doi:10.3969/j.issn.1000-436x.2014.05.009

基于枚举错误向量的 McEliece 公钥密码体制攻击方法

刘景美[1]，王延丽[1]，梁斌[1]，赵林森[2]
（1. 西安电子科技大学 综合业务网国家重点实验室，陕西 西安 710071；
2. 西安邮电大学 电子工程学院，陕西 西安 710061）

摘 要：对 McEliece（M）公钥密码体制的安全性进行研究，该体制中错误向量的汉明重量相对于码长较小，而基于 Goppa 码的 M 公钥密码体制存在低重量的公开码字。基于以上分析，提出了枚举错误向量的攻击算法。重点分析了算法中错误翻转比特个数和算法迭代次数等参数对正确解密概率的影响，利用所提算法分析了基于（1024,524,101）Goppa 码的 M 体制安全性。从算法正确解密概率和工作因子 2 个方面进行仿真分析，仿真实验表明所提算法在码重较低的情况下具有优异的性能。

关键词：Goppa 码；McEliece；低重量码字；枚举错误向量
中图分类号：TN918.1 **文献标识码**：A **文章编号**：1000-436X(2014)05-0065-05

McEliece public key cryptosystem attack algorithm based on enumeration error vector

LIU Jing-mei[1], WANG Yan-li[1], LIANG Bin[1], ZHAO Lin-sen[2]
(1. National Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China;
2. College of Electronic Engineering, Xi'an University of Post & Telecommunications, Xi'an 710061, China)

Abstract: The research on the security of McEliece (M) public key cryptosystem was presented. The Hamming weight of error vector is less than the code length, and M public key cryptosystem based on Goppa code possesses low weight public code words. Considering the above analysis, an attack algorithm based on enumeration error vector was proposed. The effect on probability of correct decryption by the numbers of error flipping bits and algorithm iteration was focused on. And the security of (1 024,524,101) Goppa-based M public key cryptosystem was analyzed. Performance analysis of the proposed algorithm from probability of correct decryption and work factor was simulated, and the experimental results show that the proposed algorithm has a good performance when the code weight is low.

Key words: Goppa code; McEliece; low weight code word; enumeration error vector

1 引言

1975 年，由 Berlekamp、MeElieee 及 VanTilborg 证明了线性分组的译码问题是 NP 完全问题，McEliece 利用了这一难题，首次构造了基于 Goppa 码的 McEliece 公钥密码体制。迄今为止已有许多学者对 M 体制的安全性进行了深入的研究[1~5]，提出信息集译码攻击[6,7]、Stern 及其改进算法的攻击[8,9]、信息重发与信息相关攻击[10,11]等。本文针对基于 Goppa 码的 M 体制中存在低重量的码字，提出枚举错误向量的攻击算法，通过仿真可知该算法具有优异的性能。

2 M 公钥密码体制简介

密钥生成过程如下

$$\boldsymbol{G}' = \boldsymbol{SGP} \tag{1}$$

收稿日期：2013-01-07；**修回日期**：2013-04-03
基金项目：国家自然科学基金资助项目(60903199)；高等学校创新引智基地基金资助项目(B08038)；中央高校基本科研业务费专项基金资助项目(K5051201014)
Foundation Items: The National Natural Science Foundation of China(60903199); The 111 Project(B08038); The Fundamental Research Funds for the Central Universities (K5051201014)

其中，$\boldsymbol{G}'$是可以纠正t个错误的$(n, k, 2t+1)$Goppa码的生成矩阵，$\boldsymbol{S}$是$k\times k$非奇异加扰矩阵，$\boldsymbol{P}$是$n\times n$置换矩阵，$\boldsymbol{S}$和$\boldsymbol{P}$都是GF(2)上的非奇异方阵，存在逆矩阵$\boldsymbol{S}^{-1}$和转置矩阵$\boldsymbol{P}^{\mathrm{T}}$($\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{P}^{-1}$)，$\boldsymbol{G}'$是公钥，$\boldsymbol{S}^{-1}$和$\boldsymbol{P}^{\mathrm{T}}$及Goppa码的校验矩阵$H$是私钥。

