高密度背包型公钥密码体制的设计

王保仓 胡予濮
(西安电子科技大学计算机网络与信息安全教育部重点实验室 西安 710071)

摘 要 该文提出了一类新的易解背包问题，基于此问题构造了一个新的加法背包型公钥密码体制。该公钥密码体制具有较高的背包密度，因此可以抵抗低密度子集和攻击。对该密码体制的其它的攻击方法进行了分析。

关键词 公钥密码体制，陷门背包，低密度子集和攻击，格基规约
中图分类号：TP309 **文献标识码**：A **文章编号**：1009-5896(2006)12-2390-04

Knapsack-Type Public-Key Cryptosystem with High Density

Wang Bao-cang Hu Yu-pu
(*Key Laboratory of Computer Networks & Information Security, Ministry of Education, Xi'an* 710071, *China*)

Abstract This article proposes a new easy knapsack problem, based on which a novel knapsack-type public key cryptosystem is derived. The cryptosystem obtains a high knapsack density, and hence it is secure against low density subset-sum attack. Some other attacks on the scheme are also analyzed.

Key words Public-key cryptosystem, Trapdoor knapsack, Low density subset-sum attack, Lattice basis reduction

1 引言

自从Merkle和Hellman[1]提出第一个背包型公钥密码以来，许多陷门背包被提了出来。背包型公钥密码的设计极大地丰富了公钥密码，在陷门背包的发展过程中，人们使用了各种各样的技术来设计陷门背包。比如，使用加法背包的公钥密码，使用紧凑背包的公钥密码，使用二次背包即矩阵覆盖的公钥密码，使用多项式的公钥密码，使用乘法背包的公钥密码，使用模背包的公钥密码，使用丢番图方程的公钥密码等。背包型公钥密码由于其加解密速度快而备受关注，然而，现有的背包型公钥密码几乎都被证明是不安全的。在对背包型公钥密码的分析过程中，主要的分析工具就是低密度子集和攻击[2]以及丢番图逼近攻击[3]等。

影响加法陷门背包的安全性的一个重要的参数是背包密度，背包密度小于0.9408的陷门背包都容易遭受低密度子集和攻击[2]。因而，要设计安全的背包型公钥密码，其背包密度必须大于0.9408。如果背包密度大于1，则该体制在实现过程中就会存在解密不唯一的问题。因此，要设计安全的背包型公钥密码必须兼顾到背包密度不能太小以及解密必须唯一这两个要求。Chor-Rivest密码体制[4]为我们提供了一种兼顾这两个要求的方法：在明文空间中只选取一个小子空间作为合法的明文，即为明文添加一些冗余信息使得解密者能够从数个解密后的明文中选出正确的明文来。然而，即便如此，Chor-Rivest密码体制仍然是不安全的[5]。

本文从一个新的易解背包问题出发来构造一个公钥密码，该公钥密码具有较高的背包密度，因而能够抵抗低密度子集和攻击。本文在第2节给出一个新的易解背包问题并且给出使用该问题构造的公钥密码，第3节对该公钥密码进行安全性分析。

2005-04-18 收到，2005-09-19 改回
国家自然科学基金(60473029)资助课题

2 一个新的背包型公钥密码

陷门背包的设计思想几乎总是从一个简单的背包问题出发来进行构造：把易解的背包问题伪装成一个看似困难的背包问题，这一伪装的方法就是陷门信息。合法的接收者Alice 由于掌握了陷门信息因而能够把该问题恢复成一个易解背包问题，通过求解该易解背包问题，Alice 能够重构明文；而对于非法的接收者 Eve 来说，他从密文恢复明文就意味着求解一个困难的背包问题。

2.1 一个新的易解背包问题

就笔者所知，用于设计公钥密码的易解背包问题目前总共有3类：超递增背包序列[1]，Graham-Shamir公钥密码所使用的背包问题[6]以及对基于丢番图方程的公钥密码进行攻击所使用的背包问题[7]。本节给出一个新的易解背包问题。

