SOME REMARKS ON WILLIAMS' PUBLIC-KEY CRYPTO FUNCTIONS

Siguna Müller
Department of Mathematics, Ross Hall 311, University of Wyoming
P.O. Box 3036, Laramie, WY 82071
email: smuller@uwyo.edu
(Submitted August 2004)

ABSTRACT

In 1984, H.C. Williams introduced a public key cryptosystem whose security is as intractable as factorization. That is, the system is provably as difficult to break as it is to find the factors of the modulus $n = pq$. By utilizing properties of the Lucas functions, this proposal is the only factorization equivalent scheme that is known which does not impose any restrictions on the primes used in the modulus.

However, Williams anticipates several restrictions on the messages without further analyzing if these are always fulfilled. By investigating simple numerical examples we found that any message not meeting these criteria cannot be encrypted and most likely directly exposes a factor of the modulus during the encryption process.

We analyze this problem encountered in the original scheme and establish the exact number of such 'dangerous' messages. Moreover, we provide a simple modification of the Williams' system which minimizes these difficulties. The modification does not complicate the system in any way. Evaluation of the proposed system can be obtained in exactly the same number of steps as in the original system. The results obtained will demonstrate that the possible lack of security due to the 'dangerous' messages is negligibly small for large moduli.

1. INTRODUCTION

1.1 Background

The security of many cryptographic techniques depends upon the intractability of the integer factorization problem. In spite of spectacular progress of recent years in developing fast factorization algorithms (cf. [5, 6, 10, 11]), an appropriately chosen, sufficiently large modulus still cannot be factorized by current techniques.

It is well-known that the RSA public-key cryptosystem can be broken if its modulus $n = pq$ can be factored into its secret keys p and q. However, it is not known if the opposite is true. This problem has led to the development of a variety of PKCSs [3, 4, 8, 16-20] whose security is *equivalent in difficulty to factoring* the modulus n, i.e., for which knowledge of the factorization of the modulus is necessary in order to retrieve plaintext from ciphertext without the use of the decryption key. As with Rabin's signature scheme [13], the proof of the equivalence to factoring of all these schemes is of a constructive nature and consequently can be converted into a chosen ciphertext attack (CCA). During the last few years, a great number of mechanisms have been developed for securing cryptoschemes against such types of attacks (see e.g. [1, 2, 9]).

All the above factorization equivalent techniques (except for [19]) require special forms of the underlying primes in the modulus. Although it is not known whether or not factorization of these special moduli is more easily tractable, schemes based on *the general factorization*

primitive certainly are to be preferred. The Williams scheme [19] utilizes certain properties of Lucas functions, thus making it the only cryptosystem equivalent to factoring that does not impose any restrictions on the primes.

1.2 Motivation

In spite of being the only system not requiring any special primes, it seems that it has never obtained a great deal of interest. Indeed, when a colleague of mine started implementing it, he boldly announced that it did not even work at all! His calculations showed that very often the required multiplicative inverses don't exist. As a consequence, not all messages can be encrypted, resp. decrypted, and moreover, this fact *allows to totally break the system.* Encryption of such 'dangerous' messages leads to an immediate factorization of the modulus into its secret keys. What had gone wrong? I implemented it myself and encountered the same 'problems'. After consulting the original paper [19], I learned that for the correctness of the scheme several restrictions on the message are necessary. Unfortunately Williams does not further analyze these restrictions, and moreover, from his paper it is not clear, if, or to what extend these impose a lack of security in the scheme. These problems encountered during the implementation were the motivation for writing this paper.

Outline of the paper: The goal of this paper is to firstly analyze Williams' encryption scheme [19]. We will determine the exact number of messages that cannot be encrypted and thereby directly expose a factor during encryption. Secondly, we will provide a simple modification of the scheme that minimizes this difficulty. In particular, when $p \equiv q \equiv 3 \bmod 4$, all factor-revealing messages are eliminated. The modified version will not only be shown to be as secure as the original, but also has the exact same complexity and performance as the original.

2. 'DANGEROUS' MESSAGES FOR THE WILLIAMS' SCHEME?

