第32卷 第16期
Vol.32 № 16
计 算 机 工 程
Computer Engineering
2006年8月
August 2006
·安全技术· 文章编号：1000—3428(2006)16—0118—02 文献标识码：A 中图分类号：TP309

RSA 算法中几种可能泄密的参数选择

谢建全[1,2]，阳春华[1]
(1. 中南大学信息科学与工程学院，长沙 410083；2. 湖南财经高等专科学校，长沙 410205)

摘 要：RSA 加密算法是目前使用较多、安全性高的一种非对称加密算法，在实际应用中要使该算法有较高的防破解强度，在大素数的选择上是有要求的。文章给出了选择高质量的大素数的有效方法，并对一些不当的选择可能造成的泄密给出了相应的证明。
关键词：大素数；RSA 算法；安全性；攻击

Several Possible Parameters Options Causing Encryption Failure in RSA Algorithm

XIE Jianquan[1,2], YANG Chunhua[1]
(1. Information Engineering College, Central South University, Changsha 410083; 2 Hunan Finance and Economics College, Changsha 410205)

【Abstract】 RSA encryption algorithm is a popular unsymmetrical encryption with high secure character. There are some special requirements in the selection of the large prime number in order to prevent the decryption. An efficient method has been introduced, which can be used to select the large prime numbers, and the some reasons which cause lower security by selecting incorrect prime number have been proved.
【Key words】 Big primes; RSA algorithms; Security; Attack

1 RSA 算法的基本描述

RSA 算法是公钥密码体制的典型算法，它具有密钥管理方便、破译难度大的优点，因而在网络中常用于敏感信息的加密。RSA 算法的具体操作是：选取 2 个大素数 p 与 q，然后算出 $m=pq$，$\varphi(m)=(p-1)(q-1)$，再选取一个正整数 e，使之满足$(e, \varphi(m))=1$，$1<e<\varphi(m)$，其中()表示求最大公约数；再求出正整数 d，使之满足 $1<d<\varphi(m)$，且使 $de\equiv 1 \bmod \varphi(m)$。然后用$\{m, e\}$作为公钥，而用$\{m,d\}$作为私钥。

用 RSA 体制加密时，先将明文数字化，然后进行分组，每组的长度不超过 $log(m)$，再每组单独加密和解密。加密过程如下：假设要加密的明文组为 $n(0\leq n<m)$，将 n 的 e 次方用 m 除后得到余数 c，c 为密文，这就是加密过程；将密文 c 的 d 次方用 m 除后得到的余数刚好为 n，即明文，这就是解密过程。

2 可能导致 RSA 失密的几种大素数的选择

如果要对 RSA 算法加密后的信息进行破译，最直接的方法是根据公开密钥$\{m, e\}$中的 m，对 m 进行因子分解，算出 p 和 q，再算出 $\varphi(m)$和 d。

实际应用中 m 是一个相当大的数，对大数的因子分解目前是难解的，这是 RSA 公钥密码体制建立的基础。因为迄今没有找到一个有效算法可用于分解因子；只要 p 和 q 足够大，用计算机进行穷举，一般在信息的有效期内也无法找出 p 或 q，如果 p 和 q 的长度为 100 位时，若用一台每秒能进行 1 亿次因子分解的高速计算机来做分解，其所需时间均为 3 800 000 年。

如果不通过对 m 进行因子分解能算出 p、q 或 d，也能对 RSA 算法加密后的信息进行破译，因为只须知道 p、q、$\varphi(m)$ 与 d 之中的一个，就可推导出其它的数。因此有时要对 RSA 加密进行破解，不一定要直接分解 m，也就是说，RSA 体制的安全性并非是无条件的。只有选择满足某些条件的高质量 p、q 及 d，才能使 RSA 的安全性尽量基于大数因子分解难上。否则，大数因子分解难，但 RSA 不难攻破。

