Improved known plaintexts attack on Domingo-Ferrer homomorphic cryptosystem[1]

A.V. Trepacheva <alina1989malina@ya.ru>
Southern Federal University,
105/42, Bolshaya Sadovaya st., Rostov-on-Don, 344006, Russia.

Abstract. This paper is devoted to known plaintexts cryptanalysis of homomorphic cryptosystem proposed by Domingo-Ferrer. In previous works it was shown that at least $d+1$ pairs (plaintext, ciphertext) are necessary to recover secret key, where d is a degree of polynomials representing ciphertexts. Here we analyze existing known plaintext attack. And also slightly modified attack on this cryptosystem is presented. It allows to reduce the necessary number of pairs meaningfully. In particular interception only of two pairs may be enough for successful key recovering with overwhelming probability. The running time of our attack depends polynomially on d and logarithmically on plaintexts space size as well as for previous attack. We provide the results of computer experiments.

Key words: known plaintext cryptanalysis; homomorphic encryption; cloud computations.

1. Introduction

Homomorphic encryption (HE) is a cryptographic primitive supporting the additional property in comparison with ordinary encryption: *HE allows computing over encrypted data.* Let's explain what this means. We assume that plaintexts space P and ciphertexts space C are rings with operations $+_P, \cdot_P$ and $+_C, \cdot_C$ correspondingly. And let E, D be encryption and decryption functions of cryptosystem ε. The last one is homomorphic if for $\forall x, y \in P$ and $\forall E(x), E(y) \in C$ the following properties are satisfied:

$$D(E(x) +_C E(y)) = x +_P y, \quad (1)$$

$$D(E(x) \cdot_C E(y)) = x \cdot_P y. \quad (2)$$

So the result of computations over ciphertexts will be an encryption of computations result over underlying plaintexts.

[1] This work is supported by grant RFBR 15-07-00597-a

Homomorphic cryptosystems (HC) are of key importance for protecting sensitive data in clouds. Computationally weak clients may outsource computations over their data while keeping this data in secret. This makes the development of new homomorphic cryptosystems and cryptanalysis of existing a hot topic.

By the present moment a variety of homomorphic cryptosystems were proposed (for example see [1-5]). RSA [1] is one of the most well known, because the product of RSA ciphertexts is an encryption of corresponding plaintexts product. But cryptosystems [1-5] are partially homomorphic, because they allow to compute over ciphertexts only functions lying in some bounded class. In particular for [1] only property (2) holds (multiplicatively homomorphic cryptosystems). Whereas for instance for [2] only (1) holds (additively homomorphic).

The simplest example of HC holding both (1), (2) was introduced in the fundamental paper [6] of Rivest, Adleman and Dertouzos. Encryption function $E: \mathbb{Z}_n \to \mathbb{Z}_p \times \mathbb{Z}_q$ works as follows $x \in \mathbb{Z}_n \to (x \bmod p, x \bmod q)$. Unfortunately, in [7] such encryption was shown to be unsecure against known plaintext attack (KPA). Beginning with [6] lots of cryptosystems with properties (1), (2) were suggested. Here two the most important groups may be highlighted. In the first group there are cryptosystems [8-11] with unlimited ciphertexts sizes growth during computing over them (their security analysis may be founded in [12,13]). Whereas cryptosystems of second group have some polynomially bounds on ciphertexts sizes growth. In this group for example there are cryptosystems [14-18] belonging to direction initiated by innovative work [14] of IBM researcher Craig Gentry.

Second group obviously is more interesting for practice. But unfortunately existing cryptosystems are not enough efficient for usage in real applications. The development of Gentry-like HCs now has mostly theoretical character. And in practice at the present moment HCs from the first group are used. For instance cryptosystems [10, 11] proposed by Domingo-Ferrer are exploited in secure packet forwarding in mobile ad hoc networks (see [19-24]). The main reason is a conceptual simplicity of constructions from [10, 11].

In the light of this the analysis of Domingo-Ferrer HCs resistance to different attacks is of value. Here we will concentrate on KPA. In [25] the authors described KPA on [10] and showed that to recover secret key an adversary A should intercept $t \geq d+1$ pairs (plaintext, ciphertext), where d is a degree of polynomials representing ciphertext. The aim of the present work to demonstrate that [10] may be broken using even two pairs (plaintext, ciphertext). We give some theoretical reasoning to this fact. And also we provide an experimental confirmation.

2. Denotations

All logarithms are base-2. A probability of event M is denoted by $\Pr(M)$, ring of integers – by $\mathbb{Z}$, ring of integers modulo n – by $\mathbb{Z}_n$, the multiplicative subgroup of $\mathbb{Z}_n$ – by $\mathbb{Z}_n^*$. An adversary trying to break cryptosystem will be denoted by A. For

symmetric cryptosystem ε: P – plaintexts space, C – ciphertexts space, K – secret keys space, $\mathbb{D}$ – probabilistic distribution over P.

We denote by $x \xleftarrow{\$} R$ a random element sampled according to uniform distribution over ring R and also by $x \xleftarrow{\mathbb{D}} R$ – random ring element generated according distribution $\mathbb{D}$ over R. Denotation $f(x) \xleftarrow{\$} R[x]$ means that all coefficients of polynomial f are random values chosen uniformly and independently from R.

3. Overview of Domingo-Ferrer cryptosystem

Let's briefly recall cryptosystem from [10]. The author sets $P = \mathbb{Z}_n$, $C \subset \mathbb{Z}_p[x] \times \mathbb{Z}_q[x]$, $K = \mathbb{Z}_p^* \times \mathbb{Z}_q^*$, where $n = p \cdot q$, $p, q -$ big primes, $p < q$, $\log p \approx \log q$, i.e. n – RSA modulus. Its factorization is a secret. Secret key is a pair $k = (r_p, r_q) \in K$. Before encryption public parameter $d \in \mathbb{Z}_+$ is fixed.

