第7卷 第1期 2007年1月
1671-1819(2007)1-0122-03
科 学 技 术 与 工 程
Science Technology and Engineering
Vol.7 No.1 Jan. 2007
© 2007 Sci. Tech. Engng.

关于Hill密码密钥空间大小的计算

戴经国 张韶华[1*] 胡玉平 羊四清
(湖南人文科技学院计算机系,娄底417000;中国船舶重工集团公司第七二二研究所信息安全研究中心[1],武汉430079)

摘 要 初步综述了一些著名密码体制的密钥空间大小的计算,着重阐述了计算Hill密码的2阶可逆加密方阵个数(即密钥空间大小)的方法,给出了2阶Hill密码的密钥空间的计算表达式。

关键词 密钥空间 Hill密码 同余 二次同余的解
中图法分类号 TP309 O156.1; **文献标识码** A

一个密码通信系统通常由明文消息空间、密文消息空间、密钥空间以及加解密变换等几个部分组成。为了保护信息的机密性抵抗密码分析、选择合适大小的密钥空间是十分重要的。显然,如果密码体制的密钥空间过小,就容易被穷举攻击所破译。密码史上一个熟知的例子是1997年6月18日美国科罗拉多州以Rocke Verser为首的一个工作小组,通过Internet网,利用数万台微机,历时4个多月穷举破译了DES。1998年7月17日美国EFF(Electronic Frontier Foundation)用一台价值25万美元的计算机,只用了56 h就穷举破译了一个DES密钥。1999年EFF又将这种穷举速度提高到了24 h。由此可见,考虑密钥空间的大小是密码体制设计人员不可忽略的重要因素。目前一些著名的密码体制的密钥空间都大到足以抵抗当今计算能力的穷举攻击。例如RSA密码体制的密钥空间大小为$\varphi(\varphi(n))$,这里n为RSA模,它可取1 024比特位(宜取2 048比特位);EIGamal密码体制的密钥空间大小为$p-2$,这里p是150位以上的十进制随机大素数且$p-1$有大的素因子;椭圆曲线密码体制的密钥空间大小为q,这里q是椭圆曲线循环子群的生成元的阶,它是160位以上的十进制素数;AES的密钥空间大小可为2^{128}、2^{192}、2^{256}不等。以上这些密码体制的密钥空间大小都有比较简单的计算表达式;但是还有一些密码体制,要计算其密钥空间的大小并不容易。例如Hill密码。Hill密码[1]是由Lester S. Hill利用模算术和矩阵变换进行加解密的多字母代换密码。虽然目前它很少被采用,但它的设计思想仍给密码设计人员以诸多启示。

最简单的Hill密码加密过程如下:

(1)明文字母按照先后顺序每两个分一组,记任意一组为[a,b]。

(2)选一个正整数k做模,再选一个2×2矩阵A,A的元素为模k的剩余类中的元。

(3)密文为[a,b]与A的乘积模k,即$[a,b]\times A \pmod k$。

显然,能正确解密上述加密算法的充要条件是A的行列式值与模k互素。一个有趣的问题是加密矩阵空间到底有多大呢?即对于事先给定的模k,在模k的剩余类中,可逆的2×2矩阵有多少个呢?记在模k的剩余类中全体可逆的2×2矩阵的集合为$GL_2(Z_k)$,记$GL_2(Z_k)$的元素个数为$|GL_2(Z_k)|$。显然,为了计算$|GL_2(Z_k)|$,首先应考虑n与k互素

2006年9月6日收到 国家自然科学项目(No60573103)基金和湖南省教育厅重点项目(6A002)资助

第一作者简介:戴经国,男,(1962—),湖南双峰人,硕士,副教授,国防科学技术大学访问学者。研究方向:网络和信息安全等,E-mail:djghzp@163.com.cn。

*通信作者简介:张韶华,男,(1971—),湖南双峰人,硕士,工程师。研究方向:数论和密码学。

时,如何求同余方程 $ab-cd\equiv n(\bmod k)$ 的解数(记为 $T_{k,n}$)。这使得探讨如何求四元二次同余方程 $ab-cd\equiv n(\bmod k)$ 的解数问题,这里 a,b,c,d 都是整数,除非特殊说明,小写英语字母均代表整数。证明了下列定理:

定理 1 $T_{k,n}=k^3\prod_{p\mid k}\left(1-\frac{1}{p^2}\right)$,其中 p 是 k 的素因数。

定理 2 $|GL_2(Z_k)|=\varphi(k)T_{k,n}=k^4\prod_{p\mid k}\left(1-\frac{1}{p}\right)\left(1-\frac{1}{p^2}\right)$,这里 $\varphi(k)$ 表示 k 的欧拉函数。

