Security Proof for Partial-Domain Hash Signature Schemes

Jean-Sébastien Coron
Gemplus Card International
34 rue Guynemer
Issy-les-Moulineaux, F-92447, France
jean-sebastien.coron@gemplus.com

Abstract. We study the security of partial-domain hash signature schemes, in which the output size of the hash function is only a fraction of the modulus size. We show that for $e = 2$ (Rabin), partial-domain hash signature schemes are provably secure in the random oracle model, if the output size of the hash function is larger than $2/3$ of the modulus size. This provides a security proof for a variant of the signature standards ISO 9796-2 and PKCS#1 v1.5, in which a larger digest size is used.

Key-words: Signature Schemes, Provable Security, Random Oracle Model.

1 Introduction

A common practice for signing with RSA or Rabin consists in first hashing the message m, then padding the hash value with some predetermined or message-dependent block, and eventually raising the result $\mu(m)$ to the private exponent d. This is commonly referred to as the "hash-and-sign" paradigm:

$$s = \mu(m)^d \mod N$$

For digital signature schemes, the strongest security notion was defined by Goldwasser, Micali and Rivest in [**?**], as *existential unforgeability under an adaptive chosen message attack*. This notion captures the property that an attacker cannot produce a valid signature, even after obtaining the signature of (polynomially many) messages of his choice.

The random oracle model, introduced by Bellare and Rogaway in [**?**], is a theoretical framework allowing to prove the security of hash-and-sign signature schemes. In this model, the hash function is seen as an oracle which outputs a random value for each new query. Bellare and Rogaway defined in [**?**] the Full Domain Hash (FDH) signature scheme, in which the output size of the hash function is the same as the modulus size. FDH is provably secure in the random oracle model assuming that inverting RSA is hard. Actually, a security proof in the random oracle model does not necessarily imply that the scheme is secure in the real world (see [?]). Nevertheless, it seems to be a good engineering principle to design a scheme so that it is provably secure in the random oracle model. Many encryption and signature schemes were proven to be secure in the random oracle model

Other hash-and-sign signature schemes include the widely used signature standards PKCS#1 v1.5 and ISO 9796-2. In these standards, the digest size is only a fraction of the modulus size. As opposed to FDH, no security proof is known for those standards. Moreover, it was shown in [?] that ISO 9796-2 was insecure if the size of the hash function was too small, and the standard was subsequently revised.

In this paper, we study the security of partial-domain hash signature schemes, in which the hash size is only a fraction of the modulus size. We show that for $e = 2$, partial-domain hash signature schemes are provably secure in the random oracle model, assuming that factoring is hard, if the size of the hash function is larger than 2/3 of the modulus size. The proof is based on a modification of Vallée's generator of small random squares [?]. This provides a security proof for a variant of PKCS#1 v1.5 and ISO 9796-2 signatures, in which the digest size is larger than 2/3 of the size of the modulus.

2 Definitions

In this section we briefly present some notations and definitions used throughout the paper. We start by recalling the definition of a signature scheme.

Definition 1 (Signature Scheme). *A signature scheme* $(\texttt{Gen}, \texttt{Sign}, \texttt{Verify})$ *is defined as follows:*

- The key generation algorithm `Gen` *is a probabilistic algorithm which given* 1^k*, outputs a pair of matching public and private keys,* (pk, sk)*.*

- The signing algorithm `Sign` *takes the message* M *to be signed, the private key* sk*, and returns a signature* $x = \texttt{Sign}_{sk}(M)$*. The signing algorithm may be probabilistic.*

- The verification algorithm `Verify` *takes a message* M*, a candidate signature* x' *and* pk*. It returns a bit* $\texttt{Verify}_{pk}(M, x')$*, equal to one if the signature is accepted, and zero otherwise. We require that if* $x \leftarrow \texttt{Sign}_{sk}(M)$*, then* $\texttt{Verify}_{pk}(M, x) = 1$*.*

In the previously introduced *existential unforgeability under an adaptive chosen message attack* scenario, the forger can dynamically obtain signatures of messages of his choice and attempt to output a valid forgery. A *valid forgery* is a message/signature pair (M, x) such that $\texttt{Verify}_{pk}(M, x) = 1$ whereas the signature of M was never requested by the forger. Moreover, in the random oracle model, the attacker cannot evaluate the hash function by himself; instead, he queries an oracle which outputs a random value for each new query.

RSA [?] is undoubtedly the most widely used cryptosystem today:

Definition 2 (RSA). *The RSA cryptosystem is a family of trapdoor permutations, specified by:*
- The RSA generator $\mathcal{RSA}$, which on input 1^k, randomly selects two distinct $k/2$-bit primes p and q and computes the modulus $N = p \cdot q$. It picks an encryption exponent $e \in \mathbb{Z}^*_{\phi(N)}$ and computes the corresponding decryption exponent d such that $e \cdot d = 1 \mod \phi(N)$. The generator returns (N, e, d).
- The encryption function $f : \mathbb{Z}^*_N \to \mathbb{Z}^*_N$ defined by $f(x) = x^e \mod N$.
- The decryption function $f^{-1} : \mathbb{Z}^*_N \to \mathbb{Z}^*_N$ defined by $f^{-1}(y) = y^d \mod N$.