要使 RSA 加密后的密文不被破译，p、q 和 d 除了要选用大素数外，还应满足如下几个条件：(1)p 和 q 必须为强素数；(2)p 与 q 的长度应该相差不多，p 与 q 之差不能太小；(3)$p-1$ 和 $q-1$ 的最大公因子应该很小；(4)一个模数 m 只能被一个人使用；(5)e 的值不能过小，且不能为 $\log_n(km+n)$，式中 n 为明文，k 为任意正整数。

3 p 和 q 选择不当可能导致失密的证明

3.1 p 和 q 必须为强素数的证明

强素数 x 的条件：

(1)存在 2 个大素数 x_1 和 x_2，使 $x_1|(x-1)$, $x_2|(x+1)$；

(2)存在 4 个大素数 r_1，s_1，r_2 及 s_2，使 $r_1|(x_1-1)$，$s_1|(x_1+1)$，$r_2|(x_2-1$，$s_2|(x_2+1)$。

采用强素数的理由如下：若 $x-1=\prod_{i=1}^{t} x_i^{a_i}$，$x_i$ 为素数，a_i 为正整数。分解式中 $x_i<B$，B 为已知一个小整数，则存在一种 $x-1$ 的分解法，使我们易于分解 m。令 $m=xy$，且 $x-1$ 满足上述条件，$x_i<B$。设 a_i 的最大值为 a，$i=1, 2,\ldots,t$，即可构造

$$R=\prod_{i=1}^{t} x_i^{a}$$

显然$(x-1)|R$。由费尔马定理 $2^R\equiv 1 \bmod x$，令 $2^R\equiv r \bmod m$，若 $r=1$，则选 3 代 2，直到出现 $r\neq 1$。此时，由 $\gcd(x-1,m)=x$，

作者简介：谢建全(1964—)，男，教授，主研方向：计算机安全策略与效率；阳春华，教授、博导
收稿日期：2005-11-29 **E-mail**：xiejianquan@sina.com

就得到 m 的分解因子 x 和 y。

3.2 p 与 q 之差要大的证明

m 是正奇整数，若 $m=x*y$，则 m 可写成 u^2-v^2，假定 $a>b$，则：$u=(x+y)/2$，$v=(x-y)/$，所以 u 和 v 为非负整数。

反过来 m 是正奇数，m 可表示为 $m=u^2-v^2=(u+v)(u-v)$，故 $x=u+v$，$y=u-v$，这里可提供一种因数分解办法。

因若 $m=xy$ 且 x、y 比较接近，则 v 将是一个很小的数，而 u 比 $\sqrt{m}$ 稍大，这样 u 可从开始每次递增 1($\sqrt{m}+1$，$\sqrt{m}+2$，...)，直到找到 $u^2-n=v^2$ 为止。由于 u 只比 $\sqrt{m}$ 稍大，只要进行几次运算就能很快确定 u 和 v。例如取两个相差小的素数 401 和 409，则 m=164 009，$\sqrt{m}\approx 405$，$405^2-m=16$，可得 $v=4$，从而确定 $u=405$，进而得到 $x=409$，$y=401$。

注意不能为保证 p 与 q 之差大而过于拉大 p 与 q 的长度差，因为针对 m 的穷举搜索难度取决于 p 与 q 的最小值，比如 p 与 q 的长度分别为 150 位和 50 位，m 虽有 200 位，但此时的攻击难度仅相当于 100 位。在实际应用中，p 与 q 的长度应只相差几位。

3.3 p-1 和 q-1 的最大公因子要小的证明

根据 RSA 算法的原理，保密的 $\varphi(m)=(p-1)(q-1)$是用于计算 e(或 d)的，故在 RSA 中具有关键作用。一旦 $\varphi(m)$被攻破，则 p、q 与 d 均很容易被攻破(例如，对 $\varphi(m)$代数运算可求 P 与 q，由 $de\equiv 1(\mathrm{mod}(\varphi(m)))$可求 e 或 d)，因此必须加强 $\varphi(m)$的抗攻击性。