定理 1 设$\boldsymbol{A}=(a_1,\cdots,a_n)$为背包向量，记$d_1=a_1$，$d_i=\gcd(a_i, d_{i-1})$，$d_n=\gcd(a_n, d_{n-1})=1$，$\boldsymbol{D}=(d_1,\cdots,d_n=1)$。如果$k\le d_{i-1}/d_i$，$i=1,\cdots,n$，则广义背包问题：

$$\sum_{i=1}^{n} a_i x_i = s,\quad 0\le x_i\le k-1 \tag{1}$$

是易解的而且至多有一个解；如果该方程有解，唯一的一个解为

$$x_n \equiv s a_n^{-1} \left(\bmod d_{n-1}\right)，\text{其中 } a_n a_n^{-1}\equiv 1\left(\bmod d_{n-1}\right)$$

$$x_i \equiv \left[\left(s-\sum_{j=i+1}^{n} a_j x_j\right)\Big/ d_i\right]\left(a_i/d_i\right)^{-1}\left(\bmod d_{i-1}/d_i\right)，\text{其中}$$

$$\left(a_i/d_i\right)\left(a_i/d_i\right)^{-1}\equiv\left(\bmod d_{i-1}/d_i\right)，\ i=2,\cdots,n$$

$$x_1=\left(s-\sum_{j=i+1}^{n} a_j x_j\right)\Big/ a_1$$

证明　设方程式(1)有解。注意到$d_{n-1}|a_i$，$i=1,\cdots,n$，因此，式(1)mod d_{n-1}就有

$$a_n x_n \equiv s \pmod{d_{n-1}}$$

再注意到$d_n=\gcd(a_n, d_{n-1})=1$，就有 $x_n \equiv s a_n^{-1} \pmod{d_{n-1}}$。假设$x_{i+1},\cdots,x_n$都已经求解出来了，现在来求解$x_i$。把$x_{i+1},\cdots,x_n$的求解结果带入到式(1)就有

$$\sum_{k=1}^{i} a_k x_k = s - \sum_{j=i+1}^{n} a_j x_j \tag{2}$$

注意到$d_i|a_k$，$k=1,\cdots,i$，式(2)左边能被d_i整除。式(1)有解，因此式(2)右边也能被d_i整除。两边同时除以d_i就有

$$\sum_{k=1}^{i} (a_k/d_i) x_k = \left(s - \sum_{j=i+1}^{n} a_j x_j\right) \Big/ d_i \tag{3}$$

对于$k=1,\cdots,i-1$，都有$d_{i-1}|a_k$，$d_i|d_{i-1}$，$d_i|a_k$，因此，$(d_{i-1}/d_i)|(a_k/d_i)$，式(3) mod d_{i-1}/d_i就有

$$(a_i/d_i) x_i \equiv \left(s - \sum_{j=i+1}^{n} a_j x_j\right) \Big/ d_i \pmod{d_{i-1}/d_i}$$

再注意到$d_i=\gcd(a_i, d_{i-1})$，$\gcd(a_i/d_i, d_{i-1}/d_i)=1$，就有

$$x_i \equiv \left[\left(s - \sum_{j=i+1}^{n} a_j x_j\right) \Big/ d_i\right] (a_i/d_i)^{-1} \pmod{d_{i-1}/d_i}$$

x_1的求解是显然的。

从上述的计算方法可以看出：该求解方法可以在关于变量个数 n 的多项式时间内完成，因此这种类型的背包问题是一个易解问题；而且，在限定 $0 \le x_i \le k-1$ 的情况下，式(1)至多有一个解。　　证毕

定理 1 给出了一个新的易解背包问题，这一问题与文献[6,7]中的背包问题以及与基于超递增背包向量的背包问题都不相同。这一易解背包问题巧妙地把一个超递增背包向量$\boldsymbol{D}=(d_1,\cdots,d_n=1)$ “揉入”到了背包向量$\boldsymbol{A}=(a_1,\cdots,a_n)$中。这就是下边的定理。