2.1 Some Preliminaries of the Williams' Scheme

Let $n = pq$ be the product of two distinct primes and let $c \in \mathbb{Z}_n^*$ be defined such that the Legendre symbols $\epsilon_p = \left(\frac{c}{p}\right)$ and $\epsilon_q = \left(\frac{c}{q}\right)$ satisfy

$$\epsilon_p \equiv -p \bmod 4, \quad \epsilon_q \equiv -q \bmod 4.$$

Additionally, a value s with $\gcd(s, n) = 1$ and $\left(\frac{s^2-c}{n}\right) = -1$ is determined. Further, let the public enciphering key e and the secret deciphering key d be chosen according to $ed \equiv \frac{m+1}{2} \bmod m$, where $m = \frac{(p-\epsilon_p)(q-\epsilon_q)}{4}$.

In the following, let $w \in \mathbb{Z}$ be the message to be encrypted. Let n, e, c, s constitute the *public key*, and p, q, m, d denote the *secret key.*

Let $b_1 = 1$, if $\left(\frac{w^2-c}{n}\right) = 1$, and $b_1 = -1$, if $\left(\frac{w^2-c}{n}\right) = -1$.

Define $\alpha(w)$ as $\frac{w+\sqrt{c}}{w-\sqrt{c}}$, respectively as $\frac{(w+\sqrt{c})(s+\sqrt{c})}{(w-\sqrt{c})(s-\sqrt{c})} \bmod n$, according as $b_1 = 1$ or -1.

Suppose $\gcd(w^2 - c, n) = 1$. Then, if $\alpha(w) \equiv a + b\sqrt{c} \bmod n$, it follows that

$$a = a(w) \equiv \begin{cases} \frac{w^2+c}{w^2-c} \bmod n, & \text{if } b_1 = 1, \\ \frac{(w^2+c)(s^2+c)+4csw}{(w^2-c)(s^2-c)} \bmod n, & \text{if } b_1 = -1, \end{cases}$$

$$b = b(w) \equiv \begin{cases} \frac{2w}{w^2-c} \bmod n, & \text{if } b_1 = 1, \\ \frac{2s(w^2+c)+2w(s^2+c)}{(w^2-c)(s^2-c)} \bmod n, & \text{if } b_1 = -1. \end{cases}$$

The en- and decryption functions are described by means of the sequences (cf. [14, 21]), $X_i(a) \equiv \frac{V_i(2a,1)}{2} \bmod n$, and $Y_i(a,b) \equiv bU_i(2a,1) \bmod n$, where U_i and V_i, denote the Lucas sequences of the first and second kind, respectively.

In essence, the **encryption** routine of [19] consists of calculating

$$E(w) \equiv \frac{X_e(a(w))}{Y_e(a(w), b(w))} \bmod n \tag{2.2}$$

(cf. also [15]). Further, **decryption** essentially consists of retrieving the parameters $a(w)$ and $b(w)$ from the cryptogram $[E(w), b_1, b_2]$, where $b_2 \in \{0, 1\}$ is defined by $a(w) \bmod 2$. Finally, w can be retrieved from $a(w)$ and $b(w)$ by

$$w = \begin{cases} \frac{a(w)+1}{b(w)} \bmod n, & \text{if } b_1 = 1, \\ \frac{cb(w)-s(a(w)+1)}{a(w)+1-sb(w)} \bmod n, & \text{if } b_1 = -1. \end{cases} \tag{2.3}$$

At this stage we leave out further details of the scheme as these are not necessary for our investigations. For a more specific background we refer to [19, 15] and section 4 below.

2.2 Insecure Encryption?:

Williams [19] showed that (for suitable messages) breaking the scheme is equivalent in difficulty to factoring. Nonetheless, this does not exclude the possibility that the scheme could be broken by some other means (e.g., the underlying routines frequently require the calculation of multiplicative inverses modulo n. In many cases, these won't exist). We investigated encryption and decryption for some sample moduli. The following table shows the distribution of examples of messages that cannot be encrypted, resp. decrypted, and that (for the former) very likely automatically expose the factorization of n.

			messages that cannot be encrypted		messages that cannot be decrypted	
p	q	n	number	percentage	number	percentage
13	17	221	43	19	29	13
41	43	1763	210	11	165	9
83	107	8881	663	7	569	6
151	191	28841	1195	4	1025	3

The above messages relating to the encryption process will in most cases automatically expose the secret key to any legitimate user. In that case the system will be broken without an attempted attack, but rather by simply running it during encryption.

In spite of the data provided so far, we show that these 'difficulties' actually represent no threat whatsoever under realistic settings of the scheme.

