Rao-Nam 私钥密码体制的修正

刘金龙　许宗泽

(南京航空航天大学信息科学与技术学院　南京　210016)

摘　要： 该文提出了一种非查表的错误图样生成算法。该算法通过将可纠正的错误矢量的部分信息嵌入到明文消息中，从而得到比原错误矢量具有更大汉明重量的错误图样。用该算法修正的 Rao-Nam 私钥密码体制高效实用，既无需存储错误图样，又增强了安全性。

关键词： 私钥密码体制，纠错码，错误图样

中图分类号： TP918.3　　**文献标识码：** A　　**文章编号：** 1009-5896(2005)08-1287-03

Modification of Rao-Nam's Private-Key Cryptosystem

Liu Jin-long　　Xu Zong-ze

(*College of Info. Sci. and Tech., Nanjing Univ. of Aeronaut. and Astronaut., Nanjing* 210016, *China*)

Abstract　A new kind of algorithm is proposed which can produce random error vector without looking up the syndrome-error table. An error vector with larger Hamming weight can be produced by mixing a part of the original error vector with the plaintext. A modified Rao-Nam scheme based on the new algorithm is presented, which requires no storage for random error vector. The scheme is efficient and practical, and offers a much higher security level.

Key words　Private-key cryptosystem，Error-correcting code，Error vector

1　引言

1978 年，Berlekamp 等人[1]证明了一般线性码的译码是一个 NP 完全问题. 同年，McEliece[2]根据这一结论构造了一类基于Goppa码的公钥密码体制。与其它的密码体制相比较，McEliece 体制有快速加、解密的优点。从此，纠错码和密码相结合的研究得到迅速发展。

1984 年，Rao[3]提出了一种基于代数编码的私钥密码体制，但是在选择明文和大数表决相结合的攻击算法[4]下，该体制是不安全的。1987 年 Rao 和 Nam [5]提出了一种修正方案，即预先建立伴随式错误表，然后通过查表法进行加、解密。由于错误图样需要占用 $O(n2^{n-k})$比特的存储空间，因此，Rao-Nam 私钥体制对码长和码距作了限制(码距小于等于 6bit，码长最多为 250bit)。

针对 Rao-Nam 密码体制的不足，本文提出了一种非查表的错误图样生成算法，用该算法修正了 Rao-Nam 密码体制，并对修正后的体制的性能和安全性进行了分析。

2　Rao-Nam 私钥密码体制

设 $\boldsymbol{G}$ 是$[n, k, d]$Goppa 码(或其它具有快速译码算法的线性码)的生成矩阵，$\boldsymbol{S}$ 是 $k\times k$ 阶满秩矩阵，$\boldsymbol{P}$ 是 $n\times n$ 阶置换矩阵。$\boldsymbol{G}$, $\boldsymbol{S}$, $\boldsymbol{P}$ 以及伴随式错误表构成 Rao-Nam 密码体制的私钥。加密算法可表示为 $\boldsymbol{C}=(\boldsymbol{mSG}+\boldsymbol{z})\boldsymbol{P}$。其中明文 $\boldsymbol{m}$ 是 GF(2)上的 k 比特向量，密文 $\boldsymbol{C}$ 是 GF(2)上的 n 比特向量，z 是从伴随式错误表中随机选取的错误图样。

由于这个私钥体制的安全性在很大程度上依赖于 z 的选取，所以，伴随式错误表的建立非常重要。Rao-Nam 在文献[5]中给出了下列方法：有限域GF(2)上的 n 维向量空间 $GF^n(2)$关于$[n, k]$线性码 C 有 2^{n-k} 个陪集，每个陪集对应唯一一个伴随式，在每个陪集中选定一个重量大约等于 $n/2$ 的码字，这样就形成 2^{n-k} 个重量大约等于 $n/2$ 的 n 重。当错误图样的重量约等于 $n/2$ 时，码字的每比特位被正确猜测的概率为 1/2，所以大数表决方法不能成功。

对 Rao-Nam 私钥体制已知的较好的攻击方法是 Struik-Tilburg 算法[6]。该算法通过以下 3 步即可获得等价生成矩阵 $\boldsymbol{E}$ 和错误图样集合 Z。

