Known-plaintext attack to two cryptosystems based on the BB equation

G. Alvarez, L. Hernández Encinas, and J. Muñoz Masquè

Abstract—Recently, Rama Murthy and Swamy proposed a symmetric cryptosystem based on the Brahmagupta-Bhāskara (BB) equation. The BB equation is the quadratic Diophantine equation \( nx^2 + k = y^2 \), where \( k \) is an integer and \( n \) is a positive integer such that \( \sqrt{n} \) is irrational. For the particular case \( k = 1 \), the equation is called the Pell equation. The proposed cryptosystem was modified later by the same authors in order to avoid the cryptanalysis given by Youssef. Below, a known-plaintext attack to both cryptosystems is presented.

I. PRELIMINARIES AND NOTATIONS

The Brahmagupta-Bhāskara (BB) equation is the quadratic Diophantine equation

\[ nx^2 + k = y^2, \]

where \( k \) is an integer and \( n \) is a positive integer such that \( \sqrt{n} \) is irrational. For the particular case \( k = 1 \), the equation is called the Pell equation.

The goal of this paper is to show that the two cryptosystems proposed by Rama Murthy and Swamy in [2], [3] are vulnerable to the known-plaintext attack and hence, they are insecure.

First of all, we summarize the cryptosystem proposed in [2], which was subsequently improved in [3] in order to avoid the cryptanalysis appeared in [4]. The authors proposed a symmetric cryptosystem based on the Pell equation in the finite field \( GF(p) \), so that the Pell equation can be written as

\[ nX + 1 \equiv Y \pmod{p}, \quad (1) \]

where \( p \) is a prime number, and \( X \) and \( Y \) are quadratic residues modulo \( p \). In the proposed symmetrical cryptosystem there are two classes of keys: The primary key is a prime number \( p \), with \( t = \lceil \log p \rceil \) the bitlength of \( p \); the secondary key is a couple of prime numbers, \( a \) and \( b \) in the interval \( (2^t, p) \). The plaintext corresponding to a block of the original message of \( t \) bits, \( B \), is the integer \( n \) defined as \( n = (B)_{10} + 1 < p \), where \( (B)_{10} \) is the decimal expression of the block \( B \). The ciphertext, with bitlength \( 2(t + 1) \), corresponding to \( n \) is the pair of integers, \((s, d)\), given by

\[
\begin{align*}
s &\equiv aX + bY \pmod{p}, \\
d &\equiv aX - bY \pmod{p}.
\end{align*}
\]

The deciphering process to obtain \( n \) from \((s, d)\) takes into account the following expressions

\[
\begin{align*}
X &\equiv (2a)^{-1}(s + d) \pmod{p}, \\
Y &\equiv (2b)^{-1}(s - d) \pmod{p}, \\
n &\equiv X^{-1}(Y - 1) \pmod{p}.
\end{align*}
\]

Once the cryptanalyst has obtained the values \( a^{-1} = \frac{a_1}{a_1a_2}, b^{-1} = \frac{b_1}{b_1b_2} \), \( X_1, X_2 \in \mathbb{Q} \), a similar system is solved by using the equations derived from \( n_1 \) or \( n_2 \) and a new pair of plaintext-ciphertext: \((n_3, (s_3, d_3))\). Then new rational values for \( a^{-1} = \frac{a_3}{a_2a_3}, b^{-1} = \frac{b_3}{b_2b_3} \) are obtained, and from then the value of

\[
\gcd(a_1a_2 - a_1a_2, b_1b_2 - b_1b_2) = p'
\]

provides a multiple of the prime \( p \).