定理 2　设$\boldsymbol{A}=(a_1,\cdots,a_n)$为背包向量且满足定理 1 中的条件，$\boldsymbol{D}=(d_1,\ldots,d_n=1)$为定理 1 中所定义的向量，则$(d_n,\cdots,d_1)$构成一个超递增序列。

证明　就是证明对于 $i=n-1,n-2,\cdots,1$，$d_i > \sum_{j=i+1}^{n} d_j$，

$d_{n-1} > d_n = 1$ 显然。假设已经证明了 $d_i > \sum_{j=i+1}^{n} d_j$，只需证明 $d_{i-1} > \sum_{j=i}^{n} d_j$ 即可。事实上，由 $k \le d_{i-1}/d_i$ 知，

$$d_{i-1} \ge k d_i \ge 2 d_i > d_i + \sum_{j=i+1}^{n} d_j = \sum_{j=i}^{n} d_j$$　　证毕

在构造新的背包体制之前，先给出下述结论。

定理 3　设 U={14,17,19,22,23,26,28,29,30, 31,34,37,38, 39,40,41,42,43,44,46,47,48}，$V=\{s|s\in Z,\ a>50\}$，$W=U\cup V$。任取 W 中元素 d，对任意不同的 i, $j=0,\cdots,7$，则 $i^2 \ne j^2 \pmod{d}$。

证明　对 U 中的每个数逐一检验即知结论对于 U 中的数成立；若 d 取自 V，结论显然成立。　　证毕

结合定理 1 和定理 3，就有如下结果。

定理 4　设$\boldsymbol{A}=(a_1,\cdots,a_n)$为背包向量，记$d_1=a_1$，$d_i=\gcd(a_i, d_{i-1})$，$d_n=\gcd(a_n, d_{n-1})=1$，$\boldsymbol{D}=(d_1,\cdots,d_n=1)$。如果 $d_{i-1}/d_i \in W$，$i=2,\cdots,n$，则广义背包问题

$$\sum_{i=1}^{n} a_i x_i = s,\quad 0 \le x_i \le 7$$

是易解的而且至多有一个解；如果该方程有解，唯一的一个解满足

$$x_n^2 \equiv s a_n^{-1} \pmod{d_{n-1}},\ \text{其中}\ a_n a_n^{-1} \equiv 1 \pmod{d_{n-1}}$$

$$x_i^2 \equiv \left[\left(s - \sum_{j=i+1}^{n} a_j x_j^2\right) \Big/ d_i\right] (a_i/d_i)^{-1} \pmod{d_{i-1}/d_i},\ \text{其中}$$

$(a_i/d_i)(a_i/d_i)^{-1} \equiv \pmod{d_{i-1}/d_i}$，$i=2,\cdots,n$

$$x_1^2 = \left(s - \sum_{i=2}^{n} a_i x_i^2\right) \Big/ a_1$$

事实上，一旦求解出了 $x_i^2 \pmod{d_{i-1}/d_i}$，唯一的 $0 \le x_i \le 7$ 就能求解出来了。

2.2　新的背包型公钥密码的设计

正如引言中所述，背包型公钥密码体制设计的关键就在于如何兼顾背包密度不能太小以及解密必须唯一这两个要求。本节利用 2.1 节的结果构造一个新的背包体制。

密钥生成：随机选取背包向量$\boldsymbol{A}=(a_1,\cdots,a_n)$，记$d_1=a_1$，$d_i=\gcd(a_i, d_{i-1})$，$d_n=\gcd(a_n, d_{n-1})=1$，$\boldsymbol{D}=(d_1,\cdots,d_n=1)$。而且满足 $d_{i-1}/d_i \in W$。随机选取模数 $M > 49\sum_{i=1}^{n} a_i$，以及与M互素的乘法因子w: $0<w<M$。计算 $b_i \equiv a_i w \pmod{M}$ 以及w关于mod$M$的逆元w^{-1}。$\boldsymbol{B}=(b_1,\cdots,b_n)$为公钥。向量$\boldsymbol{D}=(d_1,\cdots,\ d_n=1)$，$w^{-1}$和$M$为私钥。