加密算法：发送方将明文分成长为k的比特组，设u是一k bit的明文，利用接收方的公钥$\boldsymbol{G}'$做如下的运算，其中，$\boldsymbol{e}$是随机生成的汉明重量为t的错误向量。

$$x=u\boldsymbol{G}'+\boldsymbol{e} \tag{2}$$

解密算法：接收端收到密文x后，做如下的运算。

1) 计算

$$x'=\boldsymbol{x}\boldsymbol{P}^{\mathrm{T}}=u\boldsymbol{S}\boldsymbol{G}\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}+\boldsymbol{e}\boldsymbol{P}^{\mathrm{T}}=u\boldsymbol{S}\boldsymbol{G}+\boldsymbol{e}\boldsymbol{P}^{\mathrm{T}}$$

2) 利用私钥校验矩阵$\boldsymbol{H}$纠出错误向量$\boldsymbol{e}\boldsymbol{P}^{\mathrm{T}}$，求出$u'=u\boldsymbol{S}$；

3）计算$u'\boldsymbol{S}^{-1}=u\boldsymbol{S}\boldsymbol{S}^{-1}=u$，恢复出明文$u$。

3 枚举错误向量攻击

由M体制的加密算法可知，错误向量的汉明重量相对于纠错码的码长要小得多，可以通过枚举出所有可能的错误向量来破解密文，错误向量所有可能的取值是$\binom{n}{t}$；对采用(1 024, 524)Goppa码的M体制，$\mathrm{lb}\binom{n}{t}=284$，这个数值是相当大的，通过枚举出所有的错误向量来破解密文是十分困难的，但基于Goppa码的M体制中存在低重量的码字，可以利用如下的算法来解密低重量的码字对应的密文，该算法具有较好的性能。

算法描述如下。

1) 运用高斯行变换，把公钥$\boldsymbol{G}'$变换成$\boldsymbol{G}'=[IZ]$的形式，则变换后的生成矩阵$\boldsymbol{G}'$对应的校验矩阵$\boldsymbol{H}', \boldsymbol{H}'=[\boldsymbol{Z}^{\mathrm{T}}\boldsymbol{I}]$。

2) 假设有低重量的码字（汉明重量等于w）与错误向量$\boldsymbol{e}$中的1在比特位置上没有重叠，选取密文x中未被选过的t个1进行比特翻转得到s（例如x=1001<u>0</u>1，t=2，则s=<u>0</u>000<u>1</u>），计算向量$\boldsymbol{z}=\boldsymbol{s}\boldsymbol{H}'^{\mathrm{T}}$。

3) 如果有j bit被错误翻转，此时s含有的错误比特数是$2j$，其中，j个错误是在上一步的t个1翻转成0的比特位置上，另外的j个错误在s中w个1的比特位上，选取汉明重量等于$2j$的向量$\boldsymbol{e}'$，此时j个1从t个1翻转成0的比特位上选取，另外的j个1从s中的w个1比特位上选取，若$\boldsymbol{z}+\boldsymbol{e}'\boldsymbol{H}'^{\mathrm{T}}=0$，则跳到算法下一步，若$\boldsymbol{z}+\boldsymbol{e}'\boldsymbol{H}'^{\mathrm{T}}\neq 0$，则重复这一步，当$\sum_{j=0}^{i}\left(\binom{t}{j}\binom{w}{j}\right)$个向量$\boldsymbol{e}'$被选取过后，没有新的可选的$\boldsymbol{e}'$，说明在步骤2)中被错误翻转的比特数大于上限$i$，此时，算法跳转到步骤2)，进行下一轮迭代。

