A NEW RELATED-MESSAGE-ATTACK ON RSA

ODED YACOBI AND YACOV YACOBI

Abstract. Coppersmith, Franklin, Patarin, and Reiter have shown that given two RSA cryptograms $x^e \bmod N$, and $(ax+b)^e \bmod N$ for any known constants $a, b \in \mathbb{Z}_N$ one can compute x in $O(e \log^2 e)$ $\mathbb{Z}_N$-operations with some positive error probability. We show that given e cryptograms $c_i \equiv (ax + b \cdot i)^e \bmod N$, $i = 0, 1, ...e-1$, for any known constants $a, b \in \mathbb{Z}_N$, where $\gcd(a, N) = \gcd(b, N) = \gcd(e!, N) = 1$, one can deterministically compute x in $O(e)$ $\mathbb{Z}_N$-operations using

$$x \equiv a^{-1} b[(b^e e!)^{-1} \sum_{i=0}^{e-1} \binom{e-1}{i} \cdot c_i \cdot (-1)^{e-1+i} - \frac{e-1}{2}] \bmod N.$$

Other applications of the new technique are briefly noted at the end of this paper, including evidence that a certain class of polynomial reductions from discrete-log problems to bi-linear Diffie-Hellman problem does not exist.

1. Introduction

In [CFPR] it was shown that given two RSA cryptograms $x^e \bmod N$, and $(ax+b)^e \bmod N$ for any known constants $a, b \in \mathbb{Z}_N$ one can compute x in $O(e \log^2 e)$ $\mathbb{Z}_N$-operations with some positive error probability. We show that given e cryptograms $c_i \equiv (ax+b \cdot i)^e \bmod N$, $i = 0, 1, ...e-1$, for any known constants $a, b \in \mathbb{Z}_N$, where $\gcd(a, N) = \gcd(b, N) = \gcd(e!, N) = 1$, one can deterministically compute x in $O(e)$ $\mathbb{Z}_N$-operations using

$$x \equiv a^{-1} b[(b^e e!)^{-1} \sum_{i=0}^{e-1} \binom{e-1}{i} \cdot c_i \cdot (-1)^{e-1+i} - \frac{e-1}{2}] \bmod N$$

(if any of the above gcd conditions doesn't hold then the system is already broken).

It remains an open problem whether the new approach can improve more general explicit linear dependence or the general case of implicit linear dependence, ie, $\sum_{i=1}^{k} a_i x_i = a_0$, for known scalars a_i, $i = 0, 1, 2, ...k$, whose current complexity is $O(e^{k/2} k^2)$ [CFPR]. Other applications of the new technique are briefly noted at the end of this paper, including evidence that a certain class of polynomial reductions from discrete-log problems to bi-linear Diffie-Hellman problem does not exist.

Related message attacks can be avoided altogether if before RSA-encryption the message, M, is transformed using the OAEP (Bellare-Rogoway) function into $M' = [M \oplus G(r) \;||\; r \oplus H(M \oplus G(r))]$, where r is a truely-random nonce, H is a secure hash function, G is a random number generator function, $||$ is the concatenation sign, and $\oplus$ is the bit-by-bit exclusive-or operation. Once M' is RSA-decrypted the net message M can be efficiently extracted.

Date: July 14, 2004.

Key words and phrases. RSA cryptosystem, Discrete-Log, finite difference, related message attack.

2. Finite Differences

We use upper-case letters to denote indeterminate variables and lower-case letters to denote particular values of those variables. Let $h(X) \equiv \sum_{i=0}^{n} a_i X^i \bmod N$ where $N = pq$ is a "safe" RSA composite (p and q are large primes, with some additional restrictions) $a_i \in \mathbb{Z}_N$, $a_n \neq 0$, $n \leq \varphi(N)$, and define

$$\Delta(X) \equiv h(X+1) - h(X) \bmod N.$$

The next lemma is not needed for our main result, related message attack on RSA. It is needed only when $\deg(h)$ is not known (may happen in some examples in section 4 and in the appendix). We nevertheless include it here for the sake of completeness.

