

An IND-CPA cryptosystem from Demytko's primitive

David Galindo, Sebastià Martín,
Paz Morillo and Jorge L. Villar
Dep. Matemàtica Aplicada IV.
Universitat Politècnica de Catalunya
{dgalindo,sebas,m.paz,jvillar}@mat.upc.es

Abstract —

We propose an elliptic curve scheme over the ring \mathbb{Z}_{n^2} , which is efficient and semantically secure in the standard model. It is based on factoring, and it has expansion factor 2 (previous schemes with these features present expansion factors greater or equal than 4). Demytko's primitive has been used to obtain efficiency and probabilistic encryption. Semantic security of this scheme is based on a new decisional assumption, namely, the Decisional Small Root Assumption. Confidence on this assumption is also discussed. Keywords: public-key cryptography, semantic security, expansion factor, elliptic curves, Demytko's scheme.

I. INTRODUCTION

In 1984, Goldwasser and Micali [10] defined a new security notion that any encryption scheme should satisfy namely, indistinguishability of encryptions or semantic security against chosen plaintext attack (IND-CPA), and they proposed a scheme with this property. This notion informally says that a ciphertext does not leak any useful information about the plaintext, except its length, to a passive polynomial-time attacker. Nowadays, security models deal with active adversaries, so the standard security requirement for general purpose cryptosystems is indistinguishability of encryptions against chosen ciphertext attack (IND-CCA) [17].

A great part of existing practical IND-CCA schemes rely on the Random Oracle Model (ROM). Only Cramer and Shoup cryptosystem [4] and its variants [5] are known to be IND-CCA in the standard model, that is, based only on number theoretical assumptions. These assumptions come from existing IND-CPA schemes in the standard model. At the present, there is no IND-CCA elliptic curve scheme from the factoring problem in this model.

A relevant parameter for practical purposes is the *expansion factor*, that is, the ratio between the lengths of the ciphertext and the plaintext. The use of large expansion factors leads to decrease the effective bandwidth in secure communications. Some of the known IND-CPA cryptosystems in the standard model achieve expansion factor 2, that is optimal when the encryption uses as much randomness as the message length. Nevertheless, all known IND-CPA elliptic curve cryptosystems based on factoring have expansion factors greater or equal than 4.

In this paper we propose an efficient IND-CPA elliptic curve cryptosystem with expansion factor 2. The design of our scheme is based on [9] but using as underlying primitive Demytko's scheme [6], instead of [11]. This enables to use elliptic curves with arbitrary parameters to design the scheme, in

contrast with [9], where only supersingular curves were possible. This fact would be interesting in case that some particular families of elliptic curves were found to be insecure for cryptographic purposes.

The new proposed cryptosystem uses elliptic curves over the ring \mathbb{Z}_{n^2} , where n is an RSA modulus. Its semantic security is based on a new decisional assumption, namely the Decisional Small Root Assumption. In some sense, this assumption is analogous to the one on which Catalano et al. scheme [2] is based.

Our proposal is efficient, specially in ciphering. Although it is slower than Catalano et al. cryptosystem [2], ours is much faster than the existing IND-CPA elliptic curve schemes based on factoring, such as [8, 15]. Moreover, if our scheme is implemented with a small exponent, its efficiency in encryption is similar to El Gamal like elliptic curve schemes.

The rest of the paper is organised as follows. Section II briefly recalls Demytko's scheme. In Section III, we describe the new scheme and prove it is semantically secure under a new assumption. Then, we argue why one should be confident on this new assumption. The computational cost of the new scheme is discussed in Section IV. Finally, Section V contains some further research.

For a brief description of the results about elliptic curves over the ring \mathbb{Z}_{n^2} used in this paper, see [9].

II. DEMYTKO'S SCHEME REVISITED

Demytko proposed in [6] an elliptic curve RSA based scheme. He uses a randomly chosen elliptic curve $E_n(a, b)$ over the ring \mathbb{Z}_n , where $n = pq$ is an RSA modulus. Let $t_p = p + 1 - |E_p(a, b)|$, $t_q = q + 1 - |E_q(a, b)|$ and e an (small) integer such that

$$\gcd(e, p + 1 \pm t_p) = \gcd(e, q + 1 \pm t_q) = 1. \quad (1)$$

Demytko considered the *quadratic twist*, \mathcal{E} , of $E_n(a, b)$. This set consists on four elliptic curves, with orders $(p + 1 \pm t_p) \cdot (q + 1 \pm t_q)$.

