The Rabin scheme revisited

Michele Elia (Politecnico di Torino)
joint work with Matteo Piva (University of Trento)
Davide Schipani (University of Zurich)

Bunny 2
Trento, 12 settembre 2011

Outline

Rabin scheme and roots of polynomials

- In 1979, Michael Rabin suggested a variant of RSA with public-key exponent 2, which he showed to be as secure as factoring.

Rabin scheme and roots of polynomials

- In 1979, Michael Rabin suggested a variant of RSA with public-key exponent 2, which he showed to be as secure as factoring.

  Let $N = pq$ be a product of two prime numbers.

Rabin scheme and roots of polynomials

- In 1979, Michael Rabin suggested a variant of RSA with public-key exponent 2, which he showed to be as secure as factoring.

  Let $N = pq$ be a product of two prime numbers.
- **Encryption** of a message $m \in \mathbb{Z}_N^*$ is

$$C = m^2 \bmod N$$

Rabin scheme and roots of polynomials

- In 1979, Michael Rabin suggested a variant of RSA with public-key exponent 2, which he showed to be as secure as factoring.

  Let $N = pq$ be a product of two prime numbers.
- **Encryption** of a message $m \in \mathbb{Z}_N^*$ is

$$C = m^2 \bmod N$$

- **Decryption** is performed by solving the equation

$$x^2 = C \bmod N \quad , \tag{1}$$

which has four roots in $\mathbb{Z}_N^*$.

Rabin scheme and roots of polynomials

Rabin scheme and roots of polynomials

Key points

Rabin scheme and roots of polynomials

Key points

- To solve equation (1) is "easy" if the factors of N are known

Rabin scheme and roots of polynomials

Key points

- To solve equation (1) is "easy" if the factors of N are known
- To solve equation (1) is "hard" if the factors of N are not known

Rabin scheme and roots of polynomials

Key points

- To solve equation (1) is "easy" if the factors of N are known
- To solve equation (1) is "hard" if the factors of N are not known
- **To solve the equation $x^2 - C = 0 \bmod N$ is equivalent to factor N**

Rabin scheme and roots of polynomials

Key points

- To solve equation (1) is "easy" if the factors of N are known
- To solve equation (1) is "hard" if the factors of N are not known
- **To solve the equation $x^2 - C = 0 \bmod N$ is equivalent to factor N**

Key issue (at the decryption stage)

Rabin scheme and roots of polynomials

Key points

- To solve equation (1) is "easy" if the factors of N are known
- To solve equation (1) is "hard" if the factors of N are not known
- **To solve the equation $x^2 - C = 0 \bmod N$ is equivalent to factor N**

Key issue (at the decryption stage)

- Once the four roots x_1, x_2, x_3, x_4 of equation (1) are known, how do we identify the original message?

Rabin scheme and roots of polynomials

Key points

- To solve equation (1) is "easy" if the factors of N are known
- To solve equation (1) is "hard" if the factors of N are not known
- **To solve the equation $x^2 - C = 0 \bmod N$ is equivalent to factor N**

Key issue (at the decryption stage)

- Once the four roots x_1, x_2, x_3, x_4 of equation (1) are known, how do we identify the original message?
- The further information should be computed from m without knowing the factors of N (or any information leading to easy factorization)

Preliminaries: Chinese Remainder Theorem (CRT)

- Every element a of $\mathbb{Z}_N$ is uniquely identified by its remainders a_p and a_q with respect to p and q.
- a is reconstructed by the CRT as

$$a = a_p\psi_1 + a_q\psi_2 \bmod N$$

- ψ_1 and ψ_2, obtained from the extended Euclidean algorithm, are defined by

$$\begin{cases} \psi_1 = 1 \bmod p, & \psi_1 = 0 \bmod q \\ \psi_2 = 0 \bmod p, & \psi_2 = 1 \bmod q, \end{cases}$$

and satisfy

$$\begin{cases} \psi_1\psi_2 = 0 \bmod N \\ \psi_1^2 = \psi_1 \bmod N \\ \psi_2^2 = \psi_2 \bmod N \end{cases} .$$