3. PARAMETER ANALYSIS OF THE WILLIAMS SCHEME

Without going into detail, we summarize the conditions of Theorem 3.1 of [19] that are essential for correct encryption and decryption. These are given by the following constrains: $\gcd(w^2-c,n)=1$, and $\gcd(w,n)=1$ for $b_1=1$, and $\gcd(w+s,n)=1$ and $\gcd(sw+c,n)=1$ for $b_1=-1$.

Clearly, by inspecting the encryption and decryption routines, these conditions can be specified as follows.

Lemma 1: *Let $w\in\mathbb{Z}_n^*$ be the message. Then the encryption procedure [19] can be carried out whenever both $\gcd(w^2-c,n)=1$ and $\gcd(b(w),n)=1$. Further, necessary conditions for the decryption function in [19] are $\gcd(b(w),n)=1$ for $b_1=1$, and $\gcd(a(w)+1-sb(w),n)=1$ for $b_1=-1$.*

We now establish the exact number of the messages for which encryption or decryption is not possible. Consider first the case that $\left(\frac{w^2-c}{n}\right)\neq 0$.

Lemma 2: *Suppose that $\gcd(w^2-c,n)=1$ for all messages w. Then the following conditions hold.*

1. $\#_{b_e}:=\#\{w\in\mathbb{Z},:\ \gcd(b(w),n)>1\}=3\frac{p+q}{2}-2.$
2. $\#_{b_d}:=\#\{w\in\mathbb{Z},:\ \gcd(b(w),n)>1\ \textit{and}\ b_1=1\}=\frac{p+q}{2}.$
3. $\#_{den}:=\#\{w\in\mathbb{Z},:\ \gcd(a(w)+1-sb(w),n)>1\ \textit{and}\ b_1=-1\}=\frac{p+q}{2}-1.$

Proof: We will only consider the first case. The other two can be treated in a similar way. Since $(a(w)+1)(a(w)-1)\equiv cb(w)^2 \bmod n$, it follows that $b(w)\equiv 0 \bmod p$ or q, exactly when $a(w)\pm 1\equiv 0 \bmod p$ or q. For $b_1=1$ this is the case for $\gcd(w,n)>1$. Similarly, for $b_1=-1$ we obtain $\gcd(ws+c,n)>1$ or $\gcd(w+s,n)>1$. The number of $w\in\mathbb{Z}_n$ with $\gcd(w,n)>1$ is $p+q-1$. This number can also be obtained in the other two cases. Now we have to distinguish the messages w with corresponding $b_1=1$ from those with $b_1=-1$. For a fixed ϵ with $|\epsilon|=1$, it is known (cf. [7]) that the number of $w\in\mathbb{Z}_p$ with $\left(\frac{w^2-c}{p}\right)=\epsilon$ is ± 1 plus the number of $w\in\mathbb{Z}_p$ with $\left(\frac{w^2-c}{p}\right)=-\epsilon$. Therefore, with an exception of one w, the case $b_1=1$ occurs as often as $b_1=-1$. Hence, by regarding the above three quantities for $b_1=1$ respectively $b_1=-1$, we obtain the desired number as $\frac{(p+q-1)-1}{2}+\frac{2(p+q-1)}{2}$ or $\frac{p+q-1}{2}+\frac{2(p+q-1)-1}{2}$, which yields the above assertion.

Corollary 1: *The number of messages w, $0\leq w<n$ with $\gcd(w^2-c,n)=1$ that cannot be encrypted by Williams' system equals $\#_{b_e}=3\frac{p+q}{2}-2$ and the number of those that cannot be decrypted, equals $p+q-1$.*

Those messages w will automatically factorize n unless one of the above $\gcd's$ in Lemma 2 is actually equal to n. In extending the above results also for the case that $\gcd(w^2-c,n)>1$, and considering only messages that expose a proper factor, one gets then the following general formula.