An *inverting algorithm* $\mathcal{I}$ for RSA gets as input (N, e, y) and tries to find $y^d \mod N$. Its success probability is the probability to output $y^d \mod N$ when (N, e, d) are obtained by running $\mathcal{RSA}(1^k)$ and y is set to $x^e \mod N$ for some x chosen at random in $\mathbb{Z}^*_N$.

The Full-Domain-Hash scheme (FDH) [?] was the first practical and provably secure signature scheme based on RSA. It is defined as follows: the key generation algorithm, on input 1^k, runs $\mathcal{RSA}(1^k)$ to obtain (N, e, d). It outputs (pk, sk), where the public key pk is (N, e) and the private key sk is (N, d). The signing and verifying algorithms use a hash function $H : \{0,1\}^* \to \mathbb{Z}^*_N$ which maps bit strings of arbitrary length to the set of invertible integers modulo N.

$\texttt{SignFDH}_{N,d}(M)$	$\texttt{VerifyFDH}_{N,e}(M, x)$
$y \leftarrow H(M)$	$y \leftarrow x^e \mod N$
return $y^d \mod N$	if $y = H(M)$ then return 1 else return 0.

The following theorem [?] proves the security of FDH in the random oracle model, assuming that inverting RSA is hard. It provides a better security bound than [?].

Theorem 1. *Assume that there is no algorithm which inverts RSA with probability greater than ε within time t. Then the success probability of a FDH forger making at most q_{hash} hash queries and q_{sig} signature queries within running time t' is less than ε', where*

$$\begin{aligned} \varepsilon' &= 4 \cdot q_{sig} \cdot \varepsilon \\ t' &= t - (q_{hash} + q_{sig} + 1) \cdot \mathcal{O}(k^3) \end{aligned}$$

We say that a hash-and-sign signature scheme is a *partial-domain hash signature scheme* if the encoding function $\mu(m)$ can be written as:

$$\mu(m) = \gamma \cdot H(m) + f(m) \tag{1}$$

where γ is a constant, H a hash function and f some function of m. A typical example of a partial-domain hash signature scheme is the ISO 9796-2 standard with full message recovery [?]:

$$\mu(m) = \texttt{4A}_{\texttt{16}} \| m \| H(m) \| \texttt{BC}_{\texttt{16}}$$

The main result of this paper is to show that for $e = 2$, partial-domain hash signature schemes are provably secure, if the hash size is larger than 2/3 of the modulus size. In the following, we recall the Rabin-Williams signature scheme [?]. It uses a padding function $\mu(m)$ such that for all m, $\mu(m) = 6 \mod 16$.

- Key generation: on input 1^k, generate two $k/2$-bit primes p and q such that $p = 3 \mod 8$ and $q = 7 \mod 8$. The public key is $N = p \cdot q$ and the private key is $d = (N - p - q + 5)/8$.

- Signature generation: compute the Jacobi symbol

$$J = \left(\frac{\mu(m)}{N}\right)$$

The signature of m is $s = \min(\sigma, N - \sigma)$, where:

$$\sigma = \begin{cases} \mu(m)^d \mod N & \text{if } J = 1 \\ (\mu(m)/2)^d \mod N & \text{otherwise} \end{cases}$$

- Signature verification: compute $\omega = s^2 \mod N$ and check that:

$$\mu(m) \stackrel{?}{=} \begin{cases} \omega & \text{if } \omega = 6 \mod 8 \\ 2 \cdot \omega & \text{if } \omega = 3 \mod 8 \\ N - \omega & \text{if } \omega = 7 \mod 8 \\ 2 \cdot (N - \omega) & \text{if } \omega = 2 \mod 8 \end{cases}$$

3 Security of Partial-domain Hash Signature Schemes

To prove the security of a signature scheme against chosen message attacks, one must be able to answer the signature queries of the attacker. In FDH's security proof, when answering a hash query, one generates a random $r \in \mathbb{Z}_N$ and answers $H(m) = r^e \mod N$ so that the signature r of m is known. Similarly, for partial-domain hash signature schemes, we should be able to generate a random r such that:

$$\mu(m) = \gamma \cdot H(m) + f(m) = r^e \mod N$$

with $H(m)$ being uniformly distributed in the output space of the hash function. For example, if we take $\mu(m) = H(m)$ where $0 \leq H(m) \leq N^\beta$ and $\beta < 1$, one should be able to generate a random r such that $r^e \mod N$ is uniformly distributed between 0 and N^β.