Lemma 1. *If $n < \min\{p,q\}$ then (i) $n = \deg(h) > 0$ implies $\deg(\Delta) = \deg(h) - 1$; (ii) $\deg(h) = 0$ iff $\Delta = \mathbf{0}$ (the zero polynomial).*

Proof. For $n > 0$, $\Delta(X) \equiv \sum_{i=1}^{n} a_i[(X+1)^i - X^i]$
$\equiv \sum_{i=1}^{n} a_i[\sum_{j=0}^{i} \binom{i}{j} X^j - X^i]$
$\equiv \sum_{i=1}^{n} a_i[\sum_{j=0}^{i-1} \binom{i}{j} X^j]$ $(*)$

(i) From $(*)$ $\deg(\Delta) \leq \deg(h) - 1$. Since $n < \min\{p,q\}$, and $a_n \neq 0 \bmod N$ we conclude that $a_n \binom{n}{n-1} \neq 0 \bmod N$, and therefore $\deg(\Delta) = \deg(h) - 1$.

(ii) All congruences are mod N, and therefore we omit the "mod N" notation. If $n = 0$ then $h \equiv a_0$, therefore $\Delta \equiv \mathbf{0}$. Let $\Delta(X) \equiv \sum_{i=0}^{n-1} \delta_i X^i$. If $\Delta \equiv \mathbf{0}$ (the zero polynomial) then from $(*)$, $\Delta(X) \equiv \sum_{i=0}^{n-1} \delta_i X^i \equiv \sum_{i=1}^{n} a_i[\sum_{j=0}^{i-1} \binom{i}{j} X^j]$, $\delta_{n-1} \equiv a_n \binom{n}{n-1}$, so since $n < q$ we conclude that $\delta_{n-1} \equiv 0$ implies $a_n \equiv 0$. $\delta_{n-2} \equiv a_n \binom{n-1}{n-2} + a_{n-1} \binom{n-2}{n-2} \equiv a_{n-1}$. So $\delta_{n-2} \equiv 0$ implies $a_{n-1} \equiv 0$. And in general $a_i \equiv \delta_{i-1}$, for $i = n, ...1$, hence $\deg(h) = 0$. □

Definition 1. *Let $\Delta^{(0)}(X) = h(X)$, and let $\Delta^{(i)}(X) \equiv \Delta^{(i-1)}(X+1) - \Delta^{(i-1)}(X) \bmod N$, $i = 1, 2, ...$*

Lemma 2. *Let $0 \leq k \leq n = \deg(h)$. $\Delta^{(k)}(X) \equiv \sum_{i=0}^{k} \binom{k}{i} \cdot h(X+i) \cdot (-1)^{k-i} \bmod N$.*

Proof. By induction on k. □

Let $0 \leq k \leq n = \deg(h)$, and let $T^{(k)}_{a_n, a_{n-1}}(X)$ denote the two leading terms of $\Delta^{(k)}(X)$.

Lemma 3. $T^{(k)}_{a_n, a_{n-1}}(X) = \frac{(n-1)!}{(n-k)!} X^{n-k-1}(a_n n(X + k(n-k)/2) + a_{n-1}(n-k))$.

Proof. Induction on k. Basis: $T^{(0)}_{a_n, a_{n-1}}(X) = \sum_{i=n-1}^{n} a_i X^i$; We verify one more step, $k = 1$, that is needed later.

$$T^{(1)}_{a_n, a_{n-1}}(X) = X^{n-2}(a_n n(X + \frac{n-1}{2}) + a_{n-1}(n-1)).....(*)$$

$\Delta^{(1)}(X) = h(X+1) - h(X)$, whose two leading terms are equal to $T^{(1)}_{a_n, a_{n-1}}(X)$ above.

Induction hypothesis: The two leading terms of $\Delta^{(k-1)}(X)$ are
$T^{(k-1)}_{a_n, a_{n-1}}(X) = \frac{(n-1)!}{(n-k+1)!} X^{n-k}(a_n n(X + (k-1)(n-k+1)/2) + a_{n-1}(n-k+1))$.