加密　八进制明文$\boldsymbol{X}=(x_1,\cdots,x_n)$，$0 \le x_i \le 7$ 被加密为

$$c = \sum_{i=1}^{n} b_i x_i^2$$

解密　首先计算

$$s \equiv c w^{-1} \equiv \sum_{i=1}^{n} w^{-1} b_i x_i^2 \equiv \sum_{i=1}^{n} a_i x_i^2 \pmod{M}$$

注意到 $M > 49\sum_{i=1}^{n} a_i$，此时就有 $s = \sum_{i=1}^{n} a_i x_i^2$，根据定理 4 求解该方程就恢复出了明文$\boldsymbol{X}=(x_1,\cdots,x_n)$。

2.3　参数选取的进一步说明

(1) 为提高安全性，也可以对向量 $\boldsymbol{B}$ 进行随机置换并且把置换后的向量作为公钥。

(2) 选取 n=120。

(3) 背包向量$\boldsymbol{A}=(a_1,\cdots,a_n)$的选取。随机从$U$={14,17,19, 22,23,26,28,29,30,31,34,37,38,39,40,41,42,43,44,46,47,48}中选出(可以重复) n个数$g_1,\cdots,g_n$，令 $d_i = \prod_{k=i}^{n} g_k$，然后再随机选取与$g_1,\cdots,g_n$都互素的$n-1$ 个数$h_2,\cdots,h_n$。令$a_1=d_1$，$a_i=h_i d_i$，$i=2,\cdots,n$。$h_2,\cdots,h_n$的选取要使得$a_1,\cdots,a_n$具有相同的比特长度。显然此时的$\boldsymbol{A}=(a_1,\cdots,a_n)$就满足密钥生成中的要求。

2.4　算法的计算复杂度

设 M 的二进制长度为 k，向量 $\boldsymbol{B}$ 和 $\boldsymbol{A}$ 中的数的二进制长度大约也是 k。注意到在加密过程中，只需要计算 $c=\sum_{i=1}^{n} b_i x_i^2$。因此加密算法的复杂度为 $O(k)$，即该算法的加密计算复杂度为线性的，而现在最常用的公钥加密算法

RSA[8]和El Gamal[9]的加密计算复杂度均为 3 次的。在同等条件下，本文提出的加密方案具有更快的加密速度。在解密过程中，解密者首先要计算 1 次模乘法运算 $s\equiv cw^{-1}\pmod M$，其计算复杂度为$O(k^2)$，使用定理 4 求解明文$\boldsymbol{X}$除了进行简单的查表和加减运算外(这些运算的计算复杂度为$O(k)$)，还要计算 $2n$ 个乘法运算 $a_ix_i^2$，$\left[\left(s-\sum_{j=i+1}^{n}a_jx_j^2\right)\Big/d_i\right](a_i/d_i)^{-1}\pmod{d_{i-1}/d_i}$和 n 个除法运算 $\left(s-\sum_{j=i+1}^{n}a_jx_j^2\right)\Big/d_i$。注意到 x_i^2，d_{i-1}/d_i 都是很小的数，因此，$2n$个乘法运算的计算复杂度为$O(nk)$；n个除法运算的计算复杂度为$O(nk^2)$。因此，解密算法的计算复杂度为$O(nk^2)$。当取n=120 时，$M<49\times48^n=49\times48^{120}$，$k\approx\log_2 M\approx676$。解密大约需要 $nk^2\approx5.5\times10^7$ 次运算，而RSA和El Gamal的解密复杂度均为三次的。以RSA-1024 为例，其解密大约需要 $k^3=1024^3\approx1\times10^9$ 次运算。此时，本文的加密方案的解密速度大约是RSA-1024 的 18 倍。