(1) 随机选取明文 $\boldsymbol{x}$，采用相关消息攻击获得所有 N(错误图样数目)个不同的密文 $Y^{(x)}$。令 $Z^{(x)}$是对应的不同错误图样，则 $Y^{(x)}=\{\boldsymbol{x}\boldsymbol{G}+z: z\in Z^{(x)}\}$。构造定向图 $T^{(x)}=(Y^{(x)},Z^{(x)}+Z^{(x)})$，其顶点为 $y_1^{(x)}$，$y_2^{(x)},\cdots,y_N^{(x)}$，从顶点 $y_i^{(x)}$到顶点 $y_j^{(x)}$的边的重量为 $y_i^{(x)}+y_j^{(x)}=z_i^{(x)}+z_j^{(x)}$。

(2) 设 $\boldsymbol{u}_v$ 是第 v 个 k 比特单位坐标向量，$v\in\{1,2,\cdots,k\}$，计算明文 $\boldsymbol{x}_v=\boldsymbol{x}+\boldsymbol{u}_v$。加密每一个 $\boldsymbol{x}_v$，直到获得 N 个不同的密文集 $Y^{(x_v)}$。构造定向图 $T^{(x_v)}=(Y^{(x_v)},Z^{(x_v)}+Z^{(x_v)})$，对每一个 v，计算一个从图 $T^{(x_v)}$ 到 $T^{(x)}$的映射 f'，且该映射保持图的边的重量不变，即 $f'(y_i^{x_v})+f'(y_j^{x_v})=y_i^{(x)}+y_j^{(x)}$。

2004-03-23 收到，2004-07-14 改回

(3) 对于任意的 $v\in\{1,2,\cdots,k\}$，计算 $\boldsymbol{E}$ 的第 v 行：$e_v=y_i^{x_v}+f'(y_j^{x_v})$。计算集合 $Z=\{y-\boldsymbol{xE}: y\in Y^{(x)}\}$。

Struik-Tilburg算法的工作因子主要由第(2)步计算f'的复杂度决定。该算法需要 $O(kN\log N)$次加密运算，工作因子为 $O(kN^2\log N)$比特运算，存储因子为 $O(nN)$比特。这说明当码的 Hamming 重量小于等于 6bit 且码长最多为 250bit 时，N 不是很大，Struik-Tilburg 算法不需要很高的工作因子就能得到码的私钥矩阵和错误图样集合，因此 Rao-Nam 私钥体制是不安全的。

3 错误图样生成算法和修正的 Rao-Nam 私钥体制

设 C 是可以纠正 t 个错误且具有快速译码算法的二进制 $[n, k]$线性码，符号 $\boldsymbol{G}, \boldsymbol{S}, \boldsymbol{P}, \boldsymbol{m}$ 的含义同上，错误图样生成算法如下：

输入 $\boldsymbol{m}, \boldsymbol{G}, \boldsymbol{S}, \boldsymbol{P}$。

步骤 1 随机生成 Hamming 重量小于等于 t 的 n 比特向量 $\boldsymbol{e}$；

步骤 2 取 $\boldsymbol{e}$ 的前 k 比特，设为 $\boldsymbol{e}^*$，如果 $\boldsymbol{e}^*=0$，转步骤 1；

步骤 3 计算 $\boldsymbol{m}+\boldsymbol{e}^*$，如果为 0，转步骤 1；

步骤 4 计算错误图样 z，$z=\boldsymbol{e}^*\boldsymbol{SG}+\boldsymbol{e}$。

输 出 z。

用错误图样生成算法修正后的 Rao-Nam 私钥体制的加密算法可表示为

$$\boldsymbol{c}=((\boldsymbol{m}+\boldsymbol{e}^*)\boldsymbol{SG}+\boldsymbol{e})\boldsymbol{P}$$

解密算法如下：

步骤 1 计算 $\boldsymbol{c}^*$，$\boldsymbol{c}^*=\boldsymbol{cP}^{\mathrm{T}}=(\boldsymbol{m}+\boldsymbol{e}^*)\boldsymbol{SG}+\boldsymbol{e}$；

步骤 2 通过快速译码算法得到错误向量 $\boldsymbol{e}$ 和其前 k 比特向量 $\boldsymbol{e}^*$以及信息组 $\boldsymbol{m}^*$，$\boldsymbol{m}^*=(\boldsymbol{m}+\boldsymbol{e}^*)\boldsymbol{S}$；