Let $T^{(k-1)}_{a_n, a_{n-1}}(X) = \alpha X^{n-k+1} + \beta X^{n-k}$, namely, $\alpha = \frac{(n-1)!}{(n-k)!} a_n n$, and $\beta = \frac{(n-1)!}{(n-k)!}[a_n nk(n-k)/2 + a_{n-1}(n-k)]$. The proof can be completed by showing that

$T^{(1)}_{\alpha,\beta}(X) = T^{(k)}_{a_n,a_{n-1}}(X)$, namely, we compute the first difference of $T^{(k-1)}_{a_n,a_{n-1}}(X)$ substituting α for a_n and β for a_{n-1} in (*) to get the claim. □

3. Related-messages attack

Here we analyze the case where $h(X)$ is the RSA encryption function.

Corollary 1. *The case $h(X) = X^e \bmod N$, where $e \geq 3$, is a special case of the previous lemma, with $a_n = 1, a_{n-1} = 0$, and $T^{(e-1)}_{1,0} \equiv e!(X + (e-1)/2) (\bmod N)$.*

Theorem 1. *Let (e, N) be any public RSA key with $gcd(e!, N) = 1$, $e \geq 3$, and let $m_i \equiv x + i \bmod N$, $i = 0, 1, 2, ..., e-1$. Let $c_i \equiv m_i^e \bmod N$ be known cryptograms. Then $x \equiv e!^{-1} \sum_{i=0}^{e-1} \binom{e-1}{i} \cdot c_i \cdot (-1)^{e-1+i} - (e-1)/2 \mod N$, which is computable in $O(e)$ operations in $\mathbb{Z}_N$.*

Proof. $\Delta^{(e-1)}(X) \equiv T^{(e-1)}_{1,0}(X)$. For any particular value x of X, the lhs can be computed using the previous theorem ie,

$$\Delta^{(e-1)}(x) \equiv \sum_{i=0}^{e-1} \binom{e-1}{i} \cdot c_i \cdot (-1)^{e-1+i} \bmod N,$$

and the rhs is given by the above corollary. $e!^{-1}$ exist by our assumptions. Since $\binom{e-1}{i}$ can be computed from $\binom{e-1}{i-1}$ using one multiplication and one division, this computation takes $O(e)$ operations in $\mathbb{Z}_N$. □

Corollary 2. *If $m_i \equiv ax + bi \bmod N$, $i = 0, 1, 2...e-1$, for known a and b, with $\gcd(a, N) = \gcd(b, N) = 1$, we can likewise compute x. Given cryptogram $c_i \equiv (ax + b \cdot i\ )^e \bmod N$ we can transform it into $c'_i \equiv c_i \cdot b^{-e} \equiv (y + i)^e \bmod N$, where $y \equiv xab^{-1} \bmod N$. So*

$$x \equiv a^{-1} b [b^e e!)^{-1} \sum_{i=0}^{e-1} \binom{e-1}{i} \cdot c_i \cdot (-1)^{e-1+i} - \frac{e-1}{2}] \mod N.$$

4. Other cryptographic applications of finite differences

The results of section 2 can be easily generalized to polynomials $h \in \mathbb{Z}[X]$, and $h \in \mathbb{Z}_q[X]$, where q is a prime.

(1) DL mod a prime (given g^x in any multiplicative group, find x). Suppose that given g^x an oracle returns the value of some polynomial $h(x)$ instead of x. The straight forward way of solving for x would be to interrogate the oracle $n + 1$ times, where $n = \deg(h)$, extrapolate h, then factor h to find its root x. This has complexity almost quadratic in n. Since we can interrogate the oracle with any g^{x+i}, we can do it in linear time in n, using the finite difference method modulo a prime.

(2) DL modulo a "safe" composite: Here we cannot use the alternative method proposed in (1) above, since efficient polynomial factorization is not known to be possible modulo a "safe" composite modulus, while the new method using finite differences still works with the same efficiency as modulo a prime. In particular (as a special case) suppose some DL oracle returns its results RSA-encrypted. We can create linearly related messages (eg, if $c_0 \equiv g^x \bmod p$ we can create $c_i \equiv c_1 \cdot g \equiv g^{x+i} \bmod p$. The indexes $x + i$ are the messages to be RSA encrypted), and use the finite-difference method to find x.