Encryption($a \in \mathbb{Z}_n, d \in \mathbb{Z}_+, p, q, k = (r_p, r_q) \in K$ **):**

- $a \in \mathbb{Z}_n \to a'(x) \in \mathbb{Z}_n[x]$, where $a'(x) = \sum_{i=1}^{d} a'_i \cdot x^i$ and for $i = \overline{2, d-1}$: $a'_i \xleftarrow{\$} \mathbb{Z}_n$, $a'_d \xleftarrow{\$} \mathbb{Z}_n \setminus \{0\}$ and $a'_1 := (a - \sum_{i=1}^{d-1} a'_i) \bmod n$.
- Ciphertext is a pair of polynomials $c = (c_p(x), c_q(x))$, where $c_p(x) := a'(r_p \cdot x) \bmod p$ and $c_q(x) := a'(r_q \cdot x) \bmod q$.

One may see that $a \equiv a'(1) (\bmod n)$ (or $a \equiv \sum_{i=1}^{d} a'_i (\bmod n)$).

Decryption($c = (c_p(x), c_q(x)), p, q, k^{-1} = (r_p^{-1}, r_q^{-1})$ **):**

- $a'_p(x) := c_p(r_p^{-1} \cdot x) \bmod p$ $a'_q(x) := c_q(r_q^{-1} \cdot x) \bmod q$ (clear $a'_p(x) \equiv a'(x) (\bmod p)$ and $a'_q(x) \equiv a'(x) (\bmod q)$).
- $a_p := a'_p(1) \bmod p$, $a_q := a'_q(1) \bmod q$ (clear $a \equiv a_p (\bmod p)$, $a \equiv a_q (\bmod q)$).
- $a := CRT(a_p, a_q, p, q)$, where $CRT(a_p, a_q, p, q)$ means the reconstruction of $a \in \mathbb{Z}_n$ by $a_p \in \mathbb{Z}_p$, $a_q \in \mathbb{Z}_q$ using Chinese reminder theorem.

In [10] the author suggested two regimes of cryptosystem working. In the first variant modulus n is public and plaintexts and ciphertexts coefficients are treated by untrusted party as elements of $\mathbb{Z}_n$. In the second case n is hidden for providing

higher level of security. And then plaintexts and ciphertexts coefficients are treated as elements of $\mathbb{Z}$. Here we will consider only the first case.

Homomorphic properties: Let's suppose there are plaintexts $a_1, a_2 \in \mathbb{Z}_n$ and $c_1 = (c_{p,1}(x), c_{q,1}(x))$, $c_2 = (c_{p,2}(x), c_{q,2}(x))$ – its encryptions made on the same key $k = (r_p, r_q)$ and for the same d. In [10] the author proves the following statements.

Statement 1. Ciphertext $c_+ = ((c_{p,1}(x) + c_{p,2}(x)) \bmod n, (c_{q,1}(x) + c_{q,2}(x)) \bmod n)$ is a correct encryption of plaintext $(a_1 + a_2) \bmod n \in \mathbb{Z}_n$ for key $k = (r_p, r_q)$ and parameter d.

Statement 2. Ciphertext $c_* = ((c_{p,1}(x) \cdot c_{p,2}(x)) \bmod n, (c_{q,1}(x) \cdot c_{q,2}(x)) \bmod n)$ is a correct encryption of plaintext $(a_1 \cdot a_2) \bmod n \in \mathbb{Z}_n$ for key $k = (r_p, r_q)$ and parameter $2 \cdot d$.

One may see that multiplication of ciphertexts causes an unbounded growth of their sizes (the size is doubled). So in general this HC isn't good for practice. But its simplicity makes it good for applications requiring only computations of some special functions (see [19-24]).

Remark 1. In practice for example $\log n \approx 2048$ may be chosen. Then the size S of ciphertext is $2048 \cdot d$ bits. This implies that $d \le 500$ should be chosen to obtain $S \le 10^6$ bits. Such setting seems reasonable because in all latest HCs [14-18] S is usually about 10^6 bits. Larger value of S will make homomorphic computations too much expensive. But of course it is suitable only if additive homomorphism is necessary. But if multiplicative homomorphism will be exploited then d should be smaller.

4. Cryptanalysis of Domingo-Ferrer cryptosystem

4.1 Existing KPA

Here we briefly discuss existing results [25] concerning known plaintexts analysis of Domingo-Ferrer cryptosystem [10]. Let's suppose A has t pairs $(a_i \in P, c_i \in C), i = \overline{1, t}$, where c_i is an encryption of a_i and all c_i are produced for the same n, $k = (r_p, r_q)$ and d. Ciphertexts c_i are pairs $(c_{p,i}(x) \in \mathbb{Z}_p[x], c_{q,i}(x) \in \mathbb{Z}_q[x])$, where $c_{p,i}(x) = \sum_{j=1}^{d} c_{p,i,j} \cdot x^j$, $c_{q,i}(x) = \sum_{j=1}^{d} c_{q,i,j} \cdot x^j$.

A needs to recover $p, q, k^{-1} = (r_p^{-1}, r_q^{-1})$ using n and $(a_i \in P, c_i \in C), i = \overline{1, t}$.

Remark 2. Here we consider the case of public n. So before recovering p, q A works with polynomials $c_{p,i}(x), c_{q,i}(x)$ modulo n. In [25] the authors also propose

an attack for hidden n. And in this case coefficients $c_{p,i,j}$, $c_{q,i,j}$ are treated as integers at the first step of KPA.