Theorem 1: *The number of messages* w, $0 \le w < n$ *that expose a proper factor when being encrypted by Williams' system equals*

$$\begin{cases} 3^{\frac{p+q}{2}} - 5, & \text{if } p \equiv q \equiv 1 \bmod 4, \\ 3^{\frac{3p+7q}{2}} - 5, & \text{if } p \equiv 1 \bmod 4,\ q \equiv 3 \bmod 4, \\ 7^{\frac{p+q}{2}} - 9, & \text{if } p \equiv q \equiv 3 \bmod 4. \end{cases}$$

Proof: Observe that $p \equiv -\left(\frac{c}{p}\right) \bmod 4$, and similarly for q. Then the first assertion immediately follows from the number $\#_{b_e}$ of Lemma 2 by subtracting the messages $w \equiv 0 \bmod n$, $w \equiv -c \bmod n$, and $w \equiv -\frac{c}{s} \bmod n$, for which $\gcd(b(w), n)$ is equal to n. The other two cases can be seen by additionally counting the number of messages in $\mathbb{Z}_n$ with $\gcd(w^2 - c, n) = p$ respectively q. Finally, for the third case the four messages w with $w^2 \equiv c \bmod n$ need to be excluded as $\gcd(w^2 - c, n) = n$ but not a proper factor.

Consequence: It is now obvious that the number of messages that will factorize n will become negligibly small for sufficiently large moduli. Consequently, for realistic settings, the Williams' scheme is resistant against the 'attacks' described in section 2.2.

4. THE MODIFIED SCHEME

Although the number of messages that either cannot be encrypted or decrypted in Williams scheme is very small whenever the primes p and q are sufficiently large, the scheme could be designed in a more advantageous way. We establish a very simple modification of the scheme which will minimize the number of the 'dangerous' messages. In detail, we show that the number $\#_{b_e}$ then does not need to be considered. Particularly, this very simple modification does make sense since it runs in the exact same number of steps as does the original.

4.1 The Modified Encryption and Decryption Routines

We present a simple modification of Williams' scheme that does not require the condition $\gcd(b(w), n) = 1$ in Lemma 1. Bearing in mind that this restriction for $b(w)$ was particularly caused by the calculation of the inverse of $Y_e(a(w), b(w) = b(w)U_e(2a(w), 1)$ modulo n in equation (2), this leads to the question about the zeros of $X_e(a(w)) = V_e(2a(w), 1)$ modulo p or q.

We note that the choice of the parameters $a = a(w)$ and $b = b(w)$ implies that $a^2 - cb^2 \equiv 1 \bmod n$.

Proposition 1: *Suppose that* $n = pq$ *is the product of two primes* p, q *and that* a, b, *and* c *are integers which satisfy* $a^2 - cb^2 \equiv 1 \bmod n$ *and* $p \equiv -\left(\frac{c}{p}\right) \bmod 4$. *Under these conditions* $X_e(2a) = V_e(2a, 1)$ *is always coprime to* n.

Proof: Let $\alpha, \overline{\alpha} = a \pm b\sqrt{c}$ be the distinct roots of $x^2 - Px + Q = x^2 - 2ax + 1 \equiv 0 \bmod n$ with discriminant $D = 4a^2 - 4 = 4cb^2$.

It is well know (cf. [14]) that $U_{\frac{p-(D/p)}{2}}(P, Q) \equiv U_{\frac{p-(c/p)}{2}}(2a, 1) \equiv 0 \bmod p$ iff $\left(\frac{Q}{p}\right) = 1$, which obviously is the case since $Q = 1$. Moreover, the smallest k with $U_k(P, Q) \equiv 0 \bmod p$ then must be a divisor of $\frac{p-\left(\frac{D}{p}\right)}{2}$. Therefore, $k \mid \frac{p-(c/p)}{2} \bmod p$. Hence, by the hypothesis on c, k must be odd.

On the other hand, it is well known that [21] that there exists an integer e with $V_e(P,Q) \equiv 0 \bmod p$ if and only if k is even, which is a contradiction.

Remark 1: A similar statement has been shown in [19]. The result of Proposition 1 is extending Williams' in allowing general values $b = b(w) \in \mathbb{Z}_n$ instead of $b = b(w) \in \mathbb{Z}_n^*$.

By the choice of the parameters $a = a(w)$, $b = b(w)$, and c, the expression $Y_e(a,b)$ will therefore always be relatively prime to n. It thus makes sense to modify William's encryption process as follows.

Encryption: If w is the message then let

$$E(w) = \frac{Y_e(a(w), b(w))}{X_e(a(w))} \bmod n.$$

The *cryptotext* to be transmitted is the triple $C = [E(w), b_1, b_2]$, where b_2 equals 0 or 1, depending on whether $a(w)$ is even or odd.

Decryption: Upon receiving C, the receiver first calculates the values $a_0 \equiv \frac{1+E(w)^2c}{1-E(w)^2c} \bmod n$, and $b_0 \equiv \frac{2E(w)}{1-E(w)^2c} \bmod n$.