Preliminaries: Roots in $\mathbb{Z}_N$

- The equation $X^2 - C = 0$ is solvable $\bmod N$ if and only if it is solvable $\bmod p$ and $\bmod q$.
- Let u_1 be a root $\bmod p$, the second root is $-u_1$
- Let v_1 be a root $\bmod q$, the second root is $-v_1$
- The four roots can be written as

$$\begin{cases} x_1 = u_1\psi_1 + v_1\psi_2 & \bmod N \\ x_2 = u_1\psi_1 + (q - v_1)\psi_2 & \bmod N \\ x_3 = (p - u_1)\psi_1 + v_1\psi_2 & \bmod N \\ x_4 = (p - u_1)\psi_1 + (q - v_1)\psi_2 & \bmod N \end{cases} . \tag{2}$$

- $x \to x^2$ is a 4 to 1 mapping

Preliminaries

Lemma (A)

- *The four roots* x_1, x_2, x_3, x_4 *of the polynomial* $x^2 - C$ *are partitioned into two sets* $\mathfrak{R}_1 = \{x_1, x_4\}$ *and* $\mathfrak{R}_2 = \{x_2, x_3\}$ *such that the roots in the same set have different parity, i.e.* $x_1 = 1 + x_4 \bmod 2$ *and* $x_2 = 1 + x_3 \bmod 2$*.*
- *Assuming that* u_1 *and* v_1 *in equation (2) have the same parity, the residues modulo* p *and modulo* q *of each root in* $\mathfrak{R}_1$ *have the same parity, while the roots in* $\mathfrak{R}_2$ *have residues of different parity.*

Preliminaries: the mapping $x \to x^2$

By Lemma (A) each x_i is identified by the pair of bits

$$B_p = (x_i \bmod p) \bmod 2, \quad \text{and} \quad B_q = (x_i \bmod q) \bmod 2 \quad .$$

In summary we have the table

root	B_p	B_q
x_1	$u_1 \bmod 2$	$v_1 \bmod 2$
x_2	$u_1 \bmod 2$	$q - v_1 \bmod 2$
x_3	$p - u_1 \bmod 2$	$v_1 \bmod 2$
x_4	$p - u_1 \bmod 2$	$p - v_1 \bmod 2$

Preliminaries: the mapping $x \to x^2$

For example if $u_1 = v_1 = 0 \bmod 2$ and suppose x_1 and x_2 even, we have

root	B_p	B_q	$B_p + B_q \bmod 2$	$x_i \bmod 2$
x_1	0	0	0	0
x_2	0	1	1	0
x_3	1	0	1	1
x_4	1	1	0	1

A root x_i is identified by the pair of bits

$$
\begin{aligned}
b_0 &= x_i \bmod 2 \\
b_1 &= [x_i \bmod p] + [x_i \bmod q] \bmod 2
\end{aligned}
$$

Preliminaries: the mapping $x \to x^2$

Roots of unity

- $x^2 = 1 \bmod N$ has roots

$$1 \quad , \quad -1 \quad , \quad \psi_1 - \psi_2 \quad , \quad -\psi_1 + \psi_2$$

- If a root m of $x^2 - C = 0 \bmod N$ is known the four roots are
  m, $-m$, $m(\psi_1 - \psi_2) \bmod N$, and $m(-\psi_1 + \psi_2) \bmod N$
- If we know the factors of N, we may compute the roots of unity
- If we are able to compute the roots of unity, then we may factor N

Preliminaries: Legendre and Jacobi symbols

1) Legendre symbol is defined for every odd prime p as

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } \; x^2 = a \bmod p \; \text{ is solvable in } \; \mathbb{Z}_p \\ -1 & \text{if } \; x^2 = a \bmod p \; \text{ is not solvable in } \; \mathbb{Z}_p \end{cases}$$

2) Jacobi symbol is defined for every pair r, s of positive odd integers as

$$\left(\frac{a}{rs}\right) = \left(\frac{a}{r}\right)\left(\frac{a}{s}\right)$$

3) $\left(\frac{a\psi_1 + b\psi_2}{pq}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{q}\right)$