According to encryption procedure the following congruences holds:

$$c_{p,i}(r_p^{-1}) - a_i \equiv 0 (\mathrm{mod}\ p), \quad (3)$$

$$c_{q,i}(r_q^{-1}) - a_i \equiv 0 (\mathrm{mod}\ q). \quad (4)$$

So polynomials $f_i(x) = c_{p,i}(x) - a_i \in \mathbb{Z}_n[x], i = \overline{1,t}$ have a common root r_p^{-1} modulo p. Similarly $g_i(x) = c_{q,i}(x) - a_i \in \mathbb{Z}_n[x], i = \overline{1,t}$ have a common root r_q^{-1} modulo q. And please note that r_p^{-1}, r_q^{-1} are not obligatory roots of $f_i(x)$, $g_i(x)$ modulo n. So KPA should proceed in three steps:

- A recovers secret modulus p and sets $q = n / p$.
- A computes r_p^{-1} as a common root of $f_i(x), i = \overline{1,t}$ modulo p.
- A computes r_q^{-1} as a common root of $g_i(x), i = \overline{1,t}$ modulo q.

4.1.1 Recovering of modulus p

For computing p in [25] the authors propose to consider the following matrix $\mathbf{A} \in \mathbb{Z}_n^{t \times (d+1)}$:

$$\mathbf{A} = \begin{bmatrix} -a_1 & c_{p,1,1} & \dots & c_{p,1,d} \\ -a_2 & c_{p,2,1} & \dots & c_{p,2,d} \\ \dots & \dots & \dots & \dots \\ -a_t & c_{p,t,1} & \dots & c_{p,t,d} \end{bmatrix}.$$

According to (3) homogeneous system of linear equations $(\mathbf{A} \mid \mathbf{0})$ has a nontrivial solution modulo p:

$$\mathbf{v}^T = (1, r_p^{-1}, (r_p^{-1})^2, \dots, (r_p^{-1})^d).$$

Therefore for $t = d+1$ $\mathbf{A}$ is a square matrix having zero determinant modulo p. Then equality $\det(\mathbf{A}) = p \cdot s \in \mathbb{Z}_n, s \in \{0,1,\dots,q-1\}$ holds. The last one means that if $s \neq 0$ p may be recovered as follows:

$$\mathrm{p} := \mathrm{GCD}(\det(\mathbf{A}), n).$$

According to Chinese reminder theorem we have $\det(\mathbf{A}) = (\det(\mathbf{A}) \bmod q) \cdot p \cdot (p^{-1} \bmod q)$. So $s = 0$ if and only if $\det(\mathbf{A}) \bmod q = 0$. The authors of [25] prove that

$$\Pr(\det(\mathbf{A}) \bmod q \neq 0) > e^{-3/2 \cdot (\mathrm{p}-1)} \quad (5),$$

where for large p value $e^{-3/2 \cdot (\mathrm{p}-1)} \approx 1$. Thus having $d+1$ pairs (plaintext, ciphertext) A may recover p with probability ≈ 1. Asymptotical complexity of computing p using this method is $O(d^3 \cdot \log^2(n))$.

Remark 3. Inequality (5) in [25] was proven using assumptions that $c_{p,i,j} \xleftarrow{\$} \mathbb{Z}_p$ and $a_i \bmod q \xleftarrow{\$} \mathbb{Z}_q$. But of course this is correct only if probabilistic distribution $\mathbb{D}$ over P is uniform. For not uniform $\mathbb{D}$ (5) is not true. In the worst case $\mathbb{D}$ may be such that $\Pr(0) \approx 1$ and for moderate values of d $\Pr(\det(\mathbf{A}) \bmod q = 0) > 1/2$, because if the first column of $\mathbf{A}$ is a zero vector then $\det(\mathbf{A}) \bmod q = 0$ holds. So for such $\mathbb{D}$ the probability of successful cryptanalysis is not so good. In general additional study is necessary, because it is not immediately clear how to estimate $\Pr(\det(\mathbf{A}) \bmod q \neq 0)$ for arbitrary $\mathbb{D}$.

4.1.2 Recovering of r_p^{-1}, r_q^{-1}

Now we suppose $t = d+1$ and p is recovered using $(a_i \in P, c_i \in C), i = \overline{1,t}$. The first way to compute r_p^{-1} is to solve the system of linear equations $(\mathbf{A} \mid \mathbf{0})$. The second way is to compute:

$$f(x) = GCD(f_{p,1}(x), \dots, f_{p,d+1}(x)),$$

where $f_{p,i}(x) := f_i(x) \bmod p = c_{p,i}(x) - a_{p.i}$, $a_{p.i} := a_i \bmod p$. Obviously

$$f(x) = (x - r_p^{-1}) \cdot GCD(f_{p,1}^0(x), \dots, f_{p,d+1}^0(x))$$

holds, where $f_{p,i}^0(x) = f_{p,i}(x) / (x - r_p^{-1}) \in \mathbb{Z}_p[x]$, $i = \overline{1,d+1}$. If $GCD(f_{p,1}^0(x), \dots, f_{p,d+1}^0(x)) = 1$ then $f(x) = x - r_p^{-1}$ and therefore r_p^{-1} is recovered. Based on assumption that for all $i = \overline{1,d+1}$: $f_{p,i}^0(x) \xleftarrow{\$} \mathbb{Z}_p[x]$, $\deg(f_{p,i}^0(x)) = d-1$, the authors of [25] give an estimation

$$\Pr(f(x) = x - r_p^{-1}) = \Pr(GCD(f_{p,1}^0(x), \dots, f_{p,d+1}^0(x)) = 1) > (1 - 1/p^d)^{d-1}. \quad (6)$$

So for large p and moderate d the probability to recover r_p^{-1} becomes close to 1.

Remark 4. Both ways to compute r_p^{-1} have equivalent complexity $O(d^3 \cdot \log^2(p))$.

In [25] the authors didn't give a proof that all $f_{p,i}^0(x)$ are uniformly random. So here we fill this gap.

Statement 3. Let distribution $\mathbb{D}$ is uniform and let there is a polynomial $f(x) = c_p(x) - a \in \mathbb{Z}_n[x], \deg(f) = d$ constructed using pair $(a, c = (c_p(x), c_q(x))$.

