_0+2}\cdots a_{i_0+k}]/2^k$$

的某一渐近值，设为第 s 个，即有 $m_{i_0} = p_s$，$q = q_s$．

设 k 满足 $2^{k+1} > 3q^2$，亦即上面讨论的渐近条件成立．在利用 SOL-算法求 m_{i_0} 和 q 时，判别谓词 $B(i)$ 规定为：若 p_i/q_i 的二进制小数的前 k 个有效位与 $a_{i_0+1}a_{i_0+2}\cdots a_{i_0+k}$ 不同，则视 $B(i)$ 为真，否则为假．显然当 $i = s$ 时，$p_i/q_i = m_{i_0}/q$，其小数表示就

是 D-序列从 a_{i_0+1} 开始的无限循环序列，所以 $B(s)$ 为假．令

$$t = \min\{n \in N \mid B(n) \text{ 为假}\}$$

则 $t \leqslant s$，且 SOL-算法输出 p_t 和 q_t．

下面证明 $t = s$，从而说明 SOL-算法能正确地求出 m_{i_0} 和 q． 结合前面的讨论，这意味着从任何一段长为 $2|q|$ 的序列段中就可完全确定出整个 D-序列来．

若 $t < s$，则 $q_t < q_s$ 且 $p_t/q_t \neq p_s/q_s$，所以 $|p_tq_s - p_sq_t| \geqslant 1$，由此得

$$\frac{1}{q_tq_s} \leqslant \frac{|p_tq_s - p_sq_t|}{q_tq_s} = \left|\frac{p_t}{q_t} - \frac{p_s}{q_s}\right| < \frac{1}{2^k}$$

后一不等式成立的原因是 p_t/q_t 与 p_s/q_s 的前 k 位相同．从而有

$$3q_s^2 = 3q^2 < 2^{k+1} < 2q_tq_s < 2q_s^2$$

这显然是不合理的．

因 k 可取 $2|q|$，所以计算 p_i/q_i 的前 k 位的时间为 $O_B(M(|q|))$，即判别 $B(i)$ 的时间 $D = O_B(M(|q|))$，故 SOL-算法的时间为 $O_B(|q|M(|q|))$．

当 q 的位长为 64 时，尽管 D-序列的周期近乎 2^{64}，但只须知道其任一段 128 比特，就可准确无误地确定出整个 D-序列来．

以上通过对 5 种密码体制的攻击来说明 SOL-算法在密码分析中是如何应用的．之所以称之为连分数策略而不称之为连分数方法是因为它只给出了一种攻击的思路，具体的攻击方法还须结合具体的密码体制的性质来给出．

四、一个计算例子

下面通过对一个短解密指数的 RSA 体制的例子进行具体的攻击，以说明 SOL-算法的计算过程．

取 $p = 3863$，$q = 4799$，$d = 41$，则 $n = 18538537$，$\varphi(n) = 18529876$，$e = 2259741$． 表 1 是整个计算过程的记录(忽略了对各项的具体计算过程)，其中公开分数 $\alpha = e/n = 2259741/18538537$．

表 1 SOL-算法的计算纪录表

i	α_i	a_i	p_i	q_i	$B(i)$
−1	—	—	1	0	—
0	2259741/18538537	0	0	1	真
1	18538537/2259741	8	1	8	真
2	2259741/460609	4	4	33	真
3	460609/417305	1	5	41	假

当 $B(i)$ 为假时，该行的 p_i/q_i 就是秘密分数． SOL-算法输出的是 5 和 41，其中 41 为解密指数．

参 考 文 献

[1] 华罗庚，数论导引，科学出版社，北京，1979 年，第 282—284 页．

[2] A. V. Aho et al., The Design and Analysis of Computer Algorithms, Addison-Wesley, (1975).
[3] T. Okamoto, *Electron. Lett.*, **22** (1986) 11, 581—582.
[4] 李大兴，张泽增，科学通报，**35** (1990) 11，871—874。
[5] 杨义先、李世群，罗群，通信学报，**10** (1989)2，78—80.
[6] 李大兴、赵霖，通信学报，**11** (1990) 1，93—96.
[7] M. J. Wiener, *IEEE Trans. on IT*, **IT-36** (1990) 3, 553—558.
[8] R. Lidl, W. B. Muller, Advances in Cryptology-Proc. Crypto 83, Plenum Press, New York, (1984), pp. 293—301.
[9] V. Varadharajan, *Int. J. Computer Math.*, **23** (1988) 2, 237—250.
[10] S. C. Kak, A. Chatterjee, *IEEE Trans, on IT*, **IT-27**(1981)5, 647—652.
[11] S. C. Kak, *IEEE Trans. on C*, **C-24** (1985) 9, 803—809.
[12] 高宝建，通信学报，**10** (1989) 4，74—77。

CONTINUED FRACTION TACTICS FOR CRYPTANALYSIS

Li Daxing

(*Shandong University, Jinan* 250100; *State Key Lab. of Information Security, Graduate School of University of Science and Technology of China, Beijing* 100039)

Li Dawei

(*Hangzhou Institute of Electronic Technology, Hangzhou* 310037)

Abstract The continued fraction tactics for cryptanalysis based on Legendre theorem of continued fraction are proposed.

Key words Cryptanalysis; Continued fraction; Algorithm