In order to improve the previous cryptosystem, Rama Murthy and Swamy proposed in [3] a modification in the ciphering process. With the same notations as above, the ciphertext \((s, d)\), corresponding to a given plaintext \( n \), is computed as follows:

\[
\begin{align*}
s &\equiv aX + bY \pmod{p} \\
d &\equiv bX - aY \pmod{p}
\end{align*}
\]

With such a modification, the attack proposed by Youssef is
not effective as the linear system of equations (3) becomes
\[
\begin{pmatrix}
-1 & 1 & (a+b) - n_1(a-b) \\
-1 & 1 & 0 & (a+b) - n_2(a-b) \\
1 & 1 & (a-b) + n_1(a+b) & 0 \\
1 & 1 & 0 & (a-b) - n_2(a+b)
\end{pmatrix}
\begin{pmatrix}
a \\
b \\
X_1 \\
X_2
\end{pmatrix}
= \begin{pmatrix}
s_1 + d_1 \\
s_2 + d_2 \\
1 - s_1 - d_1 \\
1 - s_2 - d_2
\end{pmatrix},
\]
which cannot be solved in the unknowns \(a, b, X_1,\) and \(X_2,\)
when \(s_1 + d_1, s_2 + d_2, s_1 - d_1, s_2 - d_2, n_1,\) and \(n_2\) are known.

II. CRYPTOANALYSIS

In this section we present a cryptanalysis to the cryptosystem introduced in [3] by using a known-plaintext attack. Suppose that four distinct plaintext-ciphertext pairs \((n_i, (s_i, d_i)), i = 1, \ldots, 4,\) are known.

The cryptanalysis is based on the fact that the system of equations
\[
s_1 = aX_1 + b(n_1X_1 + 1) \\
d_1 = bX_1 - a(n_1X_1 + 1) \\
s_2 = aX_2 + b(n_2X_2 + 1) \\
d_2 = bX_2 - a(n_2X_2 + 1) \\
s_3 = aX_3 + bY_3 \\
d_3 = bX_3 - aY_3,
\]
(5)
obtained from (1) and (4), admits a unique solution “in \(\mathbb{Q}\)” whenever the following inequalities hold:
\[
D_1 = (s_1n_1 - s_2n_2 + d_1 - d_2)^2 + (d_1n_1 - d_2n_2 - s_1 + s_2)^2 \neq 0, \\
D_2 = (s_1n_2 + 1)(s_1d_2 - s_2d_1 + (n_2 - n_1)(s_1s_2 + d_1d_2)) \neq 0.
\]
In fact, as a simple computation shows, we have
\[
a = \frac{n_1d_1 - n_2d_2 + s_2 - s_1}{D_1}, \\
b = \frac{n_1s_1 - n_2s_2 + d_1 - d_2}{D_1}, \\
X_1 = \frac{n_1s_1 - n_2s_2 + d_1 - d_2}{D_2}, \\
X_2 = \frac{n_1s_1 - n_2s_2 + d_1 - d_2}{D_2}, \\
X_3 = \frac{n_1s_1 - n_2s_2 + d_1 - d_2}{D_2},
\]
(6)
(7)
(8)
(9)
(10)
Notice that the running time for solving the system (5) is negligible as explicit formulas are used. Once \(a, b, X_1, X_2, X_3,\) and \(Y_3\) are known, \(X_4\) and \(Y_4\) can be directly obtained as follows:
\[
X_4 = \frac{as_4 + bd_4}{a^2 + b^2}, \\
Y_4 = \frac{bs_4 - ad_4}{a^2 + b^2}.
\]
(12)
(13)
As the equations \(Y_i = n_iX_i + 1\) for \(i = 3, 4,\) have not been used in the system (5), they allow one to compute \(q = \text{gcd}(N_i, N_4),\) where \(N_i\) is the numerator of the fraction \(n_iX_i + 1 - Y_i\) for \(i = 3, 4,\) and \(q\) is a multiple of \(p,\) as follows from the very definition of the BB equation modulo \(p.\) Furthermore, the bitlength of \(N_i\) is less than \(4(t + 1),\) as \(N_i\) is the sum of three summands each of which being, at most, a product of two factors of bitlength \(\leq 2(t + 1),\) and other summands of less bitlength, as it is showed in equations (14) and (15).