Then $f_p^0(x) = f_p(x)/(x - r_p^{-1}) \in \mathbb{Z}_p[x]$ is uniformly random with $\deg(f_p^0(x)) = d-1$, where $f_p(x) := f(x) \bmod p$.

Proof: Let's look at $f_p(x) = \sum_{i=0}^{d} f_{p,i} \cdot x^i \in \mathbb{Z}_p[x]$. According to encryption procedure $f_{p,i} := (a'_i \cdot r_p^i) \bmod p, i = \overline{1,d}$ and $f_{p,0} := (-\sum_{i=1}^{d} a'_i) \bmod p \ (:= (-a) \bmod p)$. Using ordinary polynomial division it's easy to verify that $f_p^0(x) = f_p(x)/(x - r_p^{-1}) = \sum_{i=0}^{d-1} f_{p,i}^0 \cdot x^i$, where $f_{p,d-1}^0 \equiv r_p^d \cdot a'_d \pmod p$, $f_{p,d-2}^0 \equiv r_p^{d-1} \cdot (a'_d + a'_{d-1}) \pmod p$, …, $f_{p,1}^0 \equiv r_p^2 \cdot (a'_d + a'_{d-1} + \ldots + a'_2) \pmod p$ and $f_{p,1}^0 \equiv r_p \cdot (a'_d + a'_{d-1} + \ldots + a'_1) \pmod p \equiv r_p \cdot a \pmod p$. Coefficients $f_{p,i}^0, i = \overline{0,d-1}$ are independent random values, where $f_{p,i}^0 \xleftarrow{\$} \mathbb{Z}_p, i = \overline{1,d-2}$, $f_{p,d-1}^0 \xleftarrow{\$} \mathbb{Z}_p \setminus \{0\}$, $f_{p,0}^0 \xleftarrow{\mathbb{D}} \mathbb{Z}_p$. So obviously if $\mathbb{D}$ is uniform then $f_p^0(x) \xleftarrow{\$} \mathbb{Z}_p[x]$ and $\deg(f_p^0(x)) = d-1$. □

One may see that for not uniform $\mathbb{D}$ polynomials $f_{p,i}^0(x), i = \overline{1,d+1}$ are not uniformly random . And in this case it is not clear whether estimation (6) is true. Thus additional study should be carried out.

Let's turn on to the uniform $\mathbb{D}$. We would like to note that in this case instead of estimation (6) one may obtain the exact value of $\Pr(GCD(f_{p,1}^0(x), \ldots, f_{p,d+1}^0(x)) = 1)$. In [26] the following result based on Euclidean algorithm was proved.

Corollary 1 ([26]). Let $(d_1, \ldots, d_m)$ be an ordered m-tuple of nonnegative integers (not all zero) and for $1 \le i \le m$ let $a_i(x) \xleftarrow{\$} \mathbb{Z}_p[x]$ $\deg(a_i(x)) = d_i$, where p is a prime. Then the probability that $a_1(x), \ldots, a_m(x)$ are relatively prime is $1 - 1/p^{m-1}$.

Based on this corollary we have $\Pr(GCD(f_{p,1}^0(x), \ldots, f_{p,d+1}^0(x)) = 1) = 1 - 1/p^d$ that is ≈ 1 for large p.

Similarly $g(x) = GCD(g_{q,1}(x), \ldots, g_{q,d+1}(x)) = x - r_q^{-1}$ with probability $1 - 1/q^d$, where $g_i(x) = c_{q,i}(x) - a_i \in \mathbb{Z}_n[x], i = \overline{1,d+1}$, $g_{q,i}(x) := g_i(x) \bmod q = c_{q,i}(x) - a_{q,i}$, $a_{q,i} := a \bmod q$. And finally we obtain that the probability to recover r_p^{-1}, r_q^{-1} is equal to $(1 - 1/p^d) \cdot (1 - 1/q^d)$. It should be noted that the last one is true because according to encryption procedure for uniform $\mathbb{D}$ for $\forall i$ polynomials $f_{p,i}(x)$ and $g_{q,i}(x)$ may be considered as independent random polynomials.

Summarizing all said above we see that KPA proposed in [25] requires $t \ge d+1$ pairs (plaintext, ciphertext) to recover secret key with probability $\Pr \approx 1$. But estimation $\Pr \approx 1$ is proved only for uniform $\mathbb{D}$. The total asymptotical complexity of KPA is $O(d^3 \cdot \log^2(n))$.

4.2 Our improvement of KPA

Now we discuss how to reduce the number of pairs t necessary for successful KPA on cryptosystem [10]. First we recall the notion of resultant for two polynomials. Let there are $f(x) = \sum_{i=0}^{d_1} f_i \cdot x^i, g(x) = \sum_{i=0}^{d_2} g_i \cdot x^i \in \mathbb{Z}_n[x]$. One may compose a Sylvester matrix $\mathbf{S} \in \mathbb{Z}_n^{(d_1+d_2)\times(d_1+d_2)}$ for $f(x), g(x)$:

$$\mathbf{S} = \begin{vmatrix} f_0 & \ldots & f_{d_1} & 0 & 0 & \ldots & 0 \\ 0 & f_0 & \ldots & f_{d_1} & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & 0 & f_0 & \ldots & f_{d_1} \\ g_0 & \ldots & g_{d_2} & 0 & 0 & \ldots & 0 \\ 0 & g_0 & \ldots & g_{d_2} & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & 0 & g_0 & \ldots & g_{d_2} \end{vmatrix}. \quad (7)$$

The resultant of polynomials $f(x), g(x) \in \mathbb{Z}_n[x]$ is defined as follows: $\Theta = \mathrm{Re}\,s(f(x), g(x)) = \det(\mathbf{S}) \bmod n \in \mathbb{Z}_n$. It is well known result that $\Theta = 0$ if and only if $f(x)$ and $g(x)$ have at least one common root or factor modulo n (for details see [27]). For further discussion we need the following simple statements.