步骤 3 计算 $\boldsymbol{m}$，$\boldsymbol{m}=\boldsymbol{m}^*\boldsymbol{S}^{-1}+\boldsymbol{e}^*$。

在修正的 Rao-Nam 私钥体制中，作者建议使用码长大于 250bit，码距大于 20bit 且具有快速译码算法的纠错码。

4 修正的 Rao-Nam 私钥体制的性能分析

4.1 错误图样生成的原理和数目

由纠错码中的标准阵[7]译码方法可知，错误图样均可表示为某个码字与某个陪集首之和。在错误图样生成算法中，错误图样$z=\boldsymbol{e}^*\boldsymbol{SG}+\boldsymbol{e}$可表示为信息组$\boldsymbol{e}^*\boldsymbol{S}$生成的码字与陪集首 $\boldsymbol{e}$ 之和。从该角度说，修正后的体制与原体制的错误图样生成原理相同。

由错误图样生成算法可知，不同的错误矢量 $\boldsymbol{e}$ 对应着不同的错误图样 z。以下，分 3 种情况进行说明。

设 $\boldsymbol{e}_1, \boldsymbol{e}_2$ 是不同的错误矢量，$\boldsymbol{e}_i^*$为 $\boldsymbol{e}_i$的前 k 比特矢量，$\boldsymbol{e}_i'$为长为 n 的矢量(前 k 比特为 0，后 $n-k$ 比特与 $\boldsymbol{e}_i$相同)，i=1, 2；z_1，z_2分别为 $\boldsymbol{e}_1, \boldsymbol{e}_2$对应的错误图样。

(1) 当 $\boldsymbol{e}_1^*=\boldsymbol{e}_2^*$，$\boldsymbol{e}_1'\neq\boldsymbol{e}_2'$时，$z_1-z_2=(\boldsymbol{e}_1^*-\boldsymbol{e}_2^*)\boldsymbol{SG}+(\boldsymbol{e}_1'-\boldsymbol{e}_2')=\boldsymbol{e}_1'-\boldsymbol{e}_2'\neq 0$；

(2) 当 $\boldsymbol{e}_1^*\neq\boldsymbol{e}_2^*$，$\boldsymbol{e}_1'=\boldsymbol{e}_2'$时，$z_1-z_2=(\boldsymbol{e}_1^*-\boldsymbol{e}_2^*)\boldsymbol{SG}+(\boldsymbol{e}_1'-\boldsymbol{e}_2')=(\boldsymbol{e}_1^*-\boldsymbol{e}_2^*)\boldsymbol{SG}\neq 0$；

(3) 当 $\boldsymbol{e}_1^*\neq\boldsymbol{e}_2^*$，$\boldsymbol{e}_1'\neq\boldsymbol{e}_2'$时，根据纠错码理论[6]，在 GF(2)上有 $wt(C_1+C_2)\le wt(C_1)+wt(C_2)$或 $wt(C_1+C_2)\ge wt(C_1)-wt(C_2)$，所以

$$z_1-z_2= \Rightarrow (\boldsymbol{e}_1^*-\boldsymbol{e}_2^*)\boldsymbol{SG}+(\boldsymbol{e}_1'-\boldsymbol{e}_2')$$

$$wt(z_1-z_2)\ge wt((\boldsymbol{e}_1^*-\boldsymbol{e}_2^*)\boldsymbol{SG})-wt(\boldsymbol{e}_1'-\boldsymbol{e}_2')$$

$$\Rightarrow wt(z_1-z_2)\ge wt((\boldsymbol{e}_1^*-\boldsymbol{e}_2^*)\boldsymbol{SG})-(wt(\boldsymbol{e}_1')+wt(\boldsymbol{e}_2'))$$

$$\Rightarrow wt(z_1-z_2)\ge d_{\min}-wt(\boldsymbol{e}_1')\,wt(\boldsymbol{e}_2')$$

$$\Rightarrow wt(z_1-z_2)\ge 2t+1-t-t$$

$$\Rightarrow wt(z_1-z_2)\ge 1$$

这里 $d_{\min}$为最小码距，由以上推导可知，$z_1-z_2\neq 0$。

由于不同的错误矢量 $\boldsymbol{e}$ 可以生成不同的错误图样 z，所以固定矩阵 $\boldsymbol{S}, \boldsymbol{G}, \boldsymbol{P}$，随机产生的错误图样总数为

$$N=\sum_{i=1}^{t}\binom{n}{i}-\sum_{j=1}^{t}\binom{n-k}{j}$$

由以上的表达式可知，当码率不变时，修正方案可以通过增加码长和码距获得足够多的错误图样以适应安全性的要求；或者保持码距不变，增加码长，既可以提高码率又可以增加错误图样数目。例如，使用二进制[255, 179, 21]BCH 码时，$N\approx 2^{58}$；使用二进制[512, 422, 21]Goppa 码时，$N\approx 2^{119}$。