Up to our knowledge, no such algorithm is known for $e \geq 3$. For $e = 2$, Vallée constructed in [?] a random generator where the size of $r^2 \mod N$ is less than 2/3 of the size of the modulus. [?] used this generator to obtain proven complexity bounds for the quadratic sieve factoring algorithm. Vallée's generator has a quasi-uniform distribution; a distribution is said to be *quasi-uniform* if there is a constant ℓ such that for all x, the probability to generate x lies between $1/\ell$ and ℓ times the probability to generate x under the uniform distribution. However, quasi-uniformity is not sufficient here, as we must simulate a random oracle and therefore our simulation should be indistinguishable from the uniform distribution.

Our contribution is to modify Vallée's generator in order to generate random squares in any interval of size $N^{2/3+\varepsilon}$, with a distribution which is statistically indistinguishable from the uniform distribution. From this generator we will derive a security proof for partial-domain hash signatures, in which the digest size is at least 2/3 of the modulus size.

Remark: for Paillier's trapdoor permutation [?] with parameter $g = 1 + N$, it is easy to show that half-domain hash is provably secure in the random oracle model, assuming that inverting RSA with $e = N$ is hard.

4 Generating Random Squares in a Given Interval

4.1 Notations

We identify $\mathbb{Z}_N$, the ring of integers modulo N with the set of integers between 0 and $N-1$. We denote by $\mathbb{Z}_N^+$ the set of integers between 0 and $(N-1)/2$. We denote by Q the squaring operation over $\mathbb{Z}_N$:

$$Q(x) = x^2 \mod N$$

Given positive integers a and h such that $a + h < N$, let B be the set:

$$B = \{x \in \mathbb{Z}_N^+ \mid a \leq Q(x) \leq a + h\}$$

Our goal is to generate integers $x \in B$ with a distribution statistically indistinguishable from the uniform distribution. The *statistical distance* between two distributions X and Y is defined as the function:

$$\delta = \frac{1}{2} \sum_{\alpha} |\Pr[X = \alpha] - \Pr[Y = \alpha]|$$

We say that two ensembles $X = \{X_n\}_{n \in \mathbb{N}}$ and $Y = \{Y_n\}_{n \in \mathbb{N}}$ are *statistically indistinguishable* if their statistical distance δ_n is a negligible function of n.

4.2 Description of *B*

In this section, we recall Vallée's description of the set B. We denote by b the cardinality of B. The following lemma, which proof can be derived from equation (6) in [?], shows that b is close to $h/2$.

Lemma 1. *Let N be a ℓ-bit RSA modulus. We have for $\ell \geq 64$:*

$$\left| b - \frac{h}{2} \right| \leq 4 \cdot \ell \cdot 2^{\ell/2}$$

In the following, we assume that the bit size of N is greater than 64. As in [?], we introduce Farey sequences [?]:

Definition 3 (Farey sequence). *The Farey sequence $\mathcal{F}_k$ of order k is the ascending sequence of irreducible fractions between* 0 *and* 1 *whose denominators do not exceed k. Thus p/q belongs to $\mathcal{F}_k$ if $0 \le p \le q \le k$ and* $\gcd(p,q) = 1$.

The characteristic property of Farey sequences is expressed by the following theorem [**?**]:

Theorem 2. *If p/q and p'/q' are two successive terms of $\mathcal{F}_k$, then $q \cdot p' - p \cdot q' = 1$*

Given $p/q \in \mathcal{F}_k$, we define the *Farey interval* $I(p,q)$ as the interval of center $pN/(2q)$ and radius $N/(2kq)$. Given the terms p'/q' and p''/q'' of $\mathcal{F}_k$ which precede and follow p/q, we let $J(p,q)$ be the interval:

$$J(p,q) = \left[\frac{N(p+p')}{2(q+q')}, \frac{N(p+p'')}{2(q+q'')} \right]$$

If $p/q = 0/1$, then p/q has no predecessor and we take $p'/q' = 0/1$. Similarly, if $p/q = 1/1$, we take $p''/q'' = 1/1$. The set of intervals $J(p,q)$ forms a partition of $\mathbb{Z}_N^+$. The following lemma [**?**] shows that intervals $I(p,q)$ and $J(p,q)$ are closely related.

Lemma 2. *$I(p,q)$ contains $J(p,q)$ and its length is at most twice the length of $J(p,q)$.*

Given $p/q \in \mathcal{F}_k$ with $p/q \neq 0/1$, let x_0 be the integer nearest to the rational $pN/2q$:

$$x_0 - \frac{pN}{2q} = u_0 \quad \text{with } |u_0| \le \frac{1}{2}$$

Let $L(x_0)$ be the lattice spanned by the two vectors $(1, 2x_0)$ and $(0, N)$. Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be the two parabolas of equations:

$$\mathcal{P}_1 : \omega + u^2 + x_0^2 = a + h \quad \text{and} \quad \mathcal{P}_2 : \omega + u^2 + x_0^2 = a$$

Let P be the domain of lattice points comprised between the two parabolas:

$$P = \{(u, \omega) \in L(x_0) \mid a \le \omega + u^2 + x_0^2 \le a + h\}$$

The following lemma, which proof is straightforward, shows that the elements of B arise from the intersection of the lattice $L(x_0)$ and the domain comprised between the two parabolas (see figure **??**).