Statement 4 . If for $n = p \cdot q$ polynomials $f(x), g(x) \in \mathbb{Z}_n[x]$ have at least one common root or factor modulo p (or q) then $\Theta_p = 0$ (or $\Theta_q = 0$), where $\Theta_p := \Theta \bmod p$, $\Theta_q := \Theta \bmod q$.

Statement 5. If $n = p \cdot q$, where $p \ne q, GCD(p,q) = 1$, then $\Theta = 0$ if and only if $\Theta_p = 0$, $\Theta_q = 0$.

We skip the proof because this statements may be immediately derived from Chinese reminder theorem and congruences properties.

Let's return to KPA on cryptosystem [10]. Now we will demonstrate that interception only of two pairs (plaintext, ciphertext) may be enough to recover factorization of n and $k = (r_p, r_q)$.

4.2.1 Recovering of modulus p

Let's suppose A intercepted $(a_i, c_i = (c_{p,i}(x) \in \mathbb{Z}_p[x],\ c_{q,i}(x) \in \mathbb{Z}_q[x])), i = \overline{1,2}$, where $\deg(c_{p,i}(x)) = d, \deg(c_{q,i}(x)) = d$. Let's look at the resultant $\Theta = \mathrm{Re}\,s(f_1(x), f_2(x)) \in \mathbb{Z}_n$, where $f_i(x) = c_{p,i}(x) - a_i \in \mathbb{Z}_n[x], i = 1,2$. As we've already seen $f_1(x), f_2(x)$ have a common root r_p^{-1} modulo p. According to statement 4 $\Theta_p = 0$ and hence $\Theta = p \cdot s, s \in \{0,1,...,q-1\}$. So for $s \neq 0$ A can compute p according formula:

$$\mathrm{p} := \mathrm{GCD}(\Theta, n).$$

Please note that the last one is true because here q is prime and $\mathrm{GCD}(s,q) = 1$ for $s \neq 0, q$.

As a result we obtain that to recover p it's enough to have only two pairs $(a_i, c_i), i = \overline{1,2}$ with $\Theta \neq 0$. So it's necessary to find out how much the probability $\mathrm{Pr}_0 = \Pr(\Theta \neq 0)$ for randomly intercepted pairs. To estimate Pr_0 we should note that according to statement 5 $\Theta = 0$ if and only if $\Theta_q = 0$ and then $\mathrm{Pr}_0 = \Pr(\Theta_q \neq 0)$. Obviously $\Theta_q \neq 0$ if and only if $GCD(f_{q,1}(x), f_{q,2}(x)) = 1$, where $f_{q,i}(x) = f_i(x) \bmod q \in \mathbb{Z}_q[x], i = 1,2$. If $f_{q,1}(x), f_{q,2}(x)$ were uniformly random in $\mathbb{Z}_q[x]$ then $\mathrm{Pr}_0 = \Pr(\Theta_q \neq 0)$ would be equal to $1 - 1/q$ according to corollary 1. But unfortunately in fact $f_{q,\mathrm{i}}(x) = \sum_{j=0}^{d} f_{q,i,j} \cdot x^j, i = 1,2$ are not strictly uniform even if distribution $\mathbb{D}$ is uniform. Indeed for uniform $\mathbb{D}$ there are $f_{q,i,j} \xleftarrow{\$} \{0,1,...,p-1\}, j = \overline{1,d-1}$, $f_{q,i,\mathrm{d}} \xleftarrow{\$} \{1,...,p-1\}$ and $f_{q,i,0} \xleftarrow{\$} \mathbb{Z}_q$. Estimation

$$\mathrm{Pr}_0 \approx 1 - 1/q \quad (8)$$

we are not ready to prove now. But (8) correlates very good with computer experiments. In tables 1,2 we present practical estimation of Pr_0 for uniform $\mathbb{D}$ for different d.

Remark 5. Cryptosystem from [10] and presented KPA were implemented using Qt 1.3.1 and NTL library [28]. For practical estimation of Pr_0 two pairs (a_i, c_i) were generated randomly 10^5 times. Then the number of cases with $\Theta_q \neq 0$ was counted. The case of not uniform $\mathbb{D}$ should be studied additionally. The only thing we can say now that in the worst case $\mathbb{D}$ may be such that $\Pr(0) = \beta$, where $\beta \approx 1$ and then $\mathrm{Pr}_0 = \Pr(\Theta_q = 0) > \beta^2$ that is ≈ 1. So for such $\mathbb{D}$ this KPA fails with overwhelming probability.

Table 1. Estimations of Pr_0 *for different* p, q *and* $d = 10$.

n	p	q	Practical estimation of Pr_0	$1 - 1/q$
6	2	3	0.67	0.67
35	5	7	0.86	0.86
91	7	13	0.922	0.923
253	11	23	0.956	0.957
1517	37	41	0.97	0.97
3599	59	61	0.98	0.99
9991	97	103	0.99	0.991

Table 2. Estimations of Pr_0 *for different* p, q *and* $d = 50$.

n	p	q	Practical estimation of Pr_0	$1 - 1/q$
15	3	5	0.8	0.8
221	13	17	0.92	0.94
1147	31	37	0.954	0.972
2173	41	53	0.999	0.999
13943	103	131	0.999	0.999

The asymptotical complexity of this method to recover p is $O(d^3 \cdot \log^2(n))$.

Finally we would like to note that the idea to compute resultant of polynomials for recovering p we borrow from [29]. In [29] the author presented KPA on another Doming-Ferrer homomorphic cryptosystem [11]. Encryption in [11] works similar to [10]. Plaintext $a \in \mathbb{Z}_{n'}$ first is mapped into random polynomial $a'(x) \in \mathbb{Z}_{n'}[x]$ such that $a'(1) \equiv a (\bmod n')$, $\deg(a'(x)) = d, a'(0) = 0$. Ciphertext is a polynomial $c(x) \in \mathbb{Z}_n[x]$ such that $c(x) := a'(r \cdot x) \bmod n$, where $r \in \mathbb{Z}_n^*$ – secret key, n – big integer ($\log(n) \approx 1000$) with many small divisors, $n' \mid n$ and $\log(n') \approx 100$. Modulus n' is hidden and n is public. It should be pointed out that in spite of similarity construction from [10] is not a special case of [11] and vice versa.