4.2 错误图样 z 和密文 $\boldsymbol{c}$ 的重量

由第 3 节的介绍可知，错误图样生成算法要求 $\boldsymbol{e}^*\neq 0$ 和 $\boldsymbol{m}+\boldsymbol{e}^*\neq 0$，于是在 GF(2)上有

$$z=\boldsymbol{e}^*\boldsymbol{SG}+\boldsymbol{e}$$

$$\Rightarrow \boldsymbol{e}^*\boldsymbol{SG}=\boldsymbol{e}+z$$

$$\Rightarrow wt(\boldsymbol{e}^*\boldsymbol{SG})\le wt(\boldsymbol{e})+wt(z)$$

$$\Rightarrow wt(z)\ge wt(\boldsymbol{e}^*\boldsymbol{SG})-wt(\boldsymbol{e})$$

$$\Rightarrow wt(z)\ge d_{\min}-wt(\boldsymbol{e})$$

$$\Rightarrow wt(z)\ge 2t+1-t$$

$$\Rightarrow wt(z)\ge t+1$$

$$\boldsymbol{c}=((\boldsymbol{m}+\boldsymbol{e}^*)\boldsymbol{SG}+\boldsymbol{e})\boldsymbol{P}$$

$$\Rightarrow wt(\boldsymbol{c})\ge wt((\boldsymbol{m}+\boldsymbol{e}^*)\boldsymbol{SGP})-wt(\boldsymbol{eP})$$

$$\Rightarrow wt(\boldsymbol{c})\ge d_{\min}-wt(\boldsymbol{eP})$$

$$\Rightarrow wt(\boldsymbol{c})\ge 2t+1-t$$

$$\Rightarrow wt(\boldsymbol{c})\ge t+1$$

由以上推导可知，错误图样和密文的 Hamming 重量均大于等于 t+1，所以修正后的 Rao-Nam 私钥体制可以有效地抵制消息重发和大数表决相结合算法的攻击。

4.3 安全性分析

由第 2 节的介绍可知，在 Rao-Nam 私钥体制中，一个明文对应着可能的 2^{n-k} 个不同的密文。理论上，如果有足够的存储容量，即存储的错误图样足够的多，则 Rao-Nam 私钥体制是安全的。但是，在实际的使用中，存储容量总是有限的，所以 Rao-Nam 体制对码长和码距作了限制，即要求码长小于 250bit，码距小于等于 6bit。例如使用[63, 51, 5]BCH 码时，错误图样数目为 2^{12}；使用[127, 113, 5]BCH 码时，错误图样数目为 2^{14}；即便使用[255, 231, 7]BCH 码时，错误图样数目也仅为 2^{24}。由此可见，通过消息重发攻击或相关攻击，Struik-Tilburg 算法容易获得相关向量的全部密文，然后构造相关向量之间的映射关系(定向图)便可得到生成矩阵和错误图样集合。

由以上的分析可知，Rao-Nam 私钥密码体制中错误图样的数目受到存储容量的限制，亦即该体制能够提供的密钥量受到限制，进而影响到该体制的安全性。为了获得更高的安全性能，文献[8,9]提出了一些修正方案，其思想是基于使用结构复杂的非线性码或是修正的错误图样集合，虽然安全性有一定的增强，但是仍然不理想。这些方案未能从根本上增加错误图样数目，以达到足够的安全性。