Lemma 3. *$x = x_0 + u$ belongs to B iff there exists a unique ω such that the point (u, ω) belongs to P.*

We let $B(p,q)$ be the set of integers in $B \cap J(p,q)$. From Lemma **??** the integers in $B(p,q)$ arise from the domain of lattice points:

$$P(p,q) = \{(u, \omega) \in P \mid x_0 + u \in J(p,q)\}$$

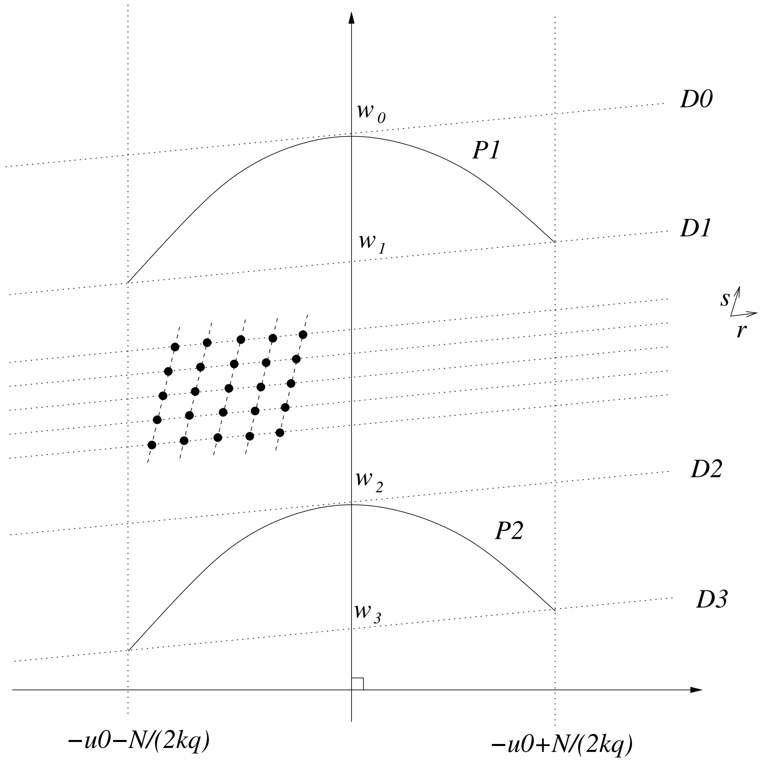

Fig. 1. The intersection between the lattice $L(x_0)$ and the domain between the two parabolas $\mathcal{P}_1$ and $\mathcal{P}_2$

From Lemma **??**, the set $P(p,q)$ is included inside the set of lattice points:

$$Q(p,q) = \{(u,\omega) \in P \mid x_0 + u \in I(p,q)\}$$

whose abcissae u are comprised between $-u_0 - N/(2kq)$ and $-u_0 + N/(2kq)$. In the following, we describe the domain $Q(p,q)$, using the following two short vectors of $L(x_0)$ (see figure **??**):

$$\boldsymbol{r} = q(1, 2x_0) - p(0, N) = (q, 2qu_0) \tag{2}$$

$$\boldsymbol{s} = q'(1, 2x_0) - p'(0, N) = (q', 2q'u_0 + N/q) \tag{3}$$

where p'/q' is the term of $\mathcal{F}_k$ which precedes p/q.

We consider the lines of the lattice parallel to vector $\boldsymbol{r}$ which intersect the domain $Q(p,q)$. These lines have a slope equal to $2u_0$. The first extremal position of these lines is the tangent D_0 to the first parabola:

$$D_0 : \ \omega - (-u_0^2 - x_0^2 + a + h) = 2u_0(u + u_0)$$

The second extremal position joins the two points of the second parabola with abscissae $-u_0-N/(2kq)$ and $-u_0+N/(2kq)$. This line D_3 has also a slope equal to $2u_0$ and satisfies the equation:

$$\omega+(u_0+\frac{N}{2kq})^2-a+x_0^2=2u_0(u+u_0+\frac{N}{2kq})$$

The two lines intersect the vertical axis at the respective points:

$$\omega_0=a-x_0^2+u_0^2+h \quad\text{and}\quad \omega_3=a-x_0^2+u_0^2-\frac{N^2}{4k^2q^2}$$