若 $p-1$ 和 $q-1$ 的最大公因子较小，由 Euler 定理可知，$\varphi(m)$大，则 $\varphi(m)$的执攻击性就强，也就是说，由于 $P-1$ 与 $q-1$ 必是大偶数，其最小素因子必均为 2，若 $P-1$ 与 $q-1$ 的最大公因子为 2，则 $\varphi(m)=(p-1)(q-1)/2$，则对其穷举攻击和对 $\varphi(m)$因子分解攻击(含因子索数性测试)必耗费巨大，即 $\varphi(m)$的抗攻击性很强。

3.4 一个模数 m 只能被一个人使用的证明

若系统中共有一个模数，只是不同的人拥有不同的 e 和 d，这样实现相对简单，但可能导致共模攻击。当同一信息用不同的公钥加密，这些公钥共模而且互质(一般情况如此)，那末该信息无需私钥就可得到恢复，从而导致泄密。其证明如下：

设 n 为信息明文，两个加密密钥分别为 e_1 和 e_2，公共模数是 m，则：

$$c_1 = m^{e_1} \bmod m$$

$$c_2 = m^{e_2} \bmod m$$

因为 $e1$ 和 $e2$ 互质，故用 Euclideam 算法能找到 r 和 s，满足：

$$r \bullet e_1 + s \bullet e_2 = 1$$

假设 r 为负数，需再用 Euclideam 算法计算 c_1^{-1}，则

$$(c_1^{-1})^{-r} \bullet c_2^{s} = n \bmod m$$

进而可算出 n，也就是说密码分析者知道 m、e_1、e_2、C_1 和 C_2，就能得到 n，即泄密。

另外，还有其它几种利用公共模数攻击的方法。总之，如果知道给定模数的一对 e 和 d，一是有利于攻击者分解模数，二是有利于攻击者计算出其它成对的 e'和 d'，而无须分解模数。解决办法只有一个，那就是不要共享模数 n。

4 e 和 d 选择不当可能导致失密的证明

在选好 p 和 q 后，要选取满足(e, $\varphi(m)$)=1 的 e 值是很容易的事，因为两个随机数互素的概率为 3/5。若采用小的 e，可加快加密的速度，但 e 过小时易遭低加密指数攻击。

(1)当 e 过小时，对小的 n，可能出现 $n^e<m$ 的情况，此时 $c=n^e \bmod m=n^e$，即未取模，由 c 直接开 e 次方就可求出明文 n；

(2)低加密指数攻击。

令网中有 3 用户为加快加密速度，均选用 $e=3$，而有不同的模 m_1、m_2、m_3，一般 m_1、m_2、m_3 互素，否则可求出其公因子，即求出构成 m_1、m_2、m_3 的两个因子 p 和 q 中的 1 个，进而导致解密密钥被破解。若有 1 用户要将明文 n 传给这 3 个用户，其密文分别为：

$c_1=n^3 \bmod m_1 \quad n< m_1$

$c_2=n^3 \bmod m_2 \quad n< m_2$

$c_3=n^3 \bmod m_3 \quad n< m_3$

设 $c=n^3 \bmod (m_1 m_2 m_3)$

利用中国余定理，可根据 m_1、m_2、m_3、c_1、c_2、c_3 求出 c。由于 $n<m_1$、$n<m_2$、$n<m_3$，可得 $n^3<m_1 \bullet m_2 \bullet m_3$，因此 $c=n^3$，故有 $n=\sqrt[3]{c}$，即失密。

可见对于较小的 e，可通过计算出 c，进而解出 n。

对于传送的信息很短时，需要用随机数字进行填充，否则即使 e 不过小，也可能导致低加密指数攻击。

在 RSA 加密过程中，对一些特殊的明文，总会出现 $n^e \bmod m=n$ 的情况，致使信息暴露，如 $n=0$，1，m-1，这些被称为 RSA 加密下的不动点。一般来说，不动点有 $[1+\gcd(e-1,p-1)]\bullet[1+\gcd(e-1,q-1)]$个，由于 p、q、e 均为奇数，$\gcd(e-1,p-1)\geqslant 2$，$\gcd(e-1,q-1)\geqslant 2$，因此不动点至少有 3×3=9 个，具体个数与 p、q、e 的选择有关。