4) If p and q are congruent to 3 modulo 4 the roots x_1 and x_2 of equation (1) have opposite Jacobi symbol

$$\left(\frac{x_1}{p}\right) = -\left(\frac{x_2}{q}\right)$$

Preliminaries: Legendre and Jacobi symbols

5) $\left(\frac{a+\mu z}{z}\right) = \left(\frac{a}{z}\right)$

6) Reciprocity law

$$\left(\frac{a}{b}\right) = \left(\frac{b}{a}\right)(-1)^{\frac{(a-1)(b-1)}{4}}$$

$$\left(\frac{2}{b}\right) = (-1)^{\frac{b^2-1}{8}}$$

1. *The properties 5) and 6) allow us to compute Legendre and Jacobi symbols by a method that mimics the Euclidean algorithm and has the same efficiency.*

Preliminaries: Dedekind sums

Definition

Let h, k be relatively prime and $k \geq 1$, a Dedekind sum is denoted by $s(h,k)$ and defined as

$$s(h,k) = \sum_{j=1}^{k} \left(\left(\frac{hj}{k}\right)\right)\left(\left(\frac{j}{k}\right)\right) \tag{3}$$

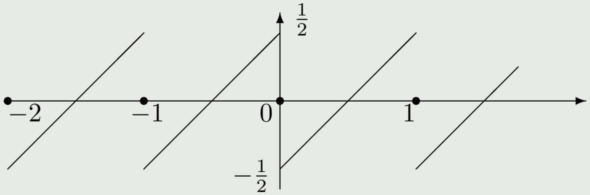
where the symbol $((x))$, defined as

$$((x)) = \begin{cases} x - \lfloor x \rfloor - \frac{1}{2} & \text{if } x \text{ is not an integer} \\ 0 & \text{if } x \text{ is an integer} \end{cases}, \tag{4}$$

denotes the well-known sawtooth function of period 1.

Preliminaries: Dedekind sums

Sawtooth function



Preliminaries: Dedekind sums

Properties

1) $h_1 = h_2 \bmod k \Rightarrow s(h_1, k) = s(h_2, k)$
2) $s(-h, k) = -s(h, k)$
3) $s(h, k) + s(k, h) = -\frac{1}{4} + \frac{1}{12}\left(\frac{h}{k} + \frac{1}{hk} + \frac{k}{h}\right)$, a property known as the reciprocity law for the Dedekind sums.
4) $12ks(h, k) = k + 1 - 2\left(\frac{h}{k}\right) \bmod 8$ for k odd, a property connecting Dedekind sums and Jacobi symbols.

The properties 1), 2), and 3) allow us to compute a Dedekind sum by a method that mimics the Euclidean algorithm and has the same efficiency.

Preliminaries: Dedekind sums

Lemma (B)

If $k = 1 \bmod 4$, then, for any h relatively prime with k, the denominator of $s(h,k)$ is odd.

Proof outline: using properties of the Dedekind sums we have

$$s(h,k) = \sum_{j=1}^{k-1} \frac{j}{k}\left(\frac{hj}{k} - \left\lfloor \frac{hj}{k} \right\rfloor - \frac{1}{2}\right) ,$$

the summation can be split into two summations such that

- the first summation has the denominator patently odd;
- the second summation, evaluated as $-\frac{1}{2}\sum_{j=1}^{k-1}\frac{j}{k} = -\frac{k-1}{4}$ is an integer by hypothesis

Preliminaries: Dedekind sums

Lemma (C)

If k is a product of two Blum primes, x_1 is relatively prime with k, and $x_2 = x_1(\psi_1 - \psi_2)$, then $s(x_1, k) + s(x_2, k) = 1 \bmod 2$.

Proof outline: by property 4) of the Dedekind sums we have

$$12Ns(z_1, N) = N + 1 - 2\left(\frac{z_1}{N}\right) \bmod 8 \qquad i = 1, 2$$

thus, summing member by member the expressions for $i = 1$ and 2, and taking into account that $N = 1 \bmod 4$ we have

$$12N[s(z_1, N) + s(z_2, N))] = 2N + 2 - 2\left[\left(\frac{z_1}{N}\right) + \left(\frac{z_2}{N}\right)\right] \bmod 8\,,$$

since $12N = 4 \bmod 8$, $2N = 2 \bmod 8$; and the sum of the two Jacobi symbols is 0. The conclusion follows from the application of Lemma (B).