\[
N_3 = (n_1n_2 + 1)(-s_1d_2 + s_2d_1) + (n_1n_3 + 1)(s_2d_3 - s_3d_2) + (n_1n_3 - s_1d_3 - s_3d_1) + n_1d_3 + s_3d_1, \\
N_4 = (n_1n_2 + 1)(-s_1d_2 + s_2d_1) + (n_1n_4 + 1)(-s_1d_4 - s_4d_1) + (n_1 - n_2)(s_2d_4 + s_4d_2) + n_1d_4 - s_4d_2 \\
+ (n_1 - n_2)(s_2d_3 + s_3d_2) + (n_1 - n_2)(s_2d_4 + s_4d_2) + (n_1 - n_2)(s_3d_4 - s_4d_3) + (n_1 + n_2)(s_3d_4 - s_4d_3).
\]
(14)
(15)
Hence the value of \(p\) is derived from \(q\) directly, without using any sophisticated factorization algorithm. Once the value of \(p\) is known, the values of \(a\) and \(b\) are obtained by reducing modulo \(p\) the rational values for \(a\) and \(b\) in (6) and (7).

Next, we show how to perform the cryptanalysis either when \(D_1\) or \(D_2\) vanishes. If \(D_1 = 0,\) then \(n_1\) and \(n_2\) are completely determined, namely,
\[
n_1 = \frac{s_2(s_2 - s_1) + d_2(d_2 - d_1)}{s_1d_2 - s_2d_1}, \\
n_2 = \frac{s_1(s_2 - s_1) + d_1(d_2 - d_1)}{s_1d_2 - s_2d_1}.
\]
(16)
(17)
Hence, it suffices to change one of the two plaintexts and to keep the other.

Moreover, the expressions \(n_2(s_2d_1 - s_1d_2) + s_1s_2 + d_1d_2\) and \((n_1(s_2d_1 - s_1d_2) - s_1s_2 - d_1d_2)\) cannot vanish simultaneously as the vanishing of both expressions imply \(n_1 = n_2,\) which contradicts the assumption. Consequently, if \(D_2 = 0,\) then either
\[
n_1 = \frac{s_1d_2 - s_2d_1 - n_2(s_1s_2 - d_1d_2)}{n_2(s_2d_1 - s_1d_2) + s_1s_2 + d_1d_2}, \\
n_2 = \frac{s_1d_2 - s_2d_1 - n_1(s_1s_2 - d_1d_2)}{n_1(s_2d_1 - s_1d_2) - s_1s_2 - d_1d_2}
\]
or

and we proceed as in the first case.

The cryptanalysis to the original cryptosystem as formulated in [2] is similar to the previous one, by using an analogous system to (5), which provides the following expressions:
\[
a = \frac{1}{2} E_1, \\
b = \frac{1}{2} n_1(s_1 - d_1) - n_2(s_2 - d_2), \\
X_1 = \frac{(s_1 + d_1)E_2}{E_1}, \\
X_2 = \frac{(s_2 + d_2)E_2}{E_1}, \\
X_3 = \frac{(s_3 + d_3)E_3 - n_1(s_1 - d_1) - n_2(s_2 - d_2)}{E_1}, \\
Y_3 = \frac{(s_3 + d_3)E_3 - n_1(s_1 - d_1) - n_2(s_2 - d_2)}{E_1},
\]
where
\[
E_1 = n_1(s_1 + d_1)(s_2 - d_2) - n_2(s_1 - d_1)(s_2 + d_2) \neq 0, \\
E_2 = (s_1 - d_1) - (s_2 - d_2) \neq 0.
\]

III. NUMERICAL EXAMPLES

A numerical example with artificially small values is presented below.
Suppose the secret keys are \( p = 88039103041, a = 82444068284, b = 74691386336 \), and let us consider the following plaintext-ciphertext pairs:

\[
n_1 = 62889964901, (s_1 = 23448928782, d_1 = 20575584876),
n_2 = 46618259218, (s_2 = 3834180954, d_2 = 17146237210),
n_3 = 58426687172, (s_3 = 20527763839, d_3 = 26549742294),
n_4 = 34101981540, (s_4 = 63745825525, d_4 = 34573143446).
\]