The second step consists of determining $\sigma = \sigma(w) = (-1)^{b_2 - X_d(a_0)}$ and $a(w), b(w)$ by $a(w) \equiv \sigma X_d(a_0)$ and $b(w) \equiv \sigma Y_d(a_0, b_0) \bmod n$.

The message w can be retrieved from $a(w)$ and $b(w)$ by means of

$$w = \begin{cases} \frac{a(w)+1}{b(w)} \bmod n, & \text{if } b_1 = 1, \\ \frac{cb(w)-s(a(w)+1)}{a(w)+1-sb(w)} \bmod n, & \text{if } b_1 = -1. \end{cases} \tag{4.1}$$

provided $\gcd(b(w), n) = 1$ for $b_1 = 1$, and $\gcd(a(w) + 1 - sb(w), n) = 1$ for $b_1 = -1$.

Lemma 3: *The encryption scheme of this section is successful for all messages $w \in \mathbb{Z}_n$ when $p \equiv q \equiv 1 \bmod 4$, and to all w that satisfy $\gcd(w^2 - c, n) = 1$ when $p \equiv 3 \bmod 4$ for some $p|n$.*

4.2 Proof of Correctness

The proof runs along the same lines as the one in [19]. We summarize Williams' underlying ideas and then extend them for our modification. As with the original scheme, the basic property is based on the congruence

$$\alpha(w)^{2ed} \equiv \pm\alpha(w) \bmod n, \tag{4.2}$$

where $\alpha(w) = a + b\sqrt{c}$, and $a = a(w), b = b(w) \in \mathbb{Z}$.

For general a and b this congruence does not always hold modulo n. However, if $a = a(w)$ and $b = b(w)$ are chosen according to the above formulas, then Williams has shown that, if $b_1 = 1$, then $\alpha(w)^{\frac{p-\epsilon(p)}{2}} \equiv \left(\frac{w^2-c}{p}\right) \bmod p$, $\alpha(w)^{\frac{q-\epsilon(q)}{2}} \equiv \left(\frac{w^2-c}{q}\right) \bmod q$. Similarly, if $b_1 = -1$, then it follows that $\alpha(w)^{\frac{p-\epsilon(p)}{2}} \equiv \left(\frac{w^2-c}{p}\right)\left(\frac{s^2-c}{p}\right) \bmod p$, $\alpha(w)^{\frac{q-\epsilon(q)}{2}} \equiv \left(\frac{w^2-c}{q}\right)\left(\frac{s^2-c}{q}\right) \bmod q$. If additionally $\left(\frac{w^2-c}{n}\right) = 1$, respectively -1 according to the value of b_1, then the choice of e and d implies the desired result $\alpha(w)^{2ed} \equiv \pm\alpha(w) \bmod n$.

By means of this $\alpha(w)$ one then has the well-known representation of the Lucas sequences, and consequently, of the underlying functions,

$$X_i(a(w)) = X_i(\alpha(w)) = \frac{V_i(2a(w),1)}{2} = \frac{\alpha(w)^i + \overline{\alpha}(w)^i}{2}, \tag{4.3}$$

$$Y_i(a(w),b(w)) = Y_i(\alpha(w)) = b(w)U_i(2a(w),1) = b(w)\frac{\alpha(w)^i - \overline{\alpha}(w)^i}{\alpha(w) - \overline{\alpha}(w)}. \tag{4.4}$$

As is well-known, the powers of $\alpha = \alpha(w)$ and $\overline{\alpha} = \overline{\alpha}(w)$ correspond to the terms of the X_i- and Y_i- functions by the characterization

$$\alpha^i = X_i(\alpha) + Y_i(\alpha)\sqrt{c} \text{ and } \overline{\alpha}^i = X_i(\alpha) - Y_i(\alpha)\sqrt{c}.$$

Moreover, by basic properties of the Lucas sequences we have $X_{2ed}(a(w)) \equiv X_d(a_0,b_0)$, and $Y_{2ed}(a(w),b(w)) \equiv Y_d(a_0,b_0) \bmod n$, where a_0, b_0 are defined by $X_{2e}(a(w)) \bmod n$ and $Y_{2e}(a(w),b(w)) \bmod n$, respectively.

Hence, by (5), $X_d(a_0) \equiv \sigma a(w) \bmod n$, $Y_d(a_0,b_0) \equiv \sigma b(w) \bmod n$.