Williams' scheme

- In 1980, Williams proposed an implementation of Rabin scheme using a parity bit and the Jacobi symbol for identifying the message.
- The decryption process is based on the observation that, setting $D = \frac{1}{2}(\frac{(p-1)(q-1)}{4} + 1)$, if $b = a^2 \bmod N$ and $\left(\dfrac{a}{N}\right) = 1$, we have $a = \pm b^D$.

Williams' scheme

Public-key

- $[N, S]$, where S is an integer such that $\left(\frac{S}{N}\right) = -1$.

Encryption

- m the message
- $[C, c_1, c_2]$ the encrypted message, where

$$c_1 = \frac{1}{2}\left[1 - \left(\frac{m}{N}\right)\right] \quad , \quad \bar{m} = S^{c_1} m \bmod N \quad ,$$

$$c_2 = \bar{m} \bmod 2 \quad , \quad C = \bar{m}^2 \bmod N \quad .$$

Decryption

- compute $m' = C^D \bmod N$ and $N - m'$,
- choose the number, m'' say, with the parity specified by c_2.
- The original message is recovered as

$$m = S^{-c_1} m'' \quad .$$

A second scheme

Public-key

- $[N]$

Encryption

- m the message
- $[C, b_0, b_1]$ the encrypted message, where

$$C = m^2 \bmod N \ , \ b_0 = m \bmod 2 \ , \ b_1 = \frac{1}{2}\left[1 + \left(\frac{m}{N}\right)\right] \ .$$

Decryption

- compute the four roots, written as positive numbers;
- take the two roots having the same parity specified by b_0, say z_1 and z_2,
- compute the numbers $\frac{1}{2}\left[1 + \left(\frac{z_1}{N}\right)\right] \ , \ \frac{1}{2}\left[1 + \left(\frac{z_2}{N}\right)\right]$
- The original message is the root corresponding to the number equal to b_1.

A scheme based on Dedekind sums

Public-key

- $[N]$

Encryption

- m the message
- $[C, b_0, b_1]$ the encrypted message, where

  $$C = m^2 \bmod N \quad , \quad b_0 = m \bmod 2 \quad , \quad b_1 = s(m, N) \bmod 2 \quad ,$$

  The Dedekind sum can be taken modulo 2 since the denominator is odd. (Lemma (B))

Decryption

- compute the four roots, written as positive numbers;
- take the two roots having the same parity specified by b_0, say z_1 and z_2,
- compute the numbers $s(z_1, N) \bmod 2$, $s(z_2, N) \bmod 2$
- The original message is the root corresponding to the number equal to b_1 (Lemma (C)).

Root identification for any pair of primes

Root identification for any pair of primes

- If p and q are not both Blum primes, the identification of m among the four roots of the polynomial $x^2 - C$ can be given, as a consequence of Lemma (A), by the pair $[b_0, b_1]$ where

$$b_0 = x_i \bmod 2 \quad \text{and} \quad b_1 = (x_i \bmod p) + (x_i \bmod q) \bmod 2 \ .$$

Root identification for any pair of primes

- If p and q are not both Blum primes, the identification of m among the four roots of the polynomial $x^2 - C$ can be given, as a consequence of Lemma (A), by the pair $[b_0, b_1]$ where

$$b_0 = x_i \bmod 2 \quad \text{and} \quad b_1 = (x_i \bmod p) + (x_i \bmod q) \bmod 2 \ .$$

- The bit b_0 can be computed at the encryption stage without knowing p and q.

Root identification for any pair of primes

- If p and q are not both Blum primes, the identification of m among the four roots of the polynomial $x^2 - C$ can be given, as a consequence of Lemma (A), by the pair $[b_0, b_1]$ where

$$b_0 = x_i \bmod 2 \quad \text{and} \quad b_1 = (x_i \bmod p) + (x_i \bmod q) \bmod 2 \quad .$$

- The bit b_0 can be computed at the encryption stage without knowing p and q.
- The bit b_1 requires, in this definition, the knowledge of p and q and cannot be directly computed knowing only N.

List

- In principle, a way to get b_1 is to publish a pre-computed binary list (or table) that has in position i the bit b_1 pertaining to the message $m = i$.
- This list does not disclose any useful information on the factorization of N, because, even if we know that the residues modulo p and modulo q have the same parity, we do not know which parity, and if these residues have different parity we do not know which parity of which residue.
- The list makes the task theoretically feasible, although its size is of exponential complexity with respect to N and thus practically unrealizable.