这 9 个明文中，对于 0 和 1，不管采用什么样的参数 e，加密后仍为 0 和 1，但明文块为 0 和 1 时，由于不包含关键信息，不存在泄密问题，也不会造成加密强度的下降；对于不动点 $n=m-1$，其位置只与 p 和 q 相关，如重要明文块刚好为 m-1，则应另选其它的 p 或 q 值；其它不动点的位置则与 p、q、e 的选择同时相关。

当 $e=\log_n(km+n)$ 时：

$$n^e = n^{\log_n(km+n)} = km+n$$

根据 RSA 算法，对明文 n 加密后的密文为：

$c=n^e \bmod m=(km+n) \bmod m=n$

可见，当 $e=\log_n(km+n)$，加密后的密文与明文一模一样，其后果等于直接用明文进行传送，因此在 p、q 已选定的情况下，选择 e 时，要防止其等于 $\log_n(km+n)$。

例：选择 $p=17$，$q=11$，$e=3$ 时，m=187

当 $n=33$ 时，$c=33^3 \bmod 187=33$

当 n=34,67,120,153,154 时，所得密文也仍为 34,67,120,153,154，即 $c=n$。

为了兼顾快速加密和抗低加密指数攻击，e 可考虑采用 16 位的素数。

由于在 RSA 体制中，d 是保密的，而 e 是公开的，因此 d 应有足够的长度以抵抗对 d 的攻击。如果 $d<\frac{1}{3}m^{1/4}$，则存在有效的算法能够解出这个指数。考虑到 d 与 e 的可互换性，应使用选择 d，再算出 e 的办法。可靠的方法是选一个大于 p、q 而小于 $\varphi(m)$，并满足 $\gcd(d,\varphi(m))=1$ 的强素数作为 d，然后算出 e，如果 e 不是很小，也不是很大(否则低速度的终端在加密时将会很慢)，同时 $e\neq\log_n(km+n)$，则 e 就为所求，否则要另选 d 和重新计算 e。 (下转第 124 页)

入 1、2、3、4 个水印后的水印图像所对应的峰值信噪比分别为 42.758 8、39.763 9、37.983 9、36.982 6。

图 4 是从图 3 中提取的水印；图 5 和图 6 分别是从图 3 受到 10%、50%剪切(剪切区域为正方形，剪切比例＝剪切区域的宽度/待测图像的宽度)、攻击后的图像中提取的水印；图 7 和图 8 分别是从图 3 受到最低 3 位、最低 4 位替换攻击后的图像中提取的水印。

图 4 未遭攻击提取的水印

图 5 遭 10%剪切攻击提取的水印

图 6 遭 50%剪切攻击提取的水印

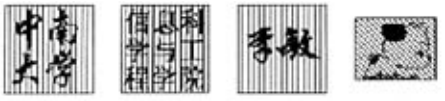

图 7 遭最低 3bits 替换攻击提取的水印

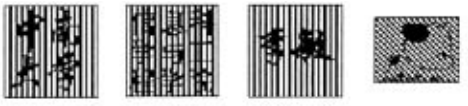

图 8 遭最低 4bits 替换攻击提取的水印

对于本算法，也许会有这样的疑问，如果可以嵌入多个水印，那么算法公开，盗版者或攻击者就可以通过再次嵌入水印而宣称对图像的所有权。事实上，这一点是不可能实现的。这是因为，原始图像的创作者是真正的所有者，攻击者没有原始图像，攻击者所谓的原始图像是包含第 1 次嵌入水印，而不包含第 2 次嵌入水印的图像，这样自然就建立了嵌入的次序。版权所有者可以从攻击者的“原始图像”中提取自己的版权水印，而攻击者却不能从作者的原始图像中提取水印，这样就证明了版权的唯一性。