4) 设一n bit的向量$\boldsymbol{q}$，有$\boldsymbol{q}=\boldsymbol{s}+\boldsymbol{e}'$，因$\boldsymbol{z}\boldsymbol{H}'^{\mathrm{T}}=\boldsymbol{s}\boldsymbol{H}'^{\mathrm{T}}+\boldsymbol{e}'\boldsymbol{H}'^{\mathrm{T}}=(\boldsymbol{s}+\boldsymbol{e}')\boldsymbol{H}'^{\mathrm{T}}$，所以$\boldsymbol{q}=\boldsymbol{x}-\boldsymbol{e}=u\boldsymbol{G}'$，$\boldsymbol{q}$是低重量的公开码字。随机地选取$\boldsymbol{q}$的$k$比特构成$\boldsymbol{q}_k$，选取公钥$\boldsymbol{G}'$对应的$\boldsymbol{k}$列构成$\boldsymbol{G}'_k$，求其逆矩阵$\boldsymbol{G}'^{-1}_k$，则计算$\boldsymbol{q}_k\boldsymbol{G}'^{-1}_k=\boldsymbol{u}\boldsymbol{G}'\boldsymbol{G}'^{-1}_k$，恢复明文$u$。

下面给出一个例子进行分析，可以清楚地看到攻击过程的运行。

① 假设明文u=100101，对应的公开码字c=100000001100000，错误向量$\boldsymbol{e}$=0000011000000010。（令t=3），密文x=1000011011000010。

② 根据步骤2)，求得$\boldsymbol{s}$=0000000011000010，知j=1。

③ 根据步骤3)，求得$\boldsymbol{e}'$=1000000000000010。

④ 然后按照步骤4)所说，求得明文u=100101。

4 性能分析

1) 概率

低重量码字c与错误向量$\boldsymbol{e}$中1 bit不重叠的概率为

$$p_{ct}=\binom{n-w}{t}\Big/\binom{n}{t} \tag{3}$$

不大于i个比特被错误翻转的概率可表示为

$$p_{j\leqslant i}=\sum_{j=0}^{i}\left(\binom{t}{t-j}\binom{w}{j}\right)\Big/\binom{w+t}{t} \tag{4}$$

正确解密的概率为

$$\begin{aligned}p_i&=p_{ct}\times p_{j\leqslant i}\\&=\left(\binom{n-w}{t}\sum_{j=0}^{i}\left(\binom{t}{t-j}\binom{w}{j}\right)\right)\Big/\left(\binom{n}{t}\binom{w+t}{t}\right)\end{aligned} \tag{5}$$

算法的每次迭代是独立的，每次成功破解密文的概率相同，后面将给出正确解密的概率与参

数间的关系图表，从图表中可以清晰地看出汉明重量 w、错误翻转比特数、迭代次数和正确解密概率的关系。

2) 复杂度

对于某一特定的密文，第一次迭代的复杂度设为 $W_{m=1}$，其中，第 1)步，高斯行变换 G' 的复杂度为 $k\times k/2\times n$；第 2)步，复杂度为 $n(n-k)$；第 3)步，假设 s 有 j 个比特被错误地翻转，e' 中 $2j$ 个 1 对应的 $\boldsymbol{H}'$ 矩阵中 $2j$ 个列相加需要 $(2j-1)(n-k)$ bit 操作，所以算法第 3)步需要的比特操作数 $(n-k)\cdot\sum_{j=0}^{i}\left((2j-1)\binom{t}{j}\binom{w}{j}\right)$；第 4)步，求 $\boldsymbol{G}'_k$ 的逆矩阵，矩阵求逆需要 αk^3 步，求 $\boldsymbol{q}_k\boldsymbol{G}'^{-1}_k$ 需要 k^2 bit 操作。如果算法第一次迭代没有成功破解密文，第 m 次迭代的复杂度设为 $W_{m\geqslant 2}$，此时迭代不需要执行算法的第 1)步；在算法的第 2)步中，让更新的 s 最多在 2 个比特上不同于上一个 s，需要 $\beta(n-k)$ bit 操作，β 的值比较小，可取到 1，这样在算法第 4)步，对应的 $\boldsymbol{G}'_k$ 最多在两列上不同于上一个 $\boldsymbol{G}'_k$，求 $\boldsymbol{G}'_k$ 的逆矩阵的比特操作可以下降到 αk^2。

对于某一特定的密文，第 m 次迭代的复杂度为

$$W_{m=1}=k\times k/2\times n+n(n-k)+(n-k)\sum_{j=0}^{i}\left((2j-1)\binom{t}{j}\binom{w}{j}\right)+\alpha k^3+k^2 \tag{6}$$

$$W_{m\geqslant 2}=\beta(n-k)+\alpha k^2+k^2+(n-k)\sum_{j=0}^{i}\left((2j-1)\binom{t}{j}\binom{w}{j}\right) \tag{7}$$