(3) We can use the finite difference technique to create evidence that one-oracle-call polynomial reductions from old DL type problems to the new Binary Diffie-Hellman problem do not exist (see appendix).

Acknowledgements:

Special thanks go to Peter Montgomery who made numerous valuable suggestions and corrections.

We also thank Don Coppersmith, Kamal Jain, Adi Shamir, and Venkie (Ramarathnam Venkatesan) for very valuable discussions on earlier applications of the finite difference technique.

5. References

[BF] Dan Boneh and M. Franklin: Identity Based Encryption from the Weil-Pairing, Proc. Crypto'01, Springer-Verlag LNCS 2139, pp. 213-229.

[CFPR]: Don Coppersmith, Matthew Franklin, Jacques Patarin, Michael Reiter: Low-Exponent RSA with related Messages, Proc. of Eurocrypt'96, LNCS 1070, pp. 1-9.

[CS] Don Coppersmith, Igor Shparlinski: On Polynomial Approximation of the Discrete Logarithm and the Diffie-Hellman Mapping, J. of Crypt. (2000) 13:339-360.

[ESY] Shimon Even, Alan Selman, Yacov Yacobi: The Complexity of Promise Problems with Applications to Public-Key Cryptography. Information and Control 61(2): 159-173 (1984),

[GS] Joachim von zur Gathen, Igor Shparlinski: Polynomial interpolation from multiples. SODA 2004:1132-1137.

[J] A. Joux: The Weil and Tate Pairings as Building blocks for Public Key Cryptosystems (survey). In ANTS 2002, LNCS 2369, pp. 20-32, 2002, Springer-Verlag.

[K] Neal Koblitz: A Course in Number Theory and Cryptography. Springer-Verlag Graduate Texts in Mathematics, 1987 ISBN 0-387-96576-9.

[R] Joseph J. Rotman: An Introduction to the theory of groups,WCB publishers,1988, ISBN 0-697-06882-X.

[RSA] Ronald Rivest, Adi Shamir, Leonard M. Adleman: A Method for Obtaining Digital Signatures and Public-Key Cryptosystems. CACM 21(2): 120-126 (1978).

[V] E. Verheul: Evidence that XTR is more secure than supersingular elliptic curve cryptosystems. In B. Pfitzmann, editor, *Proc. of Eurocrypt'01, Vol. 2045 Lecture Note in Computer Science, pp. 195-210,* Springer, 2001.

6. Appendix: Impossible reductions

We give evidence that certain 1-oracle-call polynomial reduction from Decision Diffie-Hellman in multiplicative group $\mathbb{G}_2$ (DDH_{G_2}) to Binary Diffie-Hellman in additive group $\mathbb{G}_1$ (BDH) do not exist. We restrict the discussion to single oracle call reductions, since we don't yet have a more general proof.

6.1. **Definition of the problems BDH and DDH_{G_2}.** Let $\mathbb{G}_1$ and $\mathbb{G}_2$ be cyclic groups of order q, with generators P and g, respectively. We describe elements of these groups in terms of these generators, where $\mathbb{G}_2$ is assumed multiplicative and $\mathbb{G}_1$ is assumed additive. Let $\hat{e} : \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$ be a non-degenerate bi-linear

mapping. Examples of such mappings are the modified Weil pairing [BF] and the modified Tate pairing [J]. The map must satisfy the following properties [BF]:

(1) Bilinear: We say that a map $\hat{e} : \mathbb{G}_1 \times \mathbb{G}_1 \rightarrow \mathbb{G}_2$ is bilinear if $\hat{e}(P_1 + P_2, Q) = \hat{e}(P_1, Q) \cdot \hat{e}(P_2, Q)$ and $\hat{e}(P, Q_1 + Q_2) = \hat{e}(P, Q_1) \cdot \hat{e}(P, Q_2)$. This implies $\hat{e}(aP, bQ) = \hat{e}(P, Q)^{ab}$ for all $P, Q \in \mathbb{G}_1$, and all $a, b \in \mathbb{Z}$.
(2) Non-degenerate: The map does not send all pairs in $\mathbb{G}_1 \times \mathbb{G}_1$ into the identity element in $\mathbb{G}_1$. It follows that if P is a generator of $\mathbb{G}_1$ then $\hat{e}(P, P)$ is a generator of $\mathbb{G}_2$.
(3) Computable: There is an efficient algorithm to compute $\hat{e}(R, S)$ for any $R, S \in \mathbb{G}_1$.