To break cryptosystem [11] A first should compute n' and second $(r')^{-1} := r^{-1} \bmod n'$ as a common root of polynomials $f_i(x) = c_i(x) - a_i \in \mathbb{Z}_n[x], i = \overline{1,t}$ modulo n'. According to congruences properties $(r')^{-1}$ may be used for decryption instead of r^{-1}. For recovering n' in [29] the author proposes to compute $n'' = GCD(n, \mathrm{Re}\,s(f_1, f_2), \mathrm{Re}\,s(f_3, f_3), ..., \mathrm{Re}\,s(f_{t-1}, f_t))$. Obviously

$\Pr(n'' = n') = \Pr(GCD(n/n', \mathrm{Re}\,s(f_1, f_2)/\mathrm{n}', \mathrm{Re}\,s(f_3, f_3)/\mathrm{n}', ..., \mathrm{Re}\,s(f_{t-1}, f_t)/\mathrm{n}') = 1)$

(/ is integer division) holds. Here in contrast to [10] it's not enough to take $t = 2$,

because n has many small divisors. So to estimate $\Pr_0 = \Pr(GCD(n/n', \mathrm{Re}s(f_1,f_2)/\mathrm{n}', \mathrm{Re}s(f_3,f_3)/\mathrm{n}', ..., \mathrm{Re}s(f_{t-1},f_t)/\mathrm{n}') = 1)$ one should involve a known result about the probability that randomly chosen integers are coprime. According to this result $\Pr_0 \approx 1/\zeta(\mathrm{t}/2+1)$ holds (we suppose t is even), where ζ is Riemann's zeta function. So for $t=2$ we have $\Pr_0 \approx 0,61$. That is not enough of course. To obtain $\Pr_0 \approx 1$ one should take $t > 100$.

Summarizing all said above we would like to stress out that idea of computing resultants doesn't work so good for cryptosystem [11], because A *must intercept many pairs to recover secret modulus with overwhelming probability. But for [10] computing resultant allows to decrease* t *meaningfully. Now the only case in which we while don't know how to find* p *is* $t=1$.

4.2.2 Recovering of r_p^{-1}, r_q^{-1}

For recovering r_p^{-1} A may compute

$$f(x) = GCD(f_{p,1}(x), f_{p,2}(x)) \in \mathbb{Z}_p[x],$$

where $f_{p,i}(x) := f_i(x) \bmod p$, $f_i(x) = c_{p,i}(x) - a_i \in \mathbb{Z}_n[x], i = 1,2$. For uniform $\mathbb{D}$ according to corollary 1 we obtain $\Pr(f(x) = x - r_p^{-1}) = 1 - 1/p$ that is ≈ 1 for large p. Similarly r_q^{-1} may recovered with probability $1 - 1/\mathrm{q}$. So the total probability to find r_p^{-1}, r_q^{-1} now is $\Pr_1 = (1-1/p)\cdot(1-1/\mathrm{q})$. The last one is ≈ 1 for large p,q.

The asymptotical complexity of computing r_p^{-1}, r_q^{-1} now is $O(d^2 \cdot \log^2(q))$.

To conclude we would like to present the total running time T of our KPA (time to recover p,q and r_p^{-1}, r_q^{-1}). Time measurements were done using PC with the following characteristics: Quad Core Celerone 1,7 GHz with 4 GB memory.

Table 3. Running time of KPA.

d	T for $\log n = 2^{10}, \log p = 2^9$	T for $\log n = 2^{11}, \log p = 2^{10}$
8	38 ms	112 ms
16	121 ms	387 ms
32	460 ms	1.5 s
64	1.9 s	6 s
128	9.5 s	27 s
256	52 s	2 min
512	5 min	12 min
1024	22 min	50 min

5. Conclusion

We have analysed the existing method [25] of known plaintext cryptanalysis of Domingo-Ferrer homomorphic cryptosystem [10]. This analysis shows that it provably works with overwhelming probability only for uniform probabilistic distribution $\mathbb{D}$ over plaintexts space. The case of arbitrary $\mathbb{D}$ requires the further study. Also based on results obtained in [29] we slightly modified KPA from [25]. The obtained KPA works successful even for the number t of intercepted pairs (plaintext, ciphertext) equal to 2. This is in contrast to [25] where $t \geq d+1$ must be satisfied. But unfortunately our attack also provably recovers secret parameters with probability ≈ 1 only for uniform $\mathbb{D}$. And the case of arbitrary $\mathbb{D}$ also should be studied additionally. If $\mathbb{D}$ is such that $\Pr(0) \approx 1$ than both attack fails with probability close to 1. In future we are planning to investigate the resistance of Domingo-Ferrer homomorphic cryptosystem to ciphertext only attack.

References

[1]. R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. Communications of the ACM, 1978, vol. 21, no. 2, pp. 120–126.

[2]. S. Goldwasser and S. Micali. Probabilistic encryption & how to play mental poker keeping secret all partial information. Proceedings of the fourteenth annual ACM symposium on Theory of computing. ACM, 1982, pp. 365–377.

[3]. P. Paillier. Public-key cryptosystems based on composite degree residuosity classes. Advances in cryptologyEUROCRYPT99. Springer, 1999, pp. 223–238.

[4]. D. Boneh, E.-J. Goh, and K. Nissim. Evaluating 2-dnf formulas on ciphertexts. Theory of cryptography. Springer, 2005, pp. 325–341.