3 安全性分析

本节对该加密体制进行安全性分析。

3.1 暴力攻击

一种显然的暴力攻击模式是：给定密文c，穷举搜索子$\boldsymbol{X}$=$(x_1,\cdots,x_n)$，$0\le x_i\le7$，直到$c=\sum_{i=1}^{n}b_ix_i^2$ 为止。这种穷举搜索成功的概率为 1/8^n。另外一种暴力攻击模式是：给定公钥$\boldsymbol{B}$=$(b_1,\cdots,b_n)$，穷举搜索$(w^{-1},\ M)$使背包向量 $a_i\equiv b_iw^{-1}\pmod M$ 满足定理 4 的条件。假定 M 已知，$w^{-1}\equiv a_1b_1^{-1}\pmod M$，因此，只需穷举搜索$a_1$。根据 2.3 节，这共有$|U|^n$=$22^n$种选择，因此穷举搜索$a_1$成功的概率为 1/$22^n$。

3.2 低密度子集和攻击

低密度子集和攻击的基本思想是，把求解低密度子集和问题与格理论中的基规约算法联系起来。通过低密度子集和问题来构造格基，对该基实施LLL算法[10]或者它的改进算法[11]，找到格基的一个规约基，该规约基的第一个向量很短，这个短向量以很大的概率等于低密度子集和问题的解向量。

低密度子集和问题就是求解 $\sum_{i=1}^{n}a_ix_i=s$，$0\le x_i\le k-1$。构造格基为

$$\boldsymbol{V}=\begin{pmatrix}1&0&\cdots&0&-a_1\\0&1&\cdots&0&-a_2\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1&-a_n\\0&0&\cdots&0&-s\end{pmatrix}$$

$\boldsymbol{V}$的列向量记为$v_1,\cdots,v_{n+1}$，$\boldsymbol{V}$生成了一个n+1 维的格。设$x_1,\cdots,x_n$是该问题的解，则$(x_1,\cdots,x_n,\ 0)$是格中的向量，这是因为 $(x_1,\cdots,x_n,0)=\sum_{i=1}^{n}x_iv_i+v_{n+1}$。而且 $0\le x_i\le k-1$ 说明向量$(x_1,\cdots,x_n,\ 0)$很短，对由$\boldsymbol{V}$生成的格实施格基规约算法所获得的规约基的第一个向量也很短，该向量以一定的概率等于$(x_1,\cdots,x_n,\ 0)$。低密度子集和攻击算法成功的概率依赖于背包密度 $d=n\log_2 k/\log_2\max a_i$。背包密度小于 0.9408 的陷门背包都容易遭受低密度子集和攻击[2]。但背包密度大于 1 时，该攻击算法对背包体制并不构成致命的威胁，比如Chor-Rivest背包体制[4]。

本文的加密方案使用了 $c=\sum_{i=1}^{n}b_ix_i^2$，而这里的二次方程是不能直接用格规约算法来求解的。攻击者为了利用低密度子集和攻击算法对本文的密码体制进行攻击，他就要把 $c=\sum_{i=1}^{n}b_ix_i^2$ 看作一个一次方程。即求解方程：

$$c=\sum_{i=1}^{n}b_iy_i,\quad 0\le y_i\le49$$

如果求解的每个$y_i\in\{0,\ 1,\ 4,\ 9,\ 16,\ 25,\ 36,\ 49\}$，直接令 $y_i=x_i^2$，攻击成功。换言之，在本文的加密方案中，明文空间可以看作$\{0,\ 1,\ 2,\cdots,\ 48,\ 49\}^n$而不是$\{0,\ 1,\ 2,\cdots,\ 6,\ 7\}^n$，在明文空间中只有各个分量全都属于集合$\{0,\ 1,\ 4,\ 9,\ 16,\ 25,\ 36,\ 49\}$的明文向量才是合法的明文。因此本文的密码体制的背包密度应该为