3 结论

本文提出了一种基于置乱变换的多重数字水印盲算法，通过实验可以看出，嵌入一定数量的水印后的图像，与原始图像的视觉效果基本一致，水印的隐藏效果较好。此外，本算法还具有以下优点：

(1)可嵌入多重水印，标志不同所有者(如创作者、发行者、使用者等)版权使用的合法信息；

(2)水印的嵌入和提取过程中需要个人密钥，不知道密钥的非法使用者无法正确提取水印；

(3)水印的嵌入位置只依赖于密钥，与算法无关，非法用户无法获得准确的水印嵌入位置；

(4)水印的提取无需原始图像和原始水印，是真正意义上的盲提取算法；

(5)水印信息仅隐藏到每个像素的第 3 到第 6 位，可以抵抗 LSB 攻击；

(6)算法是基于置乱变换的，可将水印的嵌入位置均匀地分布到整副图像中，即有效地分散了错误比特，能很好地抵抗剪切攻击。

一个数字水印系统要走向商业应用，其算法必须公开，即算法的安全性完全取决于密钥，而不是对算法进行保密以取得安全性。此外，应能抵抗多重水印嵌入攻击，支持多重水印来标志不同的版权所有者和使用者，并能确定水印嵌入的先后顺序，确保版权的权威性和唯一性。随着社会网络技术和多媒体技术的进一步发展，数字水印产品必将发挥越来越重要的作用。本文给出了一个基于置乱变换的多重水印嵌入的框架，本方法同样可以应用于频率域的整数 DCT 变换和 DWT 变换，以弥补空间域算法的不足。

参考文献

1 Schyndel R G V, Tirkel A Z, Osborne C F. A Digital Watermarks[C]. Proc. of IEEE Int. Conf. on Image Processing, 1994: 86-90.

2 Voyatzis G. The Using of Watermarks in the Protection of Digital Multimedia Products[J]. Proceedings of the IEEE, 1999, 87(7): 1197-1207.

3 Queluz M P, Lamy P, Martinho J M, et al. Spatial Watermark for Image Verification[C]. Proceedings of SPIE International Conf. on Security and Watermarking of Multimedia Contents, SanJose, USA, 2000: 120-130.

4 Santy P, Malay K. A Blind CDMA Image Watermarking Scheme in Wavelet Domain[C]. Proc. of 2004's International Conference on Image Processing, 2004: 2633-2636.

5 Paul B, Ma Xiaoxu. Image Adaptive Watermarking Using Wavelet Domain Singular Value Decompsition[C]. Proc. of IEEE Transactions on Circuits and System for Video Technology, 2005: 96-102.

6 肖　亮, 吴慧中, 韦志辉. 用多数字基整数实现小波域多重数字水印嵌入[J]. 计算机辅助设计与图形学学报, 2003, 24(2).

7 李　敏, 费耀平. 基于队列变换的数字图象置乱算法[J]. 计算机工程, 2005, 31(1): 148-152.

(上接第 119 页)

参考文献

1 王育民, 刘建伟. 通信网的安全——理论与技术[M]. 西安: 西安电子科技大学出版社, 2000.

2 李文卿. 数论及其应用[M]. 北京: 北京大学出版社, 2001.

3 蔡立军. 计算机网络安全技术[M]. 北京: 中国水利水电出版社, 2002.

4 陈鲁生, 沈世镒. 现代密码学[M]. 北京: 科学出版社, 2002.

5 RSA 加密算法存在的问题[Z]. http://jwc.cuit.edu.cn/JXGL/help/Cert/3RsaQues.htm.

6 于秀源. 关于 RSA 加密方法不动点的注记[J]. 计算机学报, 2001, 24(9): 998-1001.

7 颜松远. 数论密码[J]. 科学, 2003, 55(5): 50-54.

8 赵泽茂. 基于 RSA 的随机加密算法与安全可靠度分析[J]. 河海大学学报, 2002, 30(4): 75-77.