[5]. Damgård I., Jurik M. A generalisation, a simplification and some applications of Paillier's probabilistic public-key system . Public Key Cryptography. – Springer Berlin Heidelberg, 2001, pp. 119-136.

[6]. Rivest R. L., Adleman L., Dertouzos M. L. On data banks and privacy homomorphisms . Foundations of secure computation, 1978, vol. 4, no. 11, pp. 169-180.

[7]. Brickell E. F., Yacobi Y. On privacy homomorphisms . Advances in Cryptology—EUROCRYPT'87. – Springer Berlin Heidelberg, 1988, pp. 117-125.

[8]. Fellows M., Koblitz N. Combinatorial cryptosystems galore //Contemporary Mathematics, 1993, vol. 168, no. 2, pp. 51-61.

[9]. O. Zhirov, O. V. Zhirova, and S. F. Krendelev. Bezopasnye oblachnye vychisleniya s pomoshh'yu gomomorfnoj kriptografii. [Secure cloud computing using homomorphic cryptography]. Bezopasnost' informatsionnykh tekhnologij. [The security of information technologies], vol. 1, pp. 6–12, 2013 (in Russian).

[10]. J. D. i. Ferrer, A new privacy homomorphism and applications. Information Processing Letters, vol. 60, no. 5, pp. 277–282, 1996.

[11]. J. Domingo-Ferrer. A provably secure additive and multiplicative privacy homomorphism. Information Security. Springer, 2002, pp.471–483.

[12]. A. Trepacheva and L. Babenko. Known plaintexts attack on polynomial based homomorphic encryption. Proceedings of the Seventh International Conference on Security of Information and Networks. ACM, 2014.

[13]. M. R. Albrecht, P. Farshim, J.-C. Faugere, and L. Perret. Polly cracker, revisited. Advances in Cryptology–ASIACRYPT 2011. Springer, 2011, pp. 179–196.
[14]. C. Gentry. A fully homomorphic encryption scheme. Ph.D. dissertation, Stanford University, 2009.
[15]. M. Van Dijk, C. Gentry, S. Halevi, and V. Vaikuntanathan. Fully homomorphic encryption over the integers. in Advances in Cryptology–EUROCRYPT 2010. Springer, 2010, pp. 24–43.
[16]. Z. Brakerski, C. Gentry, and V. Vaikuntanathan,. (Leveled) Fully homomorphic encryption without bootstrapping. Proceedings of the 3rd Innovations in Theoretical Computer Science Conference. ACM, 2012, pp. 309–325.
[17]. N. P. Smart and F. Vercauteren. Fully homomorphic SIMD operations. Designs, Codes and Cryptography, pp. 1–25, 2011.
[18]. M. Naehrig, K. Lauter, and V. Vaikuntanathan. Can homomorphic encryption be practical?. Proceedings of the 3rd ACM workshop on Cloud computing security workshop. ACM, 2011, pp. 113–124.
[19]. L. Ertaul and J. H. Yang. Implementation of domingo ferrer's a new privacy homomorphism (df a new ph) in securing wireless sensor networks (wsn)/ Security and Management. Citeseer, 2008, pp. 498–504.
[20]. L. Ertaul, Vaidehi. Implementation of Homomorphic Encryption Schemes for Secure Packet Forwarding in Mobile Ad Hoc Networks (MANETs). IJCSNS International Journal of Computer Science and Network Security, 2007, vol. 7, no. 11 pp. 132-141.
[21]. V. Jariwala and D. Jinwala. Evaluating homomorphic encryption algorithms for privacy in wireless sensor networks. International Journal of Advancements in Computing Technology, vol. 3, no. 6, 2011.
[22]. Vaghasia and K. Bathwar.Public key encryption algorithms for wireless sensor networks in tinyos. IJITEE, 2013, vol. 2, no. 4.
[23]. Sorniotti, L. Gomez, K. Wrona, and L. Odorico. Secure and trusted in-network data processing in wireless sensor networks: a survey. Journal of Information Assurance and Security, 2007, vol. 2, no. 3, pp. 189–199.
[24]. Westhoff, J. Girao, and M. Acharya. Concealed data aggregation for reverse multicast traffic in sensor networks: Encryption, key distribution, and routing adaptation. Mobile Computing, IEEE Transactions on , 2006, vol. 5, no. 10, pp. 1417–1431.
[25]. J. H. Cheon, W.-H. Kim, and H. S. Nam. Known-plaintext cryptanalysis of the domingo-ferrer algebraic privacy homomorphism scheme. Information Processing Letters, 2006, vol. 97, no. 3, pp. 118–123.
[26]. T. Benjamin and C. D. Bennett. The probability of relatively prime polynomials. Mathematics Magazine, 2007, pp. 196–202.
[27]. Davenport, James H., Y. Siret, and E. Tournier. *Computer algebra*. London: Academic Press, 1988, 263 p.
[28]. Shoup V. NTL: A library for doing number theory. – 2001.
[29]. Wagner. Cryptanalysis of an algebraic privacy homomorphism. Information Security. Springer, 2003, pp. 234–239.

Улучшенная атака по известным открытым текстам на гомоморфную криптосистему Доминго-Феррера

А.В. Трепачева <alina1989malina@ya.ru>
Южный федеральный университет,
Россия, 344006, г. Ростов-на-Дону, ул. Большая Садовая 105/42.

Аннотация. Данная работа посвящена криптоанализу по известным открытым текстам гомоморфной криптосистемы, предложенной Доминго-Феррером. В предыдущих работах было показано, что для раскрытия секретного ключа необходимо перехватить по меньшей мере $d+1$ пару (открытый текст, шифртекст), где d – степень полиномов, являющихся шифртекстами. Здесь мы проводим анализ существующей атаки по известным открытым текстам, а также показываем, как можно её модифицировать так, чтобы значительно уменьшить нужное количество перехваченных пар. А именно, оказывается, что достаточно всего лишь двух пар для раскрытия секретного ключа. Время работы предложенной атаки так же, как и для уже существующей, зависит полиномиально от d и логарифмически от размера пространства открытых текстов. Представлены результаты компьютерных экспериментов.