Residuacity

- The Jacobi symbol, i.e. the quadratic residuacity, was used to distinguish the roots in the Rabin cryptosystem, when $p = q = 3 \bmod 4$.
- For primes congruent 1 modulo 4, Legendre symbols cannot distinguish numbers of opposite sign, therefore quadratic residuacity is not sufficient anymore to identify the roots.
- Higher power residue symbols could in principle do the desired job, but unfortunately their use unveils the factorization of N.

Polynomial function

- We may construct an identifying polynomial as an interpolation polynomial choosing a prime $P > N$.
- The polynomial

$$L(x) = \sum_{j=1}^{N-1} \left(1 - (x-j)^{P-1}\right) \left((j \bmod p) + (j \bmod q) \bmod 2\right)$$

  assumes the value 1 in $0 < m < N$, if the residues of m modulo p and modulo q have different parity, and assumes the value 0 elsewhere.
- Unfortunately, the complexity of $L(x)$ is prohibitive and makes this function practically useless.

Group isomorphism

- A possible solution is to use a function $\mathfrak{d}$ from $\mathbb{Z}_N$ into a finite group $\mathbb{G}$.

Group isomorphism

- A possible solution is to use a function $\mathfrak{d}$ from $\mathbb{Z}_N$ into a finite group $\mathbb{G}$.
- Define a function $\mathfrak{d}_1$ such that $\mathfrak{d}_1(x_1) = \mathfrak{d}(x_2)$.

Group isomorphism

- A possible solution is to use a function $\mathfrak{d}$ from $\mathbb{Z}_N$ into a finite group $\mathbb{G}$.
- Define a function $\mathfrak{d}_1$ such that $\mathfrak{d}_1(x_1) = \mathfrak{d}(x_2)$.
- The public key consists of the two functions $\mathfrak{d}$ and $\mathfrak{d}_1$.

Group isomorphism

- A possible solution is to use a function $\mathfrak{d}$ from $\mathbb{Z}_N$ into a finite group $\mathbb{G}$.
- Define a function $\mathfrak{d}_1$ such that $\mathfrak{d}_1(x_1) = \mathfrak{d}(x_2)$.
- The public key consists of the two functions $\mathfrak{d}$ and $\mathfrak{d}_1$.
- At the encryption stage both are evaluated (i.e. $\mathfrak{d}(m)$ and $\mathfrak{d}_1(m)$) and the minimum information necessary to distinguish their values is delivered together with the encrypted message. The decryption operations are obvious.

Group isomorphism

- A possible solution is to use a function $\mathfrak{d}$ from $\mathbb{Z}_N$ into a finite group $\mathbb{G}$.
- Define a function $\mathfrak{d}_1$ such that $\mathfrak{d}_1(x_1) = \mathfrak{d}(x_2)$.
- The public key consists of the two functions $\mathfrak{d}$ and $\mathfrak{d}_1$.
- At the encryption stage both are evaluated (i.e. $\mathfrak{d}(m)$ and $\mathfrak{d}_1(m)$) and the minimum information necessary to distinguish their values is delivered together with the encrypted message. The decryption operations are obvious.
- The true limitation of this scheme is that $\mathfrak{d}$ must be a one-way function, otherwise two square roots that allow us to factor N can be recovered as in the previous methods.

Group isomorphism

The following solution is based on the hardness of computing discrete logarithms.

- Given N, let $P = \mu N + 1$ be a prime (the smallest prime), that certainly exists by Dirichlet's theorem, that is congruent 1 modulo N. Let g be a primitive element generating the multiplicative group $\mathbb{Z}_P^*$.
- Define $g_1 = g^\mu$ and $g_2 = g^{\mu(\psi_1 - \psi_2)}$, and let m denote the message, as usual.
- The correspondence $x \leftrightarrow g_1^x$ defines an isomorphism between the additive group of $\mathbb{Z}_N$ and the cyclic subgroup of $\mathbb{Z}_P^*$ of order N.