Here, \( t = 36 \). The values of the expressions (6)–(13) are

\[
a = \frac{18747562383465695290983921273199736421133866424687134364991044}{769682108971496471714488330965363014221},
b = -\frac{4911602345506800038831341441408261103987631847405089117373658}{769682108971496471714488330965363014221}.
\]

\[
X_1 = \frac{10043825931118451280958295880}{631655860252748444456115505020004901984559},
X_2 = \frac{9478379379122666841732575300735297683850}{40421787962812864175034800916964},
X_3 = \frac{157913956063187111140288672550122549613975}{3073642730541927516794719785387},
X_4 = \frac{9478379379122666841732575300735297683850}{6636253689492581216074541587646},
Y_1 = \frac{473411851895651333420866287650637648841925}{3328053506208383143782735065227},
p = \gcd(N_3, N_4) = 88039103041.
\]

By using the software package MAPLE over a Intel-Pentium D PC, CPU 3.20 GHz, 2.00 GB RAM, under Windows XP (SP 2), the running time to compute these values, was 0.062 seconds.

For a prime number with bitlength \( t = 64 \), say

\[
p = 1901227095 9876747001,
\]

\[
a = 1871225235 8876810813,
\]

\[
b = 17146237210,
\]

the running time to compute the values of \( a, b, X_1, X_2, X_3, Y_3, X_4, Y_4 \), and \( \gcd(N_3, N_4) = 3p \), was 0.065 seconds.

For a larger prime of 128 bits

\[
p = 3964446899 6042275986 5633386972 192728973,
\]

the running time to break the cryptosystem was 0.078 seconds. For the secret primary key of 192 bits

\[
p = 1173414387 5675490285 0423315885 2300218332
\]

\[
0750233346 767633509,
\]

and secondary keys

\[
a = 6766631271 3194092704 6048021365 0943954694
\]

\[
3733334474 302558181,
\]

\[
b = 7684478404 6912415337 235671884 2908013378
\]

\[
1976475131 09338017,
\]

the running time to break the cryptosystem was 0.094 seconds. In this case, we obtained \( \gcd(N_3, N_4) = 11p \).

### IV. Concluding Remarks

1. As shown above, it usually suffices to consider four distinct plaintext-ciphertext pairs to recover the primary and the secondary keys of the both proposed symmetric cryptosystems in a computationally feasible way.

2. Although the values \( X, Y \) in \( nX + 1 \equiv Y \pmod{p} \) are assumed to be quadratic residues in \([2, 3]\), which is justified by the authors (personal communication) in order to avoid vulnerabilities, this property seems to play no role in the cryptanalysis proposed.

3. The cryptanalysis proposed does not involve general factorization algorithms, but only to compute the greatest common divisor of several pairs of integers and few factors much smaller than the modulus \( p \) in these greatest common divisors. As is well known (e.g. see \([1, 2]\)), the \( \gcd(u, v) \) can be computed efficiently in polynomial time, \( O((\log(\max(u, v)))^2) \), by using the Extended Euclidean Algorithm.

### Acknowledgment

Authors thank anonymous reviewers for their valuable comments and suggestions.

### References


---

**G. Álvarez** obtained his Ph.D. in Computer Science from the Polytechnic University of Madrid (Spain), in 2000. He is a researcher at the Department of Information Processing and Coding, Spanish Council for Scientific Research (CSIC). His current research interest include cryptography, Internet security, and chaotic systems.

**Luis Hernández Encinas** obtained his Ph.D. in Mathematics from the University of Salamanca (Spain), in 1992. He is a researcher at the Department of Information Processing and Coding, Spanish Council for Scientific Research (CSIC). His current research interest include cryptography, hyperelliptic curve cryptosystems, image processing, data transmission and dynamical systems.
Jaime Muñoz Masqué obtained his Ph.D. in Mathematics from the University of Salamanca (Spain), in 1983. He is a researcher at the Department of Information Processing and Coding, Spanish Council for Scientific Research (CSIC). His current research interest include cryptography, hyperelliptic curve cryptosystems, and differential geometry.