对某一特定的密文最多迭代 g 次成功[illegible]多错误翻转比特数为 i，此时正确解密[illegible]$P(m\leqslant g)$ 与复杂度 $W(m\leqslant g)$ 为

$$P(m\leqslant g)=\sum_{m=1}^{g}(1-p_i)^{m-1}p_i$$

$$W(m\leqslant g)=\sum_{m=1}^{g}W_m$$

3) 数据仿真

① 在算法的第 2)步，j bit 被错误翻转的概率为

$$p=\left(\binom{t}{t-j}\binom{w}{j}\right)\Big/\binom{w+t}{t} \tag{10}$$

图 1 与图 2 分别给出汉明重量为 101 和 161 的码字在算法第 2)步错误翻转比特的概率分布，从概率分布曲线可知，为得到高的正确解密概率，算法第 3)步中 i 的取值要尽可能大一些；汉明重量为 101 和 161 的概率分布曲线分别在 j 等于 33 和 38 处取得最大值，说明随着码字汉明重量的增加，为得到相同的正确解密的概率，i 的取值也要增加。

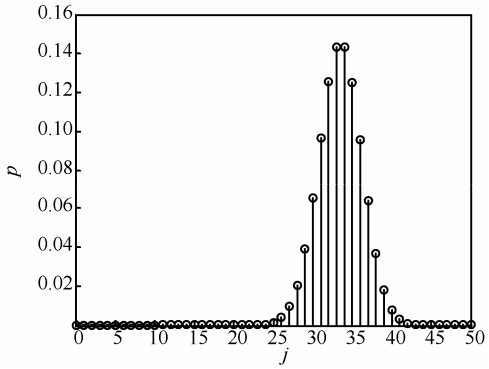


图 1　w=101 错误翻转比特的概率分布

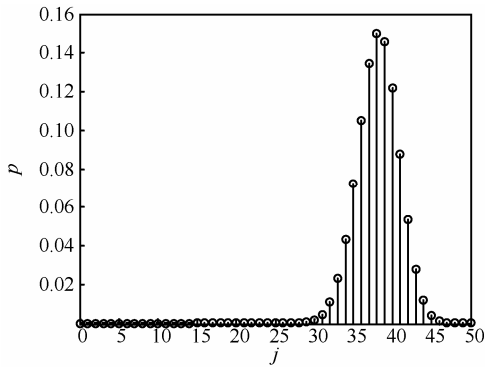


图 2　w=161 错误翻转比特的概率分布

图 3 与图 4 分别给出迭代一次正确解密的概率 $P(m=1)$ 等于 5×10^{-4}、5×10^{-5} 随着迭代次数 g 的增[illegible]解密的概率曲线，可知提高算法的迭代次数[illegible]地提高正确解密的概率，为得到相同正确[illegible]概率值，一次迭代解密的概率值越小，需要[illegible]数越大。