1 主要定理的证明

设 k,n 是给定的正整数,满足 $1<k$。若有序四元数组 (a,b,c,d) 满足同余方程 $ab-cd\equiv n(\bmod k)$,则称 (a,b,c,d) 是 $ab-cd\equiv n(\bmod k)$ 的一个解。显然,若 (a,b,c,d) 是同余方程 $ab-cd\equiv n(\bmod k)$ 的解,则 $(a+xk,b+yk,c+zk,d+wk)$ 也是同余方程 $ab-cd\equiv n(\bmod k)$ 的解。因此,在讨论同余方程 $ab-cd\equiv n(\bmod k)$ 的解的时候,一般只考虑 $a,b,c,d\in Z_k$ 的情形,这里 $Z_k=\{0,1,2,\cdots,k-1\}$,即模 k 的剩余类。约定两个四元数组 (a,b,c,d) 和 (x,y,z,w) 相等,当且仅当 $a=x,b=y,c=z,d=w$。这样,可按 $ab-cd\equiv n(\bmod k)$ 的不同解数分类,记 $T_{k,n}$ 是当 n 与 k 互素时同余方程 $ab-cd\equiv n(\bmod k)$ 满足 $a,b,c,d\in Z_k$ 的全部解数。为了计算 $T_{k,n}$,需要以下引理 1、引理 2,这两个引理的证明见参考文献[2]。

引理 1 同余方程 $px\equiv q(\bmod n)$ 当且仅当 $(p,n)\mid q$ 时有解 x,且只有唯一满足 $x\in Z_n$ 的解。

引理 2 若 $(m_i,m_j)=1$, $i\neq j$, 则 $x\equiv a_i(\bmod m_i)$, $1\leqslant i\leqslant n$ 有唯一解 $\bmod m_1,\cdots,m_n$。

(此乃著名的孙子定理,也叫中国剩余定理。)

引理 3 $T_{k,n}$ 是 k 的积性函数。

证明:用 (l,m) 表示两个整数 l 和 m 的最大公因数。设 $k=pq,(p,q)=1,(n,k)=1$。下面证明 $T_{k,n}=T_{p,n}T_{q,n}$。设 $a,b,c,d\in Z_p$ 且满足 $ab-cd\equiv n(\bmod p)$, $x,y,z,w\in Z_q$,且 $xy-zw\equiv n(\bmod q)$。由引理 2 即知,存在唯一的 r,满足 $r\in Z_{pq}$,使得 $r\equiv a(\bmod p)$, $r\equiv x(\bmod q)$;存在唯一的 s,满足 $s\in Z_{pq}$,使得 $s\equiv b(\bmod p)$, $s\equiv y(\bmod q)$;存在唯一的 t,满足 $t\in Z_{pq}$,使得 $t\equiv c(\bmod q)$, $t\equiv z(\bmod q)$;存在唯一的 u,满足 $r\in Z_{pq}$,使得 $u=d(\bmod p)$, $u\equiv(\bmod q)$。显然有序四元数组 (r,s,t,u) 满足 $rs-tu\equiv n(\bmod pq)$,而 (r,s,t,u) 被 (a,b,c,d) 和 (x,y,z,w) 唯一确定,因此 $T_{k,n}=T_{p,n}T_{q,n}$,引理 3 成立。

定理 1 的证明:由引理 3,只需考虑 k 为素数幂时的情形。设 $k=p^e$, p 是素数。设 $a,b,c,d\in Z_k$ 且满足 $ab-cd\equiv n(\bmod k)$。分两种情形讨论:

(1) 当 $(b,p)=1$ 时, $ab-cd\equiv n(\bmod k)$ 可化为 $a\equiv b_1(cd+n)(\bmod k)$,其中 $bb_1\equiv 1(\bmod k)$。由于在 Z_k 中 c 和 d 分别有 k 种取值; b_1 有 $\varphi(k)$ 种取值,而无论 b_1、c、d 取何值,对给定的 n,在 Z_k 中都唯一确定了 a 的值,因此,当 $(b,p)=1$ 时, $T_{k,n}=\varphi(k)kk=p^{3e}\left(1-\frac{1}{p}\right)$。