All the lines parallel to $\boldsymbol{r}$ that intersect $P(p,q)$ are the ones that intersect the segment $[\omega_3,\omega_0]$ on the vertical axis. We denote by $D(\nu)$ a line parallel to $\boldsymbol{r}$ which intersects the vertical axis at ordinate equal to $\omega_0-\nu N/q$. The line D_0 is the line $D(\nu_0=0)$, whereas the line D_3 is the line $D(\nu_3)$ such that:

$$\nu_3=\frac{hq}{N}+\frac{N}{4k^2q} \tag{4}$$

Eventually, we denote by $D_1=D(\nu_1)$ the line which joins the two points of the first parabola with abcissae $-u_0-N/(2kq)$ and $-u_0+N/(2kq)$, and by $D_2=D(\nu_2)$ the tangent to the second parabola, with a slope equal to $2u_0$. We have:

$$\nu_1=\frac{N}{4k^2q} \quad\text{and}\quad \nu_2=\frac{hq}{N} \tag{5}$$

A real ν is called an index if $D(\nu)$ is a line of $L(x_0)$. The difference between two consecutive indices is equal to one.

4.3 Our New Generator

In this section, we describe our new generator of integers in B. The difference with Vallée's generator is that we use different parameters for k and h, and we do not generate all the integers in B; instead we avoid a negligible subset of B.

First, we describe a generator $\mathcal{G}(p,q)$ of integers in $B(p,q)$, and we show that its distribution is statistically indistinguishable from the uniform distribution. We assume that $N\leq 2\cdot k\cdot q\cdot\sqrt{h}$, which gives $\nu_1\leq\nu_2$. Therefore the line D_1 is above the line D_2 (see figure **??**). We restrict ourselves to the integers in $B(p,q)$ such that the corresponding points $(u,\omega)\in P(p,q)$ lie on $D(\nu)$ with $\nu_1\leq\nu\leq\nu_2$. These points are the points on $D(\nu)$ whose abscissae u are such that $x_0+u\in J(p,q)$.

Generator $\mathcal{G}(p,q)$ of integers in $B(p,q)$:
1. Generate a random index ν uniformly distributed between ν_1 and ν_2.
2. Generate a point $(u,\omega)\in P(p,q)$ on $D(\nu)$ such that $x_0+u\in J(p,q)$, with the uniform distribution.
3. Output x_0+u.

The following lemma shows that under some conditions on k, h and q, the cardinality $b(p,q)$ of $B(p,q)$ is close to $h \cdot j(p,q)/N$, where $j(p,q)$ is the number of integers in the interval $J(p,q)$. Moreover, under the same conditions, the distribution induced by $\mathcal{G}(p,q)$ is statistically indistinguishable from the uniform distribution in $B(p,q)$. The proof is given in appendix **??**.

Lemma 4. *Let $\alpha > 0$ and $k = N^{\frac{1}{3}-\alpha}$. Assume that $k \geq 6$, $N^\alpha \geq 3$ and $N^{\frac{2}{3}+13\cdot\alpha} \leq h < N$. Then for all $p/q \in \mathcal{F}_k$ such that $N^{1/3-4\alpha} \leq q \leq k$, we have:*

$$\left| b(p,q) - \frac{h \cdot j(p,q)}{N} \right| \leq \frac{4h \cdot j(p,q)}{N} N^{-3\alpha} \tag{6}$$

Moreover, $\mathcal{G}(p,q)$ generates elements in $B(p,q)$ with a distribution whose distance δ_G from the uniform distribution is at most $7 \cdot N^{-3\alpha}$.

Now we construct a generator $\mathcal{V}$ of $p/q \in \mathcal{F}_k$ such that the probability to generate p/q is close to $b(p,q)/b$. It only generates $p/q \in \mathcal{F}_k$ such that $q \geq N^{1/3-4\alpha}$, so that from the previous lemma, $b(p,q)$ is nearly proportional to the number of integers in $J(p,q)$, and the distribution induced by $\mathcal{G}(p,q)$ is close to the uniform distribution.

Generator $\mathcal{V}$ of $p/q \in \mathcal{F}_k$
1. Generate a random integer $x \in \mathbb{Z}_N^+$ with the uniform distribution.
2. Determine which interval $J(p,q)$ contains x.
3. If $q \geq N^{1/3-4\alpha}$ then output $p/q \in \mathcal{F}_k$, otherwise output $\perp$.

Lemma 5. *Let denote by $\mathcal{D}$ the distribution induced by choosing $p/q \in \mathcal{F}_k$ with probability $b(p,q)/b$. Under the conditions of lemma* **??**, *the statistical distance δ_V between $\mathcal{D}$ and the distribution induced by $\mathcal{V}$ is at most $9 \cdot N^{-3\alpha}$.*

Proof. See appendix **??**.

Eventually, our generator $\mathcal{G}$ of elements in B combines the two generators $\mathcal{V}$ and $\mathcal{G}(p,q)$:

Generator $\mathcal{G}$ of $x \in B$
1. Generate y using $\mathcal{V}$.
2. If $y = \perp$, then output $\perp$.
3. Otherwise, $y = p/q$ and generate $x \in B(p,q)$ using $\mathcal{G}(p,q)$. Output x.