在本文提出的修正方案中，错误图样无需预先存储，它的生成已嵌入到加密、解密的过程之中。这为该方案使用大码距的长码提供了可行性。显然，当码率不变时，码 C 越长，则错误图样越多，安全性就越高；反之，安全性就越低。这里，作者建议使用码长大于 250bit，码距大于 20bit 且具有快速译码算法的纠错码(如 BCH 码，Goppa 码等)。这样，该体制可以产生足够多的错误图样来满足安全性要求。例如，使用二进制[255, 179, 21]BCH 码时，生成的错误图样数目约为 2^{58}，则 Struik-Tilburg 攻击算法的工作因子约为 $O(2^{129})$比特运算，存储因子约为 $O(2^{66})$；使用二进制[512, 422, 21]Goppa 码时，可产生的错误图样数目约为 2^{69}，则 Struik-Tilburg 攻击算法的工作因子约为 $O(2^{254})$比特运算，存储因子约为 $O(2^{128})$。如此大的存储量要求和异常复杂的映射关系使得 Struik-Tilburg 攻击算法对本文提出的修正方案不再凑效。

目前，作者还未发现其它的针对修正的 Rao-Nam 体制的有效攻击算法。

错误图样生成算法可以推广为 $z=f(\boldsymbol{m},\boldsymbol{e})\boldsymbol{S}\boldsymbol{G}+\boldsymbol{e}$。这里 $f(\boldsymbol{m},\boldsymbol{e})$ 是关于 $\boldsymbol{m}, \boldsymbol{e}$ 的可逆映射，其作用是将错误矢量 $\boldsymbol{e}$ 改变为具有更大 Hamming 重量的错误图样 z，并且能够保证明文的正确恢复。本文使用的是截短函数，简单易行，用其修正的 Rao-Nam 体制可以取得较好的安全性能。

修正的 Rao-Nam 体制的不足之处在于该体制是基于纠错编码理论之上，与其它的纠错密码体制一样存在着密钥规模大和消息扩展现象。针对长密钥问题，文献[10]提出使用短比特序列(称为种子密钥或种子)来确定矩阵 $\boldsymbol{G}, \boldsymbol{S}, \boldsymbol{P}$，在一定程度上解决了纠错密码体制密钥规模大的不足。在实际应用中，可以根据需要将种子密钥和高码率的码结合起来，这样可以在一定程度上弥补修正的 Rao-Nam 体制的不足。

5 结束语

本文针对 Rao-Nam 私钥密码体制存储错误图样需要大量内存的缺陷，提出了一种错误图样生成算法，该算法不仅节省了内存，而且高效实用，用该算法修正的 Rao-Nam 私钥密码体制突破了码长和码距的限制，在码率一定的情况下，使用大距离的长码可以获得很高的安全性。

参 考 文 献

[1] Berlekamp E R, McEliece R J, van Tilborg H C A. On the inherent intractability of certain coding problems. *IEEE Trans. on Information Theory*, 1978, IT-24: 384 – 386.

[2] McEliece R J. A public-key cryptosystem based on algebraic coding theory. DSN Prog. Rep. 1978: 42 – 44, 114 – 116.

[3] Rao T R N. Joint encryption and error correction schemes. Conference Proceedings-11th Annual International Symposium on Computer Architecture, Ann Arbor, Mich, USA., 1984: 240 – 241.

[4] 王新梅, 马文平, 等. 纠错密码理论. 北京: 人民邮电出版社, 2001: 137 – 139.

[5] Rao T R N, Nam K H. Private-key algebraic-code encryptions. *IEEE Trans. on Information Theory*, 1989, 35(4): 829 – 833.

[6] Struik R, Tilburg J. The Rao-Nam scheme is insecure against a chosen-plaintext attack. Advances in Cryptology-CRYPTO '87. New York: Springer-Verlag, 1988: 458 – 461.

[7] 王新梅, 肖国镇. 纠错码—原理与方法. 西安: 西安电子科技大学出版社, 2001: 63 – 63.

[8] Struik R. On the Rao-Nam scheme using nonlinear codes. Information Theory Proceedings. 1991 IEEE International Symposium, Eindhoven University of Technology, 1991: 174 – 174.

[9] Denny W F. Encryptions using linear and nonlinear codes: implementation and security considerations. [Ph.D. dissertation], The Center for Advanced Computer Studies, University of Southwestern Louisiana, Lafayette, 1988.

[10] Sun H M, Hwang T. Key generation of algebraic-code crypto systems. *Computer Math. Application*, 1994, 27(2): 99 – 106.

刘金龙：　男，1976年生，博士生，研究方向：密码学与信息安全.

许宗泽：　男，1940年生，教授，研究方向：数字通信、编码理论与应用.