The problem for which we seek assurance that it is of high complexity is the BDH problem:

Given $< P, aP, bP, cP > \in \mathbb{G}_1^4$; **Find** $\hat{e}(P, P)^{abc} \in \mathbb{G}_2$.

The main old problem under consideration is the *Decision Diffie-Hellman* in the group $\mathbb{G}_2$ (denoted DDH_{G_2}), defined as follows:

Given: $< g, g^x, g^y, g^z > \in \mathbb{G}_2^4$; **Decide:** $g^{xy} = g^z$?

6.2. **The function ζ: $\mathbb{G}_2 \rightarrow \mathbb{G}_1$.** Let α be a problem with input and output domains In(α),and Out(α), respectively, and let Turing Machine (TM) M_α solve problem α. This can be viewed as a function M_α :In(α) $\rightarrow$ Out(α). We combine machines with matching domains in the natural way. If problem α is reducible to problem β in polynomial time *using one oracle call* then there exist polynomial time TM M_1 and M_2 such that $M_\alpha = M_2 M_\beta M_1$ (where M_β is the oracle that solves problem β), and where $M_1 : In(\alpha) \rightarrow In(\beta)$, and $M_2 : Out(\beta) \rightarrow Out(\alpha)$. We are interested in the following problems (in all the cases we refer to the set of elements of groups and to functions from one set to another rather than to group morphisms, but we use short hand of the form $\zeta : \mathbb{G}_2 \rightarrow \mathbb{G}_1$): (i) BDH: $\mathbb{G}_1^4 \rightarrow \mathbb{G}_2$; (ii) $\text{DDH}_{G_2} : \mathbb{G}_2^4 \rightarrow \{yes, no\}$.

6.2.1. $< \zeta, h >$ *pairs.* We assume that $\mathbb{G}_1$ is an additive group defined by some elliptic curve E over some field $\mathbb{F}$, and that it has a generator P. Likewise, $\mathbb{G}_2$ is a multiplicative group with generator g.

Let $q = |\mathbb{G}_2| = |\mathbb{G}_1|$ and $\zeta : \mathbb{G}_2 \rightarrow \mathbb{G}_1$. There exists a function h defined over $\mathbb{F}$ such that for all $0 \leq x < q$, $\zeta(g^x) = h(x)P$. Any function h on $\mathbb{F}$ is a polynomial with $\deg(h) \leq |\mathbb{F}|$.

From Hasse's Theorem ([K], pp. 158) the difference between N, the number of points in E, and $|\mathbb{F}| + 1$ is upper bounded by $2|\mathbb{F}|^{1/2}$.

So the upper bound on $\deg(h)$ cannot be too far from N. If $\mathbb{G}_1$ is chosen so that $|\mathbb{G}_1| = q$ is a large prime divisor of N then the upper bound on $\deg(h)$ cannot be too far from q either.

6.3. **Related work.** Our work is related to a work by Coppersmith and Shparlinski [CS] on the degree of a polynomial that evaluate the DL of x, or in fact to an even earlier work, by Mullen and White [MW], showing that the DL mod a prime p, denoted $ind(x)$ in those papers, is $ind(x) \equiv (-1 + \sum_{k=1}^{p-2}(g^{-k} - 1)^{-1} x^k) \bmod p$, where g is a primituve root of a finite field of q elements, F_q. But our polynomial is $h(ind(x))$, not $h(x)$, as in [CS], and in addition we cover any modulus, including a composite.

The XTR is a computationally efficient subgroup of order $p^2 - p + 1$ of the multiplicative group $GF(p^6)^*$ of the finite field $GF(p^6)$. In [V], Verheul shows that finding an efficient injective homomorphism from the XTR subgroup into the group of points of a particular supersingular elliptic curve group over $GF(p^2)$ is at least as hard as solving the Diffie-Hellman problem in the XTR subgroup.