The following theorem, whose proof is given in appendix **??**, shows that the distribution induced by $\mathcal{G}$ is statistically indistinguishable from the uniform distribution in B.

Theorem 3. *For any $\varepsilon > 0$, letting $h = N^{\frac{2}{3}+\varepsilon}$ and $\alpha = \varepsilon/13$. If $N^\alpha \geq 3$, then the distance δ between the distribution induced by $\mathcal{G}$ and the uniform distribution in B is at most $16 \cdot N^{-3\cdot\varepsilon/13}$. The running time of $\mathcal{G}$ is $\mathcal{O}(\log^3 N)$.*

5 A Security Proof for Partial-domain Hash Signature Schemes

In this section, using the previous generator $\mathcal{G}$ of random squares, we show that partial-domain hash signature schemes are provably secure in the random oracle model, for $e = 2$, assuming that factoring is hard, if the size of the hash function is larger than 2/3 of the modulus size. Moreover, we restrict ourselves to small constants γ in (**??**), *e.g.* $\gamma = 16$ or $\gamma = 256$. This is the case for all the signature standards of the next section. We denote by k_0 the hash function's digest size. The proof is similar to the proof of theorem **??** and is given in the full version of this paper [**?**].

Theorem 4. *Let $\mathcal{S}$ be the Rabin-Williams partial-domain hash signature scheme with constant γ and hash size k_0 bits. Assume that there is no algorithm which factors a RSA modulus with probability greater than ε within time t. Then the success probability of a forger against $\mathcal{S}$ making at most q_{hash} hash queries and q_{sig} signature queries within time t' is upper bounded by ε', where:*

$$\varepsilon' = 8 \cdot q_{sig} \cdot \varepsilon + 32 \cdot (q_{hash} + q_{sig} + 1) \cdot k_1 \cdot \gamma \cdot 2^{-\frac{3}{13} \cdot k_1} \tag{7}$$

$$t' = t - k_1 \cdot \gamma \cdot (q_{hash} + q_{sig} + 1) \cdot \mathcal{O}(k^3) \tag{8}$$

and $k_1 = k_0 - \frac{2}{3}k$.

6 Application to Signature Standards

6.1 PKCS#1 v1.5 and SSL-3.02

The signature scheme PKCS#1 v1.5 [**?**] is a partial-domain hash signature scheme, with:

$$\mu(m) = \mathtt{0001_{16}} \| \mathtt{FFFF_{16}} \ldots \mathtt{FFFF_{16}} \| \mathtt{00_{16}} \| c_{\text{SHA}} \| H(m)$$

where c_{SHA} is a constant and $H(m) = \text{SHA}(m)$, or

$$\mu(m) = \mathtt{0001_{16}} \| \mathtt{FFFF_{16}} \ldots \mathtt{FFFF_{16}} \| \mathtt{00_{16}} \| c_{\text{MD5}} \| H(m)$$

where c_{MD5} is a constant and $H(m) = \text{MD5}(m)$.

The standard PKCS#1 v1.5 was not designed to work with Rabin ($e = 2$). However, one can replace the last nibble of $H(m)$ by 6 and obtain a padding scheme which is compatible with the Rabin-Williams signature scheme. The standard is then provably secure if the size of the hash-function is larger than 2/3 of the size of the modulus. This is much larger than the 128 or 160 bits which are recommended in the standard. The same analysis applies for the SSL-3.02 padding scheme [**?**].

6.2 ISO 9796-2 and ANSI x9.31

The ISO 9796-2 encoding scheme [?] is defined as follows:

$$\mu(m) = \mathtt{6A}_{\mathtt{16}} \| m[1] \| H(m) \| \mathtt{BC}_{\mathtt{16}}$$

where $m[1]$ is the leftmost part of the message, or:

$$\mu(m) = \mathtt{4A}_{\mathtt{16}} \| m \| H(m) \| \mathtt{BC}_{\mathtt{16}}$$

[?] describes an application of ISO 9796-2 with the Rabin-Williams signature scheme. Note that since $\mu(m) = 12 \mod 16$ instead of $\mu(m) = 6 \mod 16$, there is a slight change in the verification process. However, the same security bound applies: the scheme is provably secure if the size of the hash-function is larger than 2/3 of the size of the modulus. The same analysis applies for the ANSI x9.31 padding scheme [?].

7 Conclusion

We have shown that for Rabin, partial-domain hash signature schemes are provably secure in the random oracle, assuming that factoring is hard, if the size of the hash function is larger than 2/3 of the modulus size. Unfortunately, this is much larger than the size which is recommended in the standards PKCS#1 v1.5 and ISO 9796-2. An open problem is to obtain a smaller bound for the digest size, and to extend this result to RSA signatures.