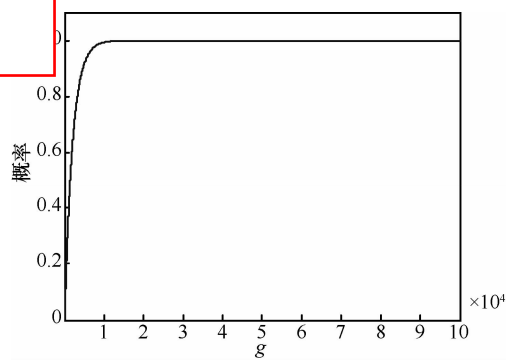


图 3　$P(m=1)=5\times10^{-4}$ 正确解密概率与迭代次数间的关系

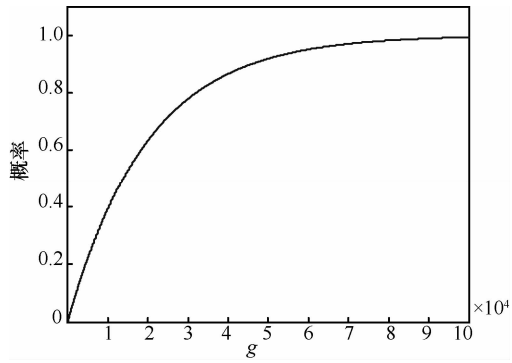


图 4 $P(m=1)=5\times10^{-5}$ 正确解密概率与迭代次数间的关系

表 1 与表 2 分别列出汉明重量等于 101、131、161 的码字在不同的 i 值，不同的迭代次数 g 下，该算法正确解密的概率和相应的工作因子。

表 1 正确解密的概率

w	i	g			
		10	1 000	10 000	100 000
101	20	6.0×10^{-8}	6.0×10^{-6}	6.0×10^{-5}	0.000 6
	30	0.006 7	0.493	**0.998 8**	—
	40	0.047 3	0.992	—	—
131	20	6.6×10^{-11}	6.6×10^{-9}	6.6×10^{-8}	6.6×10^{-7}
	30	0.000 16	0.016 4	0.152 5	0.809
	40	0.008 3	0.569 4	**0.999 7**	—
161	20	1.3×10^{-13}	1.3×10^{-11}	1.3×10^{-10}	1.3×10^{-9}
	30	3.6×10^{-6}	0.000 36	0.003 6	0.035 8
	40	0.001 2	0.117 1	0.712	**1.0**

表 2 工作因子

w	i	g			
		10	1 000	10 000	100 000
101	20	132.4	139.0	142.3	145.6
	30	149.7	156.3	**159.7**	—
	40	152.7	159.4	—	—
131	20	140.5	147.1	150.5	153.8
	30	162.3	169	172.3	175.6
	40	168.3	174.9	**178.3**	—
161	20	146.8	153.5	156.8	160.1
	30	172.1	178.8	182.1	185.4
	40	180.9	187.5	190.8	**194.2**

由表 1 表 2 可知，在迭代次数 g 取值一定的情况下，增大 i 的值，可以显著地提高正确解密的概率，但相应的工作因子上升的幅度也较大；在 i 取值一定的情况下，增大迭代次数 g，可以在一定程度上提高正确解密的概率，同时工作因子上升的幅度比较小；在 i、g 取值相同的情况下，随着码字汉明重量的增加，相应的正确解密的概率值减小，工作因子上升。汉明重量等于 101、131、161 公开码字以近于 1 的概率正确解密对应的工作因子分别是 159.7、178.3、194.2，此时的工作因子相对来说还是较高的。

5 结束语

由于基于 Goppa 码的 M 公钥密码体制存在低重量的公开码字这一不安全的因素，本文提出枚举错误向量攻击算法。从正确解密的概率和工作因子 2 方面分析了该算法的性能，并深入分析了增大算法第 3)步中错误翻转比特 i 值和算法迭代次数 g 的值对提高正确解密概率的影响，为获得较高的正确解密概率，i 与 g 的值都需要设置得较大一些；同时通过仿真给出该算法对于不同汉明重量公开码字的正确解密概率和工作因子的值，由仿真数据可知，该算法对于低重量的码字性能相对较好。该算法不仅适用于基于 Goppa 码的 M 体制，而且适用于一切存在低重量公开码字的基于纠错码的 M 公钥密码体制。

参考文献：

[1] LOIDREAU P, SENDRIER N. Weak keys in the McEliece public-key cryptosystem[J]. IEEE Transactions on Information Theory, 2001, 47(3): 1207-1211.

[2] ADAMS C M, MEIJER H. Security-related comments regarding McEliece's public-key cryptosystem[J]. IEEE Transactions on Information Theory, 1989, 35(2): 454-455.

[3] BARBIER M, BARRETO P S L M. Key reduction of McEliece's cryptosystem using list decoding[A]. Information Theory Proceedings (ISIT), 2011 IEEE International Symposium IEEE[C]. 2011. 2681-2685.

[4] BERNSTEIN D J, LANGE T, PETERS C. Wild Mceliece Incognito[M]. Berlin Heidelberg: Springer, 2011.