Ключевые слова: атака по известным открытым текстам; гомоморфное шифрование; облачные вычисления.

Литература

[1]. R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. Communications of the ACM, 1978, vol. 21, no. 2, pp. 120–126.
[2]. S. Goldwasser and S. Micali. Probabilistic encryption & how to play mental poker keeping secret all partial information. Proceedings of the fourteenth annual ACM symposium on Theory of computing. ACM, 1982, pp. 365–377.
[3]. P. Paillier. Public-key cryptosystems based on composite degree residuosity classes. Advances in cryptologyEUROCRYPT99. Springer, 1999, pp. 223–238.
[4]. D. Boneh, E.-J. Goh, and K. Nissim. Evaluating 2-dnf formulas on ciphertexts. Theory of cryptography. Springer, 2005, pp. 325–341.
[5]. Damgård I., Jurik M. A generalisation, a simplification and some applications of Paillier's probabilistic public-key system . Public Key Cryptography. – Springer Berlin Heidelberg, 2001, pp. 119-136.

[6]. Rivest R. L., Adleman L., Dertouzos M. L. On data banks and privacy homomorphisms . Foundations of secure computation, 1978, vol. 4, no. 11, pp. 169-180.
[7]. Brickell E. F., Yacobi Y. On privacy homomorphisms . Advances in Cryptology—EUROCRYPT'87. – Springer Berlin Heidelberg, 1988, pp. 117-125.
[8]. Fellows M., Koblitz N. Combinatorial cryptosystems galore //Contemporary Mathematics, 1993, vol. 168, no. 2, pp. 51-61.
[9]. Жиров А.О., Жирова О.В., Кренделев С.Ф.. Безопасные облачные вычисления с помощью гомоморфной криптографии. Безопасность информационных технологий, 2013, Т. 1, С. 6–12.
[10]. J. D. i. Ferrer, A new privacy homomorphism and applications. Information Processing Letters, vol. 60, no. 5, pp. 277–282, 1996.
[11]. J. Domingo-Ferrer. A provably secure additive and multiplicative privacy homomorphism. Information Security. Springer, 2002, pp.471–483.
[12]. A. Trepacheva and L. Babenko. Known plaintexts attack on polynomial based homomorphic encryption. Proceedings of the Seventh International Conference on Security of Information and Networks. ACM, 2014.
[13]. M. R. Albrecht, P. Farshim, J.-C. Faugere, and L. Perret. Polly cracker, revisited. Advances in Cryptology–ASIACRYPT 2011. Springer, 2011, pp. 179–196.
[14]. C. Gentry. A fully homomorphic encryption scheme. Ph.D. dissertation, Stanford University, 2009.
[15]. M. Van Dijk, C. Gentry, S. Halevi, and V. Vaikuntanathan. Fully homomorphic encryption over the integers. in Advances in Cryptology–EUROCRYPT 2010. Springer, 2010, pp. 24–43.
[16]. Z. Brakerski, C. Gentry, and V. Vaikuntanathan,. (Leveled) Fully homomorphic encryption without bootstrapping. Proceedings of the 3rd Innovations in Theoretical Computer Science Conference. ACM, 2012, pp. 309–325.
[17]. N. P. Smart and F. Vercauteren. Fully homomorphic SIMD operations. Designs, Codes and Cryptography, pp. 1–25, 2011.
[18]. M. Naehrig, K. Lauter, and V. Vaikuntanathan. Can homomorphic encryption be practical?. Proceedings of the 3rd ACM workshop on Cloud computing security workshop. ACM, 2011, pp. 113–124.
[19]. L. Ertaul and J. H. Yang. Implementation of domingo ferrer's a new privacy homomorphism (df a new ph) in securing wireless sensor networks (wsn)/ Security and Management. Citeseer, 2008, pp. 498–504.
[20]. L. Ertaul, Vaidehi. Implementation of Homomorphic Encryption Schemes for Secure Packet Forwarding in Mobile Ad Hoc Networks (MANETs). IJCSNS International Journal of Computer Science and Network Security, 2007, vol. 7, no. 11 pp. 132-141.
[21]. V. Jariwala and D. Jinwala. Evaluating homomorphic encryption algorithms for privacy in wireless sensor networks. International Journal of Advancements in Computing Technology, vol. 3, no. 6, 2011.
[22]. Vaghasia and K. Bathwar.Public key encryption algorithms for wireless sensor networks in tinyos. IJITEE, 2013, vol. 2, no. 4.
[23]. Sorniotti, L. Gomez, K. Wrona, and L. Odorico. Secure and trusted in-network data processing in wireless sensor networks: a survey. Journal of Information Assurance and Security, 2007, vol. 2, no. 3, pp. 189–199.
[24]. Westhoff, J. Girao, and M. Acharya. Concealed data aggregation for reverse multicast traffic in sensor networks: Encryption, key distribution, and routing adaptation. Mobile Computing, IEEE Transactions on , 2006, vol. 5, no. 10, pp. 1417–1431.

[25]. J. H. Cheon, W.-H. Kim, and H. S. Nam. Known-plaintext cryptanalysis of the domingo-ferrer algebraic privacy homomorphism scheme. Information Processing Letters, 2006, vol. 97, no. 3, pp. 118–123.
[26]. T. Benjamin and C. D. Bennett. The probability of relatively prime polynomials. Mathematics Magazine, 2007, pp. 196–202
[27]. Davenport, James H., Y. Siret, and E. Tournier. *Computer algebra*. London: Academic Press, 1988, 263 p.
[28]. Shoup V. NTL: A library for doing number theory. – 2001.
[29]. Wagner. Cryptanalysis of an algebraic privacy homomorphism. Information Security. Springer, 2003, pp. 234–239.