Acknowledgements

I wish to thanks the anonymous referees for their helpful comments.

References

1. ANSI X9.31, *Digital signatures using reversible public-key cryptography for the financial services industry (rDSA)*, 1998.
2. M. Bellare and P. Rogaway, *Random oracles are practical : a paradigm for designing efficient protocols*. Proceedings of the First Annual Conference on Computer and Commmunications Security, ACM, 1993.
3. M. Bellare and P. Rogaway, *The exact security of digital signatures - How to sign with RSA and Rabin*. Proceedings of Eurocrypt'96, LNCS vol. 1070, Springer-Verlag, 1996, pp. 399-416.
4. R. Canetti, O. Goldreich and S. Halevi, *The random oracle methodology, revisited*, STOC' 98, ACM, 1998.
5. J.S. Coron, D. Naccache and J.P. Stern, *On the security of RSA Padding*, Proceedings of Crypto'99, LNCS vol. 1666, Springer-Verlag, 1999, pp. 1-18.
6. J.S. Coron, *On the exact security of Full Domain Hash*, Proceedings of Crypto 2000, LNCS vol. 1880, Springer-Verlag, 2000, pp. 229-235.

7. J.S. Coron, *Security proof for partial-domain hash signature schemes.* Full version of this paper. Cryptology ePrint Archive, `http://eprint.iacr.org`.
8. S. Goldwasser, S. Micali and R. Rivest, *A digital signature scheme secure against adaptive chosen-message attacks*, SIAM Journal of computing, 17(2):281-308, april 1988.
9. G.H. Hardy and E.M. Wright, *An introduction to the theory of numbers*, Oxford science publications, fifth edition.
10. K. Hickman, *The SSL Protocol*, December 1995. Available electronically at : `http://www.netscape.com/newsref/std/ssl.html`
11. ISO/IEC 9796-2, *Information technology - Security techniques - Digital signature scheme giving message recovery, Part 2 : Mechanisms using a hash-function*, 1997.
12. A.J. Menezes, P. C. van Oorschot and S.A. Vanstone, *Handbook of Applied Cryptography*, CRC press, 1996.
13. P. Paillier, *Public-key cryptosystems based on composite degree residuosity classes*, proceedings of Eurocrypt'99, LNCS 1592, pp. 223-238, 1999.
14. R. Rivest, A. Shamir and L. Adleman, *A method for obtaining digital signatures and public key cryptosystems*, CACM 21, 1978.
15. RSA Laboratories, PKCS #1 : *RSA cryptography specifications*, version 1.5, November 1993 and version 2.0, September 1998.
16. B. Vallée, *Generation of elements with small modular squares and provably fast integer factoring algorithms*, Mathematics of Computation, vol. 56, number 194, april 1991, pp. 823-849.

A Proof of lemma ??

From the conditions of lemma **??**, we obtain:

$$\frac{hq}{N} \geq N^{9\alpha} \quad \text{and} \quad \frac{N}{k^2 q} \leq N^{6\alpha} \tag{9}$$

which gives $N \leq 2 \cdot k \cdot q \cdot \sqrt{h}$ and then $\nu_1 < \nu_2$.

Recall that $j(p,q)$ denotes the number of integers in interval $J(p,q)$. From lemma **??** the length of $J(p,q)$ is at least $N/(2kq)$ and therefore, $j(p,q) \geq N/(2kq) - 1$, which gives using $k \geq 6$:

$$\frac{j(p,q)}{q} \geq \frac{N^{3\alpha}}{3} \tag{10}$$

Let us denote by $n(\nu)$ the number of points of $P(p,q)$ on a line $D(\nu)$. The distance between the abcissae of two consecutive points of $P(p,q)$ on a line $D(\nu)$ is equal to q. Therefore, for all indices ν, we have $n(\nu) \leq \lfloor j(p,q)/q \rfloor + 1$. Moreover, for $\nu_1 \leq \nu \leq \nu_2$, $n(\nu)$ is either $\lfloor j(p,q)/q \rfloor$ or $\lfloor j(p,q)/q \rfloor + 1$. This gives the following bound for $b(p,q)$:

$$(\nu_2 - \nu_1 - 1) \cdot \left(\frac{j(p,q)}{q} - 1 \right) \leq b(p,q) \leq (\nu_3 + 1) \cdot \left(\frac{j(p,q)}{q} + 1 \right)$$

which gives using (**??**), (**??**), (**??**), (**??**) and $N^\alpha \geq 3$:

$$\left| b(p,q) - \frac{h \cdot j(p,q)}{N} \right| \leq \frac{4h \cdot j(p,q)}{N} N^{-3\alpha} \tag{11}$$