Let $\Lambda_{a,b} = \{x \in \mathbb{Z}_n^* \mid x^3 + ax + b \in \mathbb{Z}_n^*\}$. For all $m \in \Lambda_{a,b}$ there exists a unique curve in \mathcal{E} with exactly four points of the form (m, y) , where y are the four square roots of $m^3 + am + b$ modulo n . Moreover, the x -coordinate of the multiple $e\#(m, y)$, computed on the corresponding curve, is the same for the four values of y and can be computed without knowing any of them. We will hereafter denote by $e \star m$ the x -coordinate of multiple $e\#(m, y)$.

It can be proved that the map

$$\begin{aligned} \mathcal{D}_e : \Lambda_{a,b} &\longrightarrow \Lambda_{a,b} \\ m &\longmapsto e \star m \end{aligned}$$

is well defined and bijective, since an e satisfying (1) is coprime with the number of points of any of the four curves in \mathcal{E} .

In Demytko's scheme, the ciphertext c for a message $m \in \Lambda_{a,b}$ is $c = e \star m$. The ciphertext c can be efficiently computed, for almost all $m \in \Lambda_{a,b}$, as $c = \Phi_e(m) \bmod n$, where the rational function Φ_e is recursively defined as:

$$\begin{aligned}\Phi_1 &= x \\ \Phi_{2k} &= \frac{(\Phi_k^2 - a)^2 - 8b\Phi_k}{4(\Phi_k^3 + a\Phi_k + b)} \\ \Phi_{2k+1} &= \frac{-2(\Phi_k\Phi_{k+1} + a)(\Phi_k + \Phi_{k+1}) + 4b}{(\Phi_{k+1} - \Phi_k)^2} - x\end{aligned}$$

The above formulae can only fail for x corresponding to points (x, y) with order less or equal than e , that are a negligible subset of $\Lambda_{a,b}$, when e is small. Besides, such points can be totally suppressed by taking $E_n(a, b)$ such that $p+1 \pm t_p$ and $q+1 \pm t_q$ have no divisor between 3 and e (e.g. $t_p = t_q = 0, p = 2p' - 1, q = 2q' - 1$ and p', q' primes).

In [6] it is conjectured that Φ_e is a one-way trapdoor permutation with trapdoor p, q and the four inverses of e modulo $\text{lcm}(p+1 \pm t_p, q+1 \pm t_q)$.

To decrypt the ciphertext $c \in \Lambda_{a,b}$, it suffices to compute $m = d \star c$, where d is one of the four inverses of e . The Jacobi symbols $(c^3 + ac + b/p)$ and $(c^3 + ac + b/q)$ easily determine which inverse must be used.

If we restrict the above scheme to supersingular curves (i.e. $t_p = t_q = 0$), there is only one value of d involved and no Jacobi symbol computation is needed in the decryption process, and therefore the values of p and q are not explicitly used in decryption (except for improving speed by using the Chinese Remainder Theorem).

III. THE NEW SCHEME

Applying the ideas in [9], one can add semantic security to Demytko's scheme without loosing efficiency and achieving expansion factor 2.

Let $\Omega_{a,b} = \{x \in \mathbb{Z}_{n^2}^* \mid x^3 + ax + b \in \mathbb{Z}_{n^2}^*\}$. It is easy to see that $\Omega_{a,b} = \{z + mn \mid z \in \Lambda_{a,b}, m \in \mathbb{Z}_n\}$. Let us consider the function

$$\begin{aligned}\Theta_e : \Lambda_{a,b} \times \mathbb{Z}_n &\longrightarrow \Omega_{a,b} \\ (r, m) &\longrightarrow \Phi_e(r) + mn \bmod n^2\end{aligned}$$

Lemma 1 *For all e such that $\gcd(e, n(p+1 \pm t_p)(q+1 \pm t_q)) = 1$, Θ_e is well defined and bijective.*

Proof: Θ_e is well defined since $\Theta_e(r, m) \equiv \mathcal{D}_e(r) \bmod n$, and $\text{Im}(\mathcal{D}_e) = \Lambda_{a,b}$.