$$d=n\log_2 50/\log_2\max b_i$$

此时$\max b_i$<M，而 $M>49\sum_{i=1}^{n}a_i$ (参看 2.2 节)。为降低密文扩展，通常可以选择 $M\approx49\sum_{i=1}^{n}a_i$。注意到$a_1,\cdots,a_n$具有相同的比特长度，而 $a_1=d_1=\prod_{i=1}^{n}g_i<48^n$ (参看 2.3 节)。因此，$M\approx49na_1$，

$$\max b_i<M<49\times48^n$$

据此可以对背包密度做一个估计：

$$d=n\log_2 50/\log_2\max b_i>n\log_2 50\Big/\log_2\left(49\times48^n\right)$$
$$=\log_2 50\Big/\left(\log_2 48+\frac{1}{120}\log_2 49\right)>1$$

因此，本文的密码体制能够抵抗低密度子集和攻击。

3.3 评述

本文提出的公钥密码体制虽然能够抵抗低密度子集和攻击，然而这并不能说明该体制就是安全的。就像Chor-Rivest密码体制虽然也能够获得一个较高的背包密度，然而它还是不安全的[5]。

4 结束语

本文提出了一类新的易解背包问题并且据此构造了一个背包型密码体制。加密过程中使用了二次方程，从而获得了较高的背包密度，因此可以抵抗低密度子集和攻击。而且，该体制能够抵抗各种暴力攻击。

参 考 文 献

[1] Merkle R C, Hellman M E. Hiding information and signatures in trapdoor knapsacks[J]. *IEEE Trans. on Info. Theory*, 1978, IT-24(5): 525−530.

[2] Coster M J, Joux A, LaMacchia B A, *et al*.. Improved low-density subset sum algorithms[J]. *Computational Complexity*, 1992, 2(2): 111−128.

[3] Lagarias J C. Knapsack public key cryptosystems and Diophantine approximation[C]. Advances in Cryptology, Proceedings of CRYPTO ’83, New York, Plenum, 1984: 3 −23.

[4] Chor B, Rivest R L. A knapsack type public key cryptosystem based on arithmetic in finite fields[J]. *IEEE Trans. on Info. Theory*, 1988, 34(5): 901−909.

[5] Vaudenay S. Cryptanalysis of the Chor-Rivest cryptosystem[J]. *Journal of Cryptology*, 2001, 14(2): 87−100.

[6] Shamir A, Zippel R E. On the security of the Merkle-Hellman cryptographic scheme[J]. *IEEE Trans. on Info. Theory*, 1980, IT-26(3): 339−40.

[7] Laih C S, Gau M J. Cryptanalysis of a Diophantine equation oriented public key cryptosystem[J]. *IEEE Trans. on Commun*., 1997, 46(4): 511−512.

[8] Rivest R L, Shamir A, Adleman L M. A method for obtaining digital signature and public key cryptosystems[J]. *Communications of the ACM*, 1978, 21(2): 120−126.

[9] ElGamal T. A public key cryptosystem and a signature scheme based on discrete logarithms[J]. *IEEE Trans. on Info. Theory*, 1985, IT-31(3): 469−472.

[10] Lenstra A K, Lenstra H W, Lovász L. Factoring polynomials with rational coefficients[J]. *Mathematische Annualen*, 1982, 261(3): 513−534.

[11] Schnorr C. A hierarchy of polynomial time lattice basis reduction algorithm[J]. *Theoretical Computer Science*, 1987, 53(2,3): 201−224.

王保仓： 男，1979 年生，博士生，研究方向为新型快速公钥密码的设计与分析.

胡予濮： 男，1955 年生，教授，博士生导师，研究方向为网络安全与信息安全.