[5] MISOCZKI R, BARRETO P S L M. Compact McEliece keys from Goppa codes[A]. Selected Areas in Cryptography[C]. Springer Berlin Heidelberg, 2009.376-392.

[6] LEE P, BRICKELL E. An observation on the security of McEliece’s public-key cryptosystem[A]. Cryptology EUROCRYPT’88[C]. Davos, Switzerland, 1988.275-280.

[7] MCELIECE R J. A public-key cryptosystem based on algebraic[J]. Coding Thv, 1978, 4244: 114-116.

[8] GORDON R K, TEW M D, ELSHERBENI A Z. An efficient finite difference method for finding the electric potential in regions with small perturbations[A]. Antennas and Propagation Society Interna-

tional Symposium, 1992.AP-S. 1992 Digest. Held in Conjunction with: URSI Radio Science Meeting and Nuclear EMP Meeting[C]. Chicago IL, USA, IEEE, 1992.524-527.

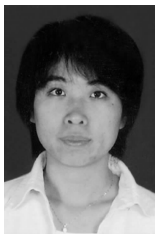
[9] HIROTOMO M, MOHRI M, MORII M. A probabilistic computation method for the weight distribution of low-density parity-check codes[A]. Information Theory 2005 ISIT 2005. Proceedings. International Symposium on[C]. Adelaide, SA, IEEE, 2005.2166-2170.

[10] BERSON T. Failure of the McEliece public-key cryptosystem under message-resend and related-message attack[A]. Cryptology CRYPTO'97[C]. California, USA, 1997.213-220.

[11] SENDRIER N. McEliece Public Key Cryptosystem[M]. Encyclopedia of Cryptography and Security, Springer Berlin Heidelberg, 2005.375-378.2011,767-768.

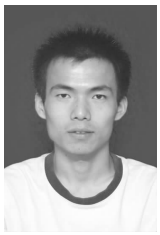
[12] SUN H. Improving the security of the McEliece public-key cryptosystem[A]. ASIACRYPT'98[C]. 1998.200-213.

作者简介：



刘景美（1979-），女，山东烟台人，博士，西安电子科技大学副教授，主要研究方向为密码分析。

王延丽（1987-），女，山东兖州人，西安电子科技大学硕士生，主要研究方向为密码分析。



梁斌（1989-），男，河南焦作人，西安电子科技大学硕士生，主要研究方向为密码分析。

赵林森（1977-），男，山东聊城人，西安邮电大学讲师，主要研究方向为网络安全与数字通信。

（上接第 64 页）

schedule for packet radio networks[J]. Information Processing Letters, 1995, 55(5): 291-295.

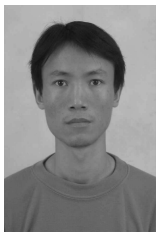
[8] LIU Y, ZHANG L, LI V O K, *et al*. Topology-transparent scheduling in mobile ad hoc networks supporting heterogeneous quality of service guarantees[A]. 2012 46th Annual Conference on Information Sciences and Systems (CISS)[C]. Princeton, NJ, 2012. 1-6.

[9] 李西洋, 范平志. 支持多类业务的移动 ad hoc 网络拓扑透明 MAC 调度码[J]. 计算机应用, 2012, 32(9): 2400-2404.
LI X Y, FAN P Z. Topology-transparent MAC scheduling code for mobile ad hoc network supporting multi-class services[J]. Journal of Computer Applications, 2012, 32(9): 2400-2404.

[10] JU J H, LI V O K. TDMA scheduling design of multihop packet radio networks based on latin squares[J]. IEEE Journal on Selected Areas in Communications, 1999, 17(8): 1345-1352.

[11] BRUALDI R. Introductory Combinatorics[M]. Upper Saddle River, NJ: Prentice-Hall, 1999.

作者简介：



李西洋（1979-），男，湖北荆州人，西南交通大学博士生，主要研究方向为序列设计和多址接入理论。

范平志（1955-），男，四川广汉人，博士，西南交通大学教授、博士生导师，主要研究方向为信息与编码、多址接入和高移动性无线通信理论。