Also, due to the bijectivity of \mathcal{D}_e , $\Theta_e(r_1, m_1) = \Theta_e(r_2, m_2)$ implies that $r_1 \equiv r_2 \bmod n$, that is $r_1 = r_2$. Therefore $m_1 = m_2$. ■

In the sequel we describe the proposed new scheme:

Key generation. Given a security parameter ℓ , choose at random two different primes p and q with ℓ bits, a random elliptic curve $E_{n^2}(a, b)$, where $n = pq$, and an integer e such that

$\gcd(e, pq) = \gcd(e, p+1 \pm t_p) = \gcd(e, q+1 \pm t_q) = 1$, where $t_p = p+1 - |E_p(a, b)|$ and $t_q = q+1 - |E_q(a, b)|$.

The public and secret keys are $\text{PK} = (n, e, a, b)$ and $\text{SK} = (p, q, d_1, d_2, d_3, d_4)$, where $d_i = e^{-1} \bmod \text{lcm}(p+1 \pm t_p, q+1 \pm t_q)$.

Encryption. To encrypt a message $m \in \mathbb{Z}_n$ we compute $c = \Theta_e(r, m)$, where r is uniformly chosen in $\Lambda_{a,b}$.

Decryption. To recover the message m from $c = \Phi_e(r) + mn$, notice that $c \equiv \mathcal{D}_e(r) \bmod n$, and the randomness r is obtained from $c \bmod n$ as in Demytko's scheme. Now, m is easily obtained from $mn = c - \Phi_e(r) \bmod n^2$.

As a particular case, the size of private key as well as decryption complexity can be reduced if supersingular curves are used. Then, $t_p = t_q = 0$ and only one value of d is needed to recover r from $c \bmod n$.

A. Semantic security

Probabilistic notation.

If X is a non-empty set, then $x \in_{\mathbb{R}} X$ denotes that x has been uniformly chosen in A . If D_1 and D_2 are two probability distributions, then the notation $D_1 \approx D_2$ means that D_1 and D_2 are polynomially indistinguishable. Notice that if g is a bijection such that g and g^{-1} can be computed in probabilistic polynomial time, then $D_1 \approx D_2$ is equivalent to $g(D_1) \approx g(D_2)$.

Our scheme is semantically secure under the following assumption:

Decisional Small Root Assumption (DSRA).

Let p, q be randomly chosen ℓ -bit long different primes, $n = pq$, $E_{n^2}(a, b)$ a randomly chosen elliptic curve and e an integer such that $\gcd(e, n) = \gcd(e, p+1 \pm t_p) = \gcd(e, q+1 \pm t_q) = 1$, where $t_p = p+1 - |E_p(a, b)|$, $t_q = q+1 - |E_q(a, b)|$.

The following probability distributions are polynomially indistinguishable

$$\begin{aligned}D_{\text{small-x}} &= (n, a, b, \Phi_e(x) \bmod n^2) \quad \text{where } x \in_{\mathbb{R}} \Lambda_{a,b} \\ D_{\text{random}} &= (n, a, b, x') \quad \text{where } x' \in_{\mathbb{R}} \Omega_{a,b}.\end{aligned}$$

Proposition 2 *The proposed scheme is semantically secure if and only if DSRA holds.*

Proof: Semantic security is equivalent to indistinguishability of encryptions, so we have to prove that for all $m_0 \in \mathbb{Z}_n$, the distributions

$$\begin{aligned}D_0 &= (n, a, b, \Phi_e(x) + m_0 n \bmod n^2) \quad \text{and} \\ D &= (n, a, b, \Phi_e(x) + mn \bmod n^2)\end{aligned}$$

where $x \in_{\mathbb{R}} \Lambda_{a,b}$, $m \in_{\mathbb{R}} \mathbb{Z}_n$, are polynomially indistinguishable, which is equivalent to

$$(n, a, b, \Phi_e(x) \bmod n^2) \approx (n, a, b, \Phi_e(x) + m' n \bmod n^2)$$

with $x \in_{\mathbb{R}} \Lambda_{a,b}$ and $m' \in_{\mathbb{R}} \mathbb{Z}_n$.