Let n' be the number of indices ν such that $\nu_1 \leq \nu \leq \nu_2$. We have $n' = \lfloor \nu_2 - \nu_1 \rfloor$ or $n' = \lfloor \nu_2 - \nu_1 \rfloor + 1$. The probability that $\mathcal{G}(p,q)$ generates an element $x \in B(p,q)$ corresponding to a point of index ν is given by:

$$\Pr[x] = P(\nu) = \frac{1}{n' \cdot n(\nu)}$$

for $\nu_1 \leq \nu \leq \nu_2$ and $P(\nu) = 0$ otherwise. The number of integers $x \in B(p,q)$ such that $\Pr[x] = 0$ is then at most:

$$(\nu_1 + \nu_3 - \nu_2 + 2) \cdot \left(\frac{j(p,q)}{q} + 1 \right) \leq N^{6\alpha} \cdot \frac{j(p,q)}{q} \tag{12}$$

For all $\nu_1 \leq \nu \leq \nu_2$, we have using (**??**), (**??**), (**??**), (**??**), (**??**) and $N^\alpha \geq 3$:

$$\left| P(\nu) - \frac{1}{b(p,q)} \right| \leq 10 \cdot \frac{N}{h \cdot j(p,q)} \cdot N^{-3\alpha} \tag{13}$$

Eventually, the statistical distance from the uniform distribution is:

$$\delta_G = \frac{1}{2} \sum_{x \in B(p,q)} \left| \Pr[x] - \frac{1}{b(p,q)} \right|$$

and we obtain using (**??**), (**??**) and (**??**):

$$\delta_G \leq 7 \cdot N^{-3\alpha}$$

B Proof of lemma ??

Let us denote $q_m = N^{1/3-4\alpha}$. For $q \geq q_m$, the probability to generate $p/q \in \mathcal{F}_k$ using $\mathcal{V}$ is $j(p,q)/|\mathbb{Z}_N^+|$. Moreover, using lemma **??**, the probability that $\mathcal{V}$ generates $\perp$ is at most:

$$\Pr[\perp] = \sum_{\mathcal{F}_k | q < q_m} \frac{2 \cdot j(p,q)}{N+1} \leq 3 \frac{q_m}{k} \leq 3 \cdot N^{-3\alpha} \tag{14}$$

Consequently, the statistical distance δ_V between $\mathcal{D}$ and the distribution induced by $\mathcal{V}$ is at most:

$$\delta_V = \frac{1}{2} \sum_{\mathcal{F}_k | q \geq q_m} \left| \frac{2 \cdot j(p,q)}{N+1} - \frac{b(p,q)}{b} \right| + \frac{1}{2} \Pr[\perp] + \frac{1}{2} \sum_{\mathcal{F}_k | q < q_m} \frac{b(p,q)}{b} \tag{15}$$

Let ℓ be the size of N in bits. From lemma **??**, we obtain for $\ell \geq 64$:

$$|b - \frac{h}{2}| \leq 4 \cdot \ell \cdot 2^{\ell/2} \leq \frac{1}{2} \cdot N^{2/3} \leq \frac{h}{2} \cdot N^{-3\alpha} \tag{16}$$

For $q \geq q_m$, we obtain from Lemma **??** and (**??**):

$$\left| \frac{b(p,q)}{b} - \frac{2 \cdot j(p,q)}{N+1} \right| \leq \frac{12 \cdot j(p,q)}{N+1} \cdot N^{-3\alpha} \tag{17}$$

This gives:

$$\sum_{\mathcal{F}_k | q < q_m} \frac{b(p,q)}{b} = 1 - \sum_{\mathcal{F}_k | q \geq q_m} \frac{b(p,q)}{b} \leq 1 - (1 - 6 \cdot N^{-3\alpha}) \cdot \sum_{\mathcal{F}_k | q \geq q_m} \frac{2 \cdot j(p,q)}{N+1}$$

From (**??**) and using:

$$\sum_{\mathcal{F}_k} \frac{2 \cdot j(p,q)}{N+1} = 1$$

we obtain:

$$\sum_{\mathcal{F}_k | q < q_m} \frac{b(p,q)}{b} \leq 9 \cdot N^{-3\alpha} \tag{18}$$

From equation (**??**) and inequalities (**??**), (**??**) and (**??**), we obtain:

$$\delta_V \leq 9 \cdot N^{-3\alpha}$$

C Proof of theorem ??

The generator $\mathcal{G}$ combines the generators $\mathcal{V}$ and $\mathcal{G}(p,q)$. Moreover, $\mathcal{V}$ generates $p/q \in \mathcal{F}_k$ such that the statistical distance δ_G of the distribution induced by $\mathcal{G}(p,q)$ from the uniform distribution in $B(p,q)$ is at most $7 \cdot N^{-3\alpha}$. Therefore the statistical distance δ of $\mathcal{G}$ from the uniform distribution in B is at most:

$$\delta \leq \delta_V + \delta_G \leq 16 \cdot N^{-3\varepsilon/13}$$