REMARKS ON A PROPOSED CRYPTANALYTIC ATTACK ON THE M.I.T. PUBLIC-KEY CRYPTOSYSTEM

Ronald L. Rivest

In this note I would like to demonstrate that the "M.I.T. Public-Key Cryptosystem" (developed by Adi Shamir, Len Adleman, and myself) (1) is essentially invulnerable against the sort of attack recently proposed by G.J. Simmons and M.J. Norris (2). (In all fairness, we point out that they made no claims that the proposed attack method had any chance of success. Here we show that it really has none.)

In our scheme, a message M is encrypted by raising it to a power e, modulo n. Here e and n are integers published by the intended recipient of the encrypted message. The recipient can decipher the received ciphertext by raising it to another power d, modulo n. The recipient has constructed n to be the product of two large prime numbers p and q, and has chosen e to be relatively prime to $(p-1)\cdot(q-1)$. The decoding exponent d is the multiplicative inverse of e, modulo $(p-1)\cdot(q-1)$. Only the recipient knows the correct decoding exponent d, since the computational difficulty (for anyone else) of computing d, given n and e, is provably equivalent to the difficulty of factoring n. Since factoring large numbers is apparently very difficult, one can be confident that publishing e and n will not enable an "enemy" to compute the corresponding decoding exponent d. Only the recipient knows the factors of n; therefore only he can compute d.

A more detailed exposition of our method is given in our paper (1). In particular, the proof that shows that computing d and factoring n are eqivalent in complexity is given there in more detail.

Being able to factor n (or equivalently, finding d) would clearly enable an "enemy" to decipher <u>every</u> message encrypted with the given e and n. However, a cryptographic system should be considered insecure if there exists any way of deciphering a large fraction of the enciphered messages (ciphertexts), even if deciphering every ciphertext is not possible. The paper by Simmons and Norris (2) suggests that such a procedure may exist for our system. The point of our note here is to demonstrate that the fraction of ciphertexts that can be successfully broken with their approach is truly insignificant — one would be better off spending one's time trying to factor n.

The proposed method is to decrypt a ciphertext C (where $C \equiv M^e \pmod n$) by successively re-encrypting C until C is again obtained. Then the original message M is the penultimate message in this list. More formally, one sets C_1 to C, and computes $C_{i+1} \equiv C_i^e \pmod n$ until $C_{i+1} = C$. Then $C_i = M$. This method will be practical only if i turns out to be relatively small (e.g. less than a million). Let's call this i the "iteration exponent" of M; then $M^{e^i} \equiv M \pmod n$.

Two questions immediately arise:

(i) Is there a significant probability that there is a small, universal, iteration exponent which works for all messages M?

(ii) Is there a significant probability that a significant fraction of the messages M have small iteration exponents?

Obviously, a positive answer to either question would imply that our system was "insecure" in a very real sense. Fortunately, we will see that both questions have very definite negative answers.

Our paper (1) makes definite suggestions as to how the prime numbers p and q should be chosen. These suggestions are relevant here, and this note should help to make those suggestions less mysterious. They were that:

(a) $p-1$ and $q-1$ should contain very large prime factors (call them p' and q', respectively), and

(b) similarly, $p'-1$ and $q'-1$ should contain very large prime factors (call them p'' and q'').

Thus, we may write

$$p = a'p'+1, \quad q = b'q'+1$$
$$p' = a''p''+1, \quad q' = b''q''+1$$

for some small a', b', a'', b''. Although (a) and (b) almost certainly hold for a randomly chosen pair of primes p and q, it is simple to construct p and q to explicitly satisfy (a) and (b). The existence of p', p'', q', and q'' will be seen to make the proposed cryptanalytic procedure quite futile.

We say that "M belongs to the exponent k, modulo n" if k is the least positive integer such that $M^k \equiv 1 \pmod n$. In order to find the iteration exponent of M we must determine:

(i) What is the exponent k to which M belongs, modulo n?

(ii) What is the exponent ℓ to which e belongs, modulo k? Then ℓ is the iteration exponent of M.