Note that the distribution on the left side is $D_{\text{small-x}}$. Besides, since $\Phi_e(x) + m' n \bmod n^2 = \Theta_e(x, m')$, and Θ_e is a bijection such that Θ_e and Θ_e^{-1} can be computed in probabilistic polynomial time, then D and D_{random} are identically distributed. ■

B. Hardness of the Small Root Problems In this subsection we argue why one should be confident on the hardness of the new decisional problem presented in this paper. From [18] (Section 3, ex. 3.7) one proves that

$$\Phi_e(x) = \frac{\nu_e(x)}{\eta_e(x)}$$

where $\nu_e(x)$ and $\eta_e(x)$ are relatively prime polynomials such that,

$$\nu_e(x) = x^{e^2} + \text{lower order terms},$$

$$\eta_e(x) = e^2 x^{e^2-1} + \text{lower order terms}.$$

Thus, given $t = \Phi_e(x_0) \bmod n^2$, x_0 is a root of the polynomial $\nu_e(x) - t\eta_e(x) \in \mathbb{Z}_{n^2}[x]$, whose degree is e^2 . Then, DSRA is equivalent to the assumption that it is infeasible deciding if the polynomial $P_e(x) = \nu_e(x) - t\eta_e(x) \in \mathbb{Z}_{n^2}[x]$, with $t \in_{\mathbb{R}} \mathbb{Z}_{n^2}$, has a root smaller than n , for random n , a and b .

Similarly, the semantic security of Catalano et al. scheme is equivalent to the difficulty of deciding if the polynomial $x^e - t \in \mathbb{Z}_{n^2}[x]$, with $t \in_{\mathbb{R}} \mathbb{Z}_{n^2}$, has a root smaller than n . The best known way to attack the above decisional problems is to solve their computational versions. The problem of finding small roots of polynomials modulo a large integer with unknown factorisation has been directly studied in the literature. The most powerful result in this area was obtained by Coppersmith in [3]. This result ensures that one can efficiently compute (i.e. in polynomial time) all roots x_0 of a polynomial $p \in \mathbb{Z}_N[x]$ with degree d such that $|x_0| < N^{1/d}$. Up to now, no improvement on this bound has been made. The result by Coppersmith implies we can find the roots $|x_0| < n^{2/e^2}$ of the polynomial $P_e(x)$. Therefore, for an $e \geq 3$, the described attack does not affect the validity of the DSRA assumption.

IV. EFFICIENCY ANALYSIS

In this section, we study the computational encryption cost of our scheme. Since operations modulo a large number are involved, we neglect the cost of performing additions, multiplications and divisions by small integers. We will express the cost in terms of multiplications mod n^2 , because modular inverses can be computed within a constant number of modular multiplications.

For an exponent of the form $e = 2^s + 1$, the computation of $\Phi_e(r) \bmod n^2$ requires $\lceil \log_2 e \rceil$ evaluations of the formula for Φ_{2k+1} and $\lceil \log_2 e \rceil - 1$ evaluations of the formula for Φ_{2k} . Both rules involve the computation of one inverse modulo n^2 . We point out that $a^{-1} \bmod n^2$ can be obtained by computing $a^{-1} \bmod n$ and then performing two multiplications modulo n^2 . Let c be the number of multiplications modulo n needed to compute $a^{-1} \bmod n$. Since the cost of multiplying two numbers mod n^2 is roughly the cost of 4 multiplications modulo n , we deduce that $a^{-1} \bmod n^2$ can be computed in $2 + c/4$ multiplications modulo n^2 .

Then, our encrypting function requires about $(12 + c/2)s$ multiplications modulo n^2 , when $e = 2s + 1$. Practical implementations, suggest that the value $c = 8$ can be taken (see [1]), so our scheme has an encryption cost of $16s$ multiplications modulo n^2 . This is a very low computational cost if we

compare with previous IND-CPA elliptic curve cryptosystems in the standard model, based on factoring [15, 8]. A lower encryption cost can be achieved if the scheme is designed over a *Montgomery form* elliptic curve (see [13]), and we estimate that the computational cost is reduced in a 50% approximately.

Thereby, if we use $e = 17$, which is a *secure* encryption exponent following the discussion in the previous section, our scheme takes between 64 and 40 products modulo n^2 , depending on the implementation used.

It is interesting to compare our scheme with existing semantically secure ECC in the standard model over finite fields. We will compare the efficiency of our scheme with the well-known El Gamal ECC scheme. We assume that El Gamal ECC is performed over \mathbb{Z}_p^* , where p is 170 bits long, and our scheme is performed over $\mathbb{Z}_{n^2}^*$, where n is 1024 bits long (cf. [12]). We will express both encryption costs in terms of multiplications modulo n^2 .