Consequently, the decrypter needs to find σ, $a(w)$, and $b(w)$, or, alternatively, σ, a_0, and b_0, since he knows d.

As in the original scheme [19], $\sigma = (-1)^{b_2 - X_d(a_0)}$. Moreover, since $X_e^2 - Y_e^2 c \equiv 1 \bmod n$ by a fundamental property of the Lucas sequences, we have

$$\alpha(w)^{2e} \equiv \frac{X_e(a(w)) + Y_e(a(w),b(w)\sqrt{c}}{X_e(a(w)) - Y_e(a(w),b(w)\sqrt{c}} \equiv \frac{1 + E(w)\sqrt{c}}{1 - E(w)\sqrt{c}} \equiv$$

$$\equiv \frac{1 + E(w)^2 c}{1 - E(w)^2 c} + \frac{2E(w)}{1 - E(w)^2 c}\sqrt{c} \equiv a_0 + b_0\sqrt{c} \bmod n. \tag{4.5}$$

Now, as $\alpha(w)^{2e} \equiv X_{2e}(\alpha(w)) + Y_{2e}(\alpha(w))\sqrt{c} \bmod n$, the receiver thus can calculate a_0, b_0 from the cryptogram,

$$a_0 \equiv \frac{1 + E(w)^2 c}{1 - E(w)^2 c} \bmod n \text{ and } b_0 \equiv \frac{2E(w)}{1 - E(w)^2 c} \bmod n,$$

as desired.

Then obviously, $a(w) \equiv \sigma X_d(a_0)$ and $b(w) \equiv \sigma Y_d(a_0,b_0) \bmod n$ and the message w can be retrieved from $a(w)$ and $b(w)$ by means of (3), provided $\gcd(b(w),n) = 1$ for $b_1 = 1$, and $\gcd(a(w) + 1 - sb(w), n) = 1$ for $b_1 = -1$.

4.3 Special Decryption

Actually, it is not necessary to impose Williams' constrains on $a(w)$ and $b(w)$ in the decryption process. Indeed, if the above conditions are violated, we can find an alternative way to retrieve w.

Lemma 4: *Let w be the message, $[E(w), b_1, b_2]$ be the cryptogram that is transmitted and $a(w)$, $b(w)$ be the values obtained in the decryption algorithm.*

(a) If $\gcd(b(w), n) = p$ *for some prime factor p of n and* $b_1 = 1$, *then* $w \equiv \frac{cpt}{a(w)-1} \bmod n$, *where* $t \in \mathbb{Z}$ *is defined via* $pt = b(w)$.

(b) If $\gcd(a(w) + 1 - sb(w), n) = p$ *for* $b_1 = -1$, *then* $w \equiv c\frac{-a(w)+1+sb(w)}{s(a(w)-1)-cb(w)} \bmod n$.

Proof: The first case follows from the definition of $b(w)$, which implies $w^2 - c = \frac{2w}{pt}$ and $a(w) - 1 \equiv \frac{2c}{w^2-c} \bmod n$. By combining those two equations we obtain the desired solution for w. Similarly the second assertion follows from the characterization of the quantities $a(w)$ and $b(w)$.

Corollary 2: *The decryption scheme of this section is successful for all messages $w \in \mathbb{Z}_n$.*

Proof: It only remains to consider the cases $b(w) \equiv 0 \bmod n$ for $b_1 = 1$ respectively $\gcd(a(w) + 1 - sb(w), n) = n$ for $b_1 = -1$. The former case is equivalent to $w \equiv 0 \bmod n$, whereas the latter is equivalent to $w \equiv -\frac{c}{w} \bmod n$.

4.4 Equivalence of Decryption and Factoring

This equivalence can be proven analogously as in [19]. For completeness we give the proof for our scheme, since it is a modification of the original one.

By the underlying properties of the encryption and decryption routines the decryption of a message cannot uniquely be obtained. Similarly as in the Rabin scheme this unambiguous decryption is the basis for establishing the equivalence to factoring. Under 'correct' decryption the original message will be obtained, while under a simple modification of this procedure the decryption routine evaluates a proper factor of n.

Assume that a cryptanalyst may be able to decrypt a certain fraction of all ciphertexts. Then the knowledge of a decryption algorithm immediately leads to the factorization of n. The procedure is the same as in the original scheme since it only relies on the values $\alpha^{2ed} \bmod p$ and $\alpha^{2ed} \bmod q$.