Let us assume for the sake of concreteness, that the primes p, q, p', q'. p'' and q'' are all larger than 10^{90}. Inasmuch as 10^{80} is an estimate of the number of elementary atomic particles in the known universe, any event which has probability 10^{-90} or less may be realistically considered as truly unlikely, or "impossible" in practice.

To begin with, we observe that a random message M, where $0 \leq M < n$, is truly unlikely to be a multiple of p or q. More precisely, the probability that $\gcd(M,n) \neq 1$ is $(p+q-1)/n$, or roughly 10^{-90}. If it were easy to find such messages M, then n would be easily factored, since $\gcd(M,n)$ would be p or q. We therefore assume that $\gcd(M,n) = 1$, i.e. that M belongs to the multiplicative group of residues which are relatively prime to n.

The size of this multiplicative group, modulo n, is just $\phi(n) = (p-1)\cdot(q-1)$. The order of an element M in this group is just the exponent k to which M belongs, modulo n. Elementary group theory tells us that k must divide $\phi(n) = (p-1)(q-1) = a'p'b'q'$. The group is an abelian group and so is the direct product of cyclic prime-power order subgroups. This product includes $Z_{p'}$ (the cyclic group on p' elements) and also $Z_{q'}$. It is then simple to see that the odds are overwhelming that $p'q'$ divides k. More precisely, the probability is only $(p'+q'-1)/p'q'$, or roughly 10^{-90} that $p'q'$ does not divide the exponent k to which M belongs. Therefore we may assume that $p'q'$ divides k, i.e. that $k = ap'q'$ for some a.

Similarly we ask for the exponent ℓ to which e belongs modulo k. If $M^k \equiv 1 \pmod{n}$, then $M^{e^{\ell} \pmod{k}} \pmod{n}$; the least ℓ such that $e^{\ell} \equiv 1 \pmod{k}$ is therefore the iteration exponent of M - by definition it is also the exponent ℓ to which e belongs, modulo k.

We can argue in a manner similar to that above that the odds are overwhelming that a random encoding exponent e will be relatively prime to $p'q'$; the chance of this not happening is $(p'+q'-1)/p'q' \approx 10^{-90}$. Note that e can be explicitly <u>chosen</u> so that $\gcd(e,p'q') = 1$, as well. Since $p'q'$ divides k, the exponent ℓ' to which e belongs, modulo $p'q'$, must divide the exponent ℓ to which e belongs, modulo k. We now show that it is essentially certain that ℓ', and therefore ℓ, must be enormous.

The exponent ℓ' to which e belongs, modulo $p'q'$, is analogous to the exponent k to which M belongs, modulo $n = p \cdot q$. Since e is virtually certain to be relatively prime to $p'q'$, it belongs to the multiplicative group of residues, modulo $p'q'$, which are relatively prime to $p'q'$. This group has order $\phi(p'q') = a"p"b"q"$, and is abelian. We can conclude, by using the same arguments used above, that the odds are overwhelming that $p"q"$ will divide the exponent ℓ' to which e belongs, modulo $p'q'$. Thus the iteration exponent ℓ of M is essentially certain to be divisible by $p"q"$, which implies that $\ell > 10^{180}$. We note that the recipient can choose e so that $p"q"$ divides ℓ' since it is simple for him to compute ℓ'. If an e is chosen for which $p"q"$ does not divide ℓ', he can simply examine other e's at random until a suitable one is found.

Conclusions

We have shown that the probability that a message M can be decrypted by successively re-encrypting the ciphertext C of M a small number of times is vanishingly small. For numbers of the size suggested, this probability is roughly 10^{-90}. Since the probability of guessing a factor of n is also of this magnitude, we conclude that a cryptanalyst should spend his time trying to factor n rather than using the proposed cryptanalytic approach — a single success then allows him to read every message rather than just the single one he was lucky enough to decrypt.

REFERENCES

1. Rivest, R.L., Shamir, A., and Adleman, L., A Method for Obtaining Digital Signatures and Public-Key Cryptosystems, M.I.T. Laboratory for Computer Science Technical Memo #82, April 1977, to appear in CACM, Feb. 1978.

2. Simmons, G.J. and Norris, M.J., Preliminary Comments on the M.I.T. Public-Key Cryptosystem, CRYPTOLOGIA 1(4)(1977), 406-414.