In El Gamal ECC the most time consuming operation is the computation of two multiples $r\#P$ and $ra\#P$, where r is a random integer which size is roughly the same as the modulus p , and a is a fixed integer. Then, using the *double and add* algorithm, the computation of these two multiples requires on average k additions of points and $2k$ doublings, where k is the bit length of r . Assuming that a point addition or doubling requires about 12 modular multiplications, then El Gamal ECC would take approximately $3 \cdot 170 \cdot 12$ multiplications modulo p . Since the time needed to perform a modular multiplication is quadratic in the size of the modulus, the ratio between the time of a multiplication modulo p and a multiplication modulo n^2 is $\frac{170^2}{(2 \cdot 1024)^2}$. It follows that the encryption time of El Gamal ECC would be equivalent to 42 multiplications modulo n^2 .

V. FURTHER RESEARCH

Starting from the proposed scheme, an IND-CCA cryptosystem in the ROM is obtained applying some generic transformation, like [16] or [7]. It would be interesting to provide IND-CCA security in the standard model to Catalano et al. scheme as well as to ours. To achieve this goal, the recent work of Cramer and Shoup [5] could provide useful ideas.

Since the publication of Paillier's cryptosystem [14], several new decisional assumptions have been formulated (e.g in [2],[8],[9]). A careful study of these decisional problems is also needed, because very few things are known at the present.

REFERENCES

- [1] R. P. Brent. Some Integer Factorization Algorithms using Elliptic Curves. *Australian Computer Science Communications* 24–26 (1986) (Republished 1998).
- [2] D. Catalano, R. Gennaro, N. Howgrave-Graham and P. Q. Nguyen. Paillier's Cryptosystem Revisited. *ACM CCS '2001 ACM Press* (2001).
- [3] D. Coppersmith. Finding a small root of a univariate modular equation. *EUROCRYPT '96, LNCS 1070* 155–165 (1996).
- [4] R. Cramer and V. Shoup. A Practical Public Key Cryptosystem Provably Secure against Adaptive Chosen Ciphertext Attack. *CRYPTO '98, LNCS 1462* 13–25 (1998).
- [5] R. Cramer and V. Shoup. Universal Hash Proofs and a Paradigm for Adaptive Chosen Ciphertext Secure Public-Key Encryption. *EUROCRYPT '2002, LNCS 2332* 45–64 (2002).
- [6] N. Demytko. A new elliptic curve based analogue of RSA. *EUROCRYPT '93, LNCS 765* 40–49 (1993).

- [7] E. Fujisaki and T. Okamoto. Secure integration of asymmetric and symmetric encryption schemes. *CRYPTO'99 LNCS* **1666** 53–68 (2000).
- [8] S. Galbraith. Elliptic curve Paillier schemes. *Journal of Cryptology* **15** (2) 129–138 (2002).
- [9] D. Galindo, S. Martín, P. Morillo and J. L. Villar. An efficient semantically secure elliptic curve cryptosystem based on KMOV. To appear at *Proceedings of WCC '03*. Also available at Cryptology ePrint Archive, Report 2002/037. <http://eprint.iacr.org/>
- [10] S. Golwasser and S. Micali. Probabilistic encryption. *Journal of Computer and System Sciences* **28** 270–299 (1984).
- [11] K. Koyama, U.M. Maurer, T. Okamoto and S.A. Vanstone. New Public-Key Schemes Based on Elliptic Curves over the Ring \mathbb{Z}_n . *CRYPTO '91, LNCS* **576** 252–266 (1991).
- [12] A. K. Lenstra and E. R. Verheul. Selecting Cryptographic Key Sizes. <http://cryptosavvy.com/cryptosizes.pdf>
- [13] Montgomery, P.L. Speeding the Pollard and Elliptic Curve Methods of Factorizations. *Math. Comp.* **48** 243–264 (1987).
- [14] P. Paillier. Public-Key Cryptosystems Based on Composite Degree Residuosity Classes. *EUROCRYPT '99, LNCS* **1592** 223–238 (1999).
- [15] P. Paillier. Trapdooring discrete logarithms on elliptic curves over rings. *ASIACRYPT '00, LNCS* **1976** 573–584 (2000).
- [16] D. Pointcheval. Chosen-Ciphertext Security for any One-Way Cryptosystem. *Proc. PKC '2000 LNCS* **1751** 129–146 (2000).
- [17] C. Rackoff and D.R. Simon. Non-interactive zero-knowledge proof of knowledge and chosen-ciphertext attack. *CRYPTO'91 LNCS* **576** 433–444 (1992).
- [18] J.H. Silverman. The arithmetic of elliptic curves. *Springer GTM* **106** (1986).