Suppose that p, q, n and c are chosen as above. The equivalence to factoring can be shown as follows. First, a number w with $\left(\frac{w^2-c}{n}\right) = -1$ is chosen. According to Lemma 4.3 of [19] there exist two and only two values of z modulo n, such that

$$\frac{w^2 + c}{w} \equiv \frac{z^2 + c}{w} \bmod n \text{ and } \left(\frac{z^2 - c}{n}\right) = 1.$$

Additionally, any of these values of z imply $\gcd(w - z, n) = p$ or q.

Notice that $\frac{w^2+c}{w} \equiv \frac{z^2+c}{w} \bmod n$ is equivalent to

$$\frac{a(w)}{b(w)} \equiv \frac{a(z)}{b(z)} \bmod n \text{ for } b_1 = 1. \tag{4.6}$$

It can immediately be verified that any $z \not\equiv w \bmod n$ that fulfills the latter condition satisfies $\left(\frac{z^2-c}{n}\right) = 1$ and therefore $\gcd(w - z, n) = p$ or q.

This idea now can be applied to obtain the factorization algorithm in the following way. By making use of property (2.12) of [19], which is

$$X_i(a_1)Y_i(a_1, b_1) \equiv X_i(a_2)Y_i(a_2, b_2) \bmod n,$$

whenever $\frac{a_1}{b_1} \equiv \frac{a_2}{b_2} \bmod n$, equation (9) yields for $i = e$,

$$X_e^{-1}(\frac{w^2+c}{w^2-w})Y_e(\frac{w^2+c}{w^2-w}, \frac{2w}{w^2-c}) \equiv$$

$$\equiv X_e^{-1}(\frac{z^2+c}{z^2-z})Y_e(\frac{z^2+z}{z^2-z}, \frac{2z}{z^2-c}) \equiv E(z) \bmod n. \tag{4.7}$$

Observe that the left hand side of the equation can be calculated by following the above encryption algorithm for $\mathbf{b_1{=}1}$ (rather the correct $b_1 = -1$) on purpose. Let $[E(w), 1, b_2]$ be the resulting cryptogram. By assumption, the cryptanalyst can apply the decryption algorithm to find the corresponding plaintext. However, the decoded cryptotext cannot be the same as w, because w was in the encryption process assigned the wrong values for $a(w)$ and $b(w)$ by the choice of $b_1 = 1$ instead of -1. Indeed, it follows from (10) that the decryption oracle returns the decryption of $E(z) \bmod n$. Since $\left(\frac{z^2-c}{n}\right) = 1$, which is the correct b_1 w.r.t. z, this decryption equals z. So the original cryptotext w will be decoded to the different value z and $\gcd(w - z, n)$ is a proper factor of n, as claimed.

5. SUMMARY

Motivated by some numerical examples of messages that either cannot be encrypted or decrypted we gave a short cryptanalysis of Williams' provable secure cryptoscheme [19] and developed the exact number of these messages. It is shown that this number becomes negligibly small for suitably large moduli. Moreover, we presented a modified version for which the number of such messages is minimized. Indeed, if $p \equiv q \equiv 1 \bmod 4$ then the modified version will never expose any factor during encryption and therefore the cryptogram of *any* message can only be obtained without the secret key by factoring the modulus. For general primes p and q of the modulus n, only the messages w with $\gcd(w^2 - c, n) > 1$ need to be excluded. Consequently, with overwhelming probability, no message will expose a factor of n during encryption. E.g., if $n = pq$ where p and q each have about 50 digits, the probability would only be less than approximately 10^{-49}.

Recently, factorization equivalent RSA modifications have become of great interest as basis for provably secure encryption schemes against chosen ciphertext attacks. While the development of the Williams scheme was originally mainly of theoretical interest, it actually presents an important basis for such security enhanced methods. The algorithms could be efficiently realized by means of rapid evaluation methods of combined Lucas sequences $U(P, Q)$, $V(P, Q)$. This can be achieved in roughly twice the time required for exponentiation, since $Q = 1$, which allows the most optimal evaluation [12, 21]. Although the algorithms are more involved than plain RSA, the Williams' scheme is therefore still quite practical. As of today, Williams system (and the proposed simple modification) represent the only factorization equivalent schemes which utilize the general factorization problem of any large number $n = pq$. Contrary to the other proposals they do not require any special form of the primes.