(2) 当 $(b,p)\neq 1$ 时,令 $b=p^rb_1$, $1\leqslant r\leqslant e$, $(b_1,p)=1$。$ab-cd\equiv n(\bmod k)$ 可化为 $ap^r\equiv b_2(cd+n)(\bmod k)$,其中 $b_2b_1\equiv 1(\bmod k)$。类似于上面的分析,可得该同余式的解数为 $\varphi(p^{e-r})\varphi(p^e)p^{e-r}p^r=\varphi(p^{e-r})p^{2e}\left(1-\frac{1}{p}\right)$。事实上,由 $1\leqslant r\leqslant e$ 及 $(n,p)=1$ 知: $(cd,p)=1$。显然在 Z_k 中 b_2 有 $\varphi(p^{e-r})$ 种取值, c 有 $\varphi(k)$ 种取值;而由引理 1 知:若 c 确定了,则在 Z_{p^r} 中,由 $cd+n\equiv 0(\bmod p^r)$ 也唯一决定了 d,因此在 Z_k 中 d 可以取 p^{e-r} 个值;而当 b_2,c,d 一旦确定,则 $a\equiv(b_2((cd+n)(\bmod p^r)))(\bmod p^{e-r})$ 也唯一确定了 a 在 $Z_{p^{e-r}}$ 中的值,因此在 Z_k 中 a 可以取 p^r 个值。所以当 $(b,p)\neq 1$ 时, $ap^r\equiv b_2(cd+n)(\bmod k)$ 的解数为 $\varphi(p^{e-r})\varphi(p^e)p^{e-r}p^r=\phi(p^{e-r})p^{2e}\left(1-\frac{1}{p}\right)$。

关于 $1\leqslant r\leqslant e$ 加起来得:

$$T_{k,n}=\sum_{r=1}^{n}\phi(p^{e-r})p^{2e}\left(1-\frac{1}{p}\right)=p^{3e-1}\left(1-\frac{1}{p}\right)。$$

综合(1)、(2)知:当 $K=p^e$ 时, $T_{k,n}=T_{p^e,n}=p^{3e}\left(1-\frac{1}{p^2}\right)$。

由引理3即知：当 $k>1$ 时，

$$T_{k,n}=k^3\prod_{p\mid k}\left(1-\frac{1}{p^2}\right)。$$

定理2的证明： 从定理1的证明过程中可以看出，当 $n,m\in Z_k$ 且都与 k 互素时，同余方程 $ab-cd\equiv n(\bmod k)$ 满足 $a,b,c,d\in Z_k$ 的全部解数与同余方程 $ab-cd\equiv m(\bmod k)$ 满足 $a,b,c,d\in Z_k$ 的全部解数相同。因此定理2成立。

例： 令 $k=26$，那么 $|GL_2(Z_{26})|=26^4\times\frac{1}{2}\times\frac{3}{4}\times\frac{12}{13}\times\frac{168}{169}=157\ 248$。

2 结 论

初步综述了一些著名密码体制的密钥空间大小的计算，并给出了 Hill 密码的2阶可逆加密方阵个数（即密钥空间大小）的表达式。虽然通过四元二次同余方程 $ab-cd\equiv n(\bmod k)$ 的解数的计算表达式可求出 Hill 密码的2阶可逆加密方阵个数，但是如何求出 Hill 密码的高阶可逆加密方阵个数仍然是一个困难的问题。如果把明文字母按照先后顺序每 $l(l>2)$ 个分一组，记任意一组为 $(a_1,\cdots,a_l)$，那么 Hill 密码的加密矩阵 A 就是一个 $l\times l$ 方阵，A 的元素为模 k 的剩余类中的元。显然，当且仅当 A 的行列式值与模 k 互素时，才能正确加解密。记在模 k 的剩余类中全体可逆的 $l\times l$ 矩阵的集合为 $GL_l(Z_k)$。显然，随着 l 的增大，求解同余方程 $|A|\equiv n(\bmod k)$ 的解数的难度也随着增大，从而计算 $|GL_l(Z_k)|$ 也就更困难。因此，如何找一个有效的算法来求出 $|A|\equiv n(\bmod k)$ 的解数，从而计算 $|GL_l(Z_k)|$ 还有待进一步探讨。另一方面，值得注意的是，对于有些密码体制（例如 Hill 密码）来说，即使其密钥空间不是很大，用穷举攻击破译密码体制本身仍然十分困难，主要原因是，很难确定每一个可能密钥的具体表达式。

参 考 文 献

1 冯登国，裴定一. 密码学导引. 北京：科学出版社，1999：8—10

2 华罗庚. 数论导引. 北京：科学出版社，1975

On the Computation of the Key Space of Hill Cipher

DAI Jing-guo, ZHANG Shao-hua[1]*, HU Yu-ping, YANG Si-qing

(Department of Computer Science, Hunan University of Humanities and Science and Technology, Loudi 417000, China;
Center for Information Security Study in Wuhan Maritime Communications Research Institute[1], Wuhan 430079, China)

[Abstract] These methods about computation of the key space based on some famous cryptosystem were summarized. The method about calculation the number of 2-order reversible encryption matrix (size of key space) based on Hill cipher was expatiated with emphasis, A computational formula of the number of the key space of order 2 of Hill cipher is given.

[Key words] Key space Hill cipher congruence solutions of the quadratic congruence