Joux [J] gives a very detailed review on known reductions between the various DL-type problems and BDH.

6.4. **Claims.**

Theorem 2. *Let $\mathbb{G}_1, \mathbb{G}_2$ be groups of prime order q, with $\mathbb{G}_1$ written additively and $\mathbb{G}_2$ written multiplicatively, and let $P \in \mathbb{G}_1 \backslash \{1\}$ and $g \in \mathbb{G}_2 \backslash \{1\}$ be their generators, respectively. Suppose we have oracle $\zeta : \mathbb{G}_2 \to \mathbb{G}_1$, whose associated function on $\mathbb{G}_1$ is h, ie for all $0 \le x < q$, $\zeta(g^x) = h(x)P$. If $\deg(h) = O(poly\log(q))$ then ζ can be used to solve DDH_{G_2} in polynomial time.*

Proof. We use the lemmas of section 2 with a twist, since here we don't have direct access to the polynomials $h(x)$. Rather we have access to the values $h(x)P$. But as shown below this is sufficient for the decision problems at hand. Since $\zeta(g^x) = h(x)P = \Delta^{(0)}(x)P$, we can compute $\Delta^{(1)}(x)P = \Delta^{(0)}(x+1)P - \Delta^{(0)}(x)P = \zeta(g^x \cdot g) - \zeta(g^x)$, and recursively compute any $\Delta^{(k)}(x)P$, $k \le \deg(h)$.

Even more efficiently, analogously to lemma 2 we can compute it using a variant of the Pascal triangle as follows (see comment at the beginning of section 4):

$$\Delta^{(k)}(x) \equiv \sum_{i=0}^{k} \binom{k}{i} \cdot h(x+i)P \cdot (-1)^{k-i} \equiv \sum_{i=0}^{k} \binom{k}{i} \cdot \zeta(g^{x+i}) \cdot (-1)^{k-i} \bmod q.$$

This can be done even if the polynomial h is unknown, however, in that case we cannot make use of lemma 3 (and it is unimportant, since here all we need is a complexity which is polynomial in $n = \deg(h)$). We proceed assuming that h and its degree, n, are unknown. For $i = 1, 2, ...$ compute $\Delta^{(i)}(x)P$ until $\Delta^{(n+1)}(x)P = \mathcal{O}$, the identity element of $\mathbb{G}_1$ (i.e. the point at infinity if $\mathbb{G}_1$ is a group defined by elliptic curve; here we use lemma 1(ii)). The value of the polynomial before this step is $\Delta^{(n)}(x) = ux + v$ for some constants u and v, and we know $\Delta^{(n)}(x)P$. Since $vP = \Delta^{(n)}(0)P$, we can compute $uxP = \Delta^{(n)}(x)P - \Delta^{(n)}(0)P$. Likewise we can compute uyP and uzP. We can also compute $uP = \Delta^{(n)}(1)P - \Delta^{(n)}(0)P$. Note that $\hat{e}(uxP, uyP) \equiv \hat{e}(uzP, uP)$ iff $g^{u^2xy} \equiv g^{u^2z}$. Since q is prime, $\gcd(u, q) = 1$ and this congruence is true iff $g^{xy} \equiv g^z$. The complexity of this process is $O(n^2) = O(poly\log(q))$. □

Corollary 3. *Since we believe that DDH_{G_2} is a hard problem, either deg(h)>O(polylog q), or a 1-oracle call reduction (from DDH_{G_2} to BDH) does not exist.*

We can prove a similar result with the Computational Diffie-Hellman in $\mathbb{G}_1$ replacing DDH_{G_2}, and from these two results via the transitivity of 1-oracle-call

polynomial reduction reach similar conclusions with respect to all DL-type problems which are higher in the hierarchy (see hierarchy in [J]).

Oded Yacobi Cornell University
E-mail address: yacobi@cornell.edu

Yacov Yacobi Microsoft Research
E-mail address: yacov@microsoft.com