ACKNOWLEDGMENTS

Research supported under APART [Austrian Programme for Advanced Research and Technology] by the Austrian Academy of Sciences.

REFERENCES

[1] R. Cramer, V. Shoup. "Design and Analysis of Practical Public-key Encryption Schemes Secure Against Adaptive Chosen Ciphertext Attack." *SIAM J. Comput.* **33.1** (2003): 167–226 (electronic).

[2] R. Cramer, V. Shoup. "Universal Hash Proofs and a Paradigm for Adaptive Chosen Ciphertext Secure Public-key Encryption." Advances in cryptology—*EUROCRYPT 2002* (Amsterdam), 45–64, Lecture Notes in Comput. Sci., 2332, Springer, 2002.

[3] J. H. Loxton, D. D. Khoo, G. J. Bird, J. Seberry. "A Cubic RSA Code Equivalent to Factorization." *Journal of Cryptology* **5** (1992): 139-150.

[4] K. Kurosawa, T. Ito, M. Takeuchi. "Public Key Cryptosystem Using a Reciprocal Number With the Same Intractability as Factoring a Large Number." *Cryptologia* **12** (1988): 225-233.

[5] A. K. Lenstra, H. W. Lenstra Jr. "The Development of the Number Field Sieve". *Lect. Notes Math.,* Number 1554, Springer, Berlin, 1993.

[6] A. K. Lenstra, H. W. Lenstra Jr., M. S. Manasse, J. M. Pollard. "The Number Field Sieve". *Lect. Notes Math.,* Number 1554, (1993): 11-42.

[7] S. Müller. "Carmichael Numbers and Lucas Tests." *Contemporary Mathematics* **225**, Proc. of the Fourth International Conference on Finite Fields and Applications, Mullin, R.C., Mullen, G.L. (eds.) Amer. Math. Soc. (1999), 193-202.

[8] S. Müller, W. B. Müller. "The Security of Public Key Cryptosystems Based on Integer Factorization." *LNCS.* 1438, Proc. of ACISP'98, Springer-Verlag (1998), 7-23.

[9] D. Pointcheval. "Chosen-ciphertext Security for Any One-way Cryptosystem". *Public key cryptography.* (Melbourne, 2000), 129–146, LNCS 1751, Springer, Berlin, 2000.

[10] C. Pomerance. "The Quadratic Sieve Factoring Algorithm." *Advances in Cryptology – Eurocrypt'84, Lecture Notes in Computer Science* **209** (1985): 169-182.

[11] C. Pomerance. "The Number Field Sieve." *Proceedings of Symposia in Applied Mathematics* **48** (1994): 465-480.

[12] H. Postl. "Fast Evaluation of Dickson Polynomials". *Contributions to General Algebra*, **6**, 223-225. B. G. Teubner: Stuttgart 1988.

[13] M.O. Rabin. "Digitalized Signatures and Public-key Functions as Intractable as Factorization." MIT/LCS/TR-212, MIT Laboratory for Computer Science, 1979.

[14] P. Ribenboim. *The Book of Prime Number Records*, Springer, Berlin, 1988.

[15] A. Salomaa. *Public Key Cryptography*, Springer, Berlin, 1990.

[16] R. Scheidler. "A Public-key Cryptosystem Using Purely Cubic Fields". *J. Cryptology* **11** (1998): 109-124.

[17] R. Scheidler, H. C. Williams. "A Public-key Cryptosystem Utilizing Cyclotomic Fields". *Designs, Codes and Cryptography* **6** (1995): 117-131.

[18] H. C. Williams. "A Modification of the RSA Public-key Encryption Procedure". *IEEE Trans. Inf. Theory*, Vol. IT-26 **6** (1980): 726-729.

[19] H. C. Williams. "Some Public-key Crypto-functions as Intractable as Factorization". *Cryptologia* **9** (1985): 223-237.

[20] H. C. Williams. "An M^3 Public-key Encryption Scheme". *Advances in Cryptology - Crypto'85, Proceedings,* Springer, Berlin (1986), 358-368.
[21] H. C. Williams. "Edouard Lucas and Primality Testing". *Canadian Mathematical Society Series of Monographs and Advanced Texts,* Vol. 22, John Wiley & Sons, 1998.

AMS Classification Numbers: 11Y16, 11B50

✠ ✠ ✠