Group isomorphism

Public-key

- $[N, g_1, g_2]$

Encryption

- m the message
- $[C, b_0, d_1, d_2, p_1, p_2]$ the encrypted message, where
  - $C = m^2 \bmod N, \;\; b_0 = m \bmod 2$,
  - p_1 is a position in the binary expansion of $g_1^m \bmod P$ whose bit d_1 is different from the bit in the corresponding position of the binary expansion of $g_2^m \bmod P$
  - p_2 is a position in the binary expansion of $g_1^m \bmod P$ whose bit d_2 is different from the bit in the corresponding position of the binary expansion of $g_2^{-m} \bmod P$.

Group isomorphism

Decryption

- compute the four roots, written as positive numbers;
- take the two roots having the same parity specified by b_0, say z_1 and z_2,
- Compute $A = g_1^{z_1} \bmod P$ and $B = g_1^{z_2} \bmod P$,
- Select the root that has the correct bits d_1 and d_2 in both the given position p_1 and p_2 of the binary expansion of A or B.

A Lemma

Lemma (D)

- *The power* $g_0 = g^\mu$ *generates a group of order* N *in* $\mathbb{Z}_P^*$*, thus the correspondence* $x \leftrightarrow g_0^x$ *establishes an isomorphism between a multiplicative subgroup of* $\mathbb{Z}_P^*$ *and the additive group of* $\mathbb{Z}_N$*.*
- *The four roots of* $x^2 = C \bmod N$*,* $C = m^2 \bmod N$ *are in a one-to-one correspondence with the four powers* $g_0^m \bmod P$*,* $g_0^{-m} \bmod P$*,* $g_0^{m(\psi_1-\psi_2)} \bmod P$ *and* $g_0^{-m(\psi_1-\psi_2)} \bmod P$*.*

Rabin signature

Public-key

- The Rabin scheme may also be used to sign a message m:
  - Let S be any root of $x^2 = m \bmod N$
  - The signature is the pair $[m, S]$
  - If the quadratic equation is not solvable, i.e. either $f_1 = \left(\frac{m}{p}\right) = -1$, or $f_2 = \left(\frac{m}{q}\right) = -1$, or both f_1 and f_2 are -1, a random padding factor U is used until $x^2 = mU \bmod N$ can be solved,
  - The signature is the triple $[m, U, S]$
- A different scheme is the Rabin-Williams signature.

We propose a Rabin signature that makes use of a deterministic padding factor.

Rabin signature

Public-key

- $[N]$

Signed message

- $[U, m, S]$, where
- $U = R^2 (f_1\psi_1 + f_2\psi_2) \bmod N$ is the padding factor, with
- R a random number, and S is any solution of the equation $x^2 = mU \bmod N$

Verification

- compute $mU \bmod N$ and $S^2 \bmod N$;
- the signature is valid if and only if these two numbers are equal.

Rabin signature

Rabin signatures with padding factors have several features

1. the signature can be done using every pair of primes, therefore it could be used with the modulo of any RSA public key, for example;
2. different signatures of the same document are different;
3. the verification needs only two multiplications, therefore it is fast enough to be used in authentication protocols.

Deterministic padding is faster than random padding and has fixed delay.

Conclusions

1) the root identification requires the delivery of additional information, which may not be easily computed, especially when not both primes are Blum primes;
2) many proposed root identification methods, based on the message semantics, have a naive character and cannot be used in many circumstances;
3) the delivery of two bits together with the encrypted message exposes the process to active attacks by maliciously modifying these bits.

Conclusions

- The Rabin scheme may come with some hindrance when used to conceal a message,
- The Rabin scheme seems effective when applied to generate electronic signature or as a hash function.

Thank you for your attention!

References

1. R. Dedekind, Schreiben an Herrn Borchardt, *J. Reine Angew. Math.*, 83, 1877, pp.265-292.
2. M. Rabin, Digitalized signature as intractable as factorization, *Technical Report MIT/LCS/TR-212*, MIT Laboratory for Computer Science, January 1978.
3. T.M. Apostol, *Introduction to Analytic Number Theory*, Springer, New York, 1976.
4. J. Hoffstein, J. Pipher, J.H. Silverman, *An introduction to mathematical cryptography*, Springer, New York, 2008.
5. J. Pieprzyk, T. Hardjono, J. Seberry, *Fundamentals of Computer Security*, Springer, New York, 2003.