On the Linear Complexity of the Naor-Reingold Sequence
Paula Bustillo, Domingo Gómez, Jaime Gutierrez, and Álvar Ibeas

1 Introduction

In this paper we provide a bound for the linear complexity of the so-called Naor-Reingold sequence of elements in a finite prime field. This sequence is presented in [4] as a primitive for cryptographic protocols.

For a prime $p$, we denote by $\mathbb{F}_p$ the field with $p$ elements and identify them with the integers in the range $\{0, \ldots, p-1\}$. We write $\mathbb{F}_p^*$ for the group of units in this field.

Let $p, t$ be prime numbers, $n$ a positive integer, and $g \in \mathbb{F}_p^*$ an element of multiplicative order $t$. Then, a vector $\mathbf{a} = (a_0, \ldots, a_{n-1}) \in (\mathbb{F}_t^*)^n$ defines the following finite sequence in the subgroup $\langle g \rangle$:

$$f_{\mathbf{a}}(x) := g^{\varphi_{\mathbf{a}}(x)},$$

where $\varphi_{\mathbf{a}}(x) := a_0^{x_0} \cdot \cdots \cdot a_{n-1}^{x_{n-1}} \in \mathbb{F}_t^*$ and $x = \sum_{i=0}^{n-1} x_i 2^i$ is the binary representation of $x$. Note that we have required that the order of $g$ is prime, which is not necessary for the definition of the sequence. It would be interesting to deal with this general case as well.

It has been shown that, if the decisional Diffie-Hellman assumption holds, then in general the index $x$ is not enough to compute in polynomial time $f_{\mathbf{a}}(x)$, even if an attacker performs polynomially many oracle calls (see [4, Theorem 4.1]). A bound on the discrepancy of the Naor-Reingold sequence is given in [6] and the article [3] investigates its period.

We recall that the linear complexity of an $N$-element sequence over a ring:

$$f(x), \quad x = 0, \ldots, N-1$$

is the order $L$ of the shortest linear recurrence

$$f(x+L) = a_{L-1}f(x+L-1) + \cdots + a_1f(x+1) + a_0f(x), \quad x = 0, \ldots, N-L-1.$$
2 Preliminaries

Throughout the paper the implied constants in the symbols ‘$O$’ and ‘$\gg$’ are absolute, and log denotes the binary logarithm. We state now an immediate consequence of the proof of Lemma 2 from [1]:

**Lemma 1** Let $K \subseteq \mathbb{F}_t^*$ be a set of cardinality $\#K = K$, and write

$$L_r(K, h) = \# \{ (k, y) \in K \times \mathbb{F}_t^* \mid rk = y, \ 0 \leq y < h \}.$$ 

There exists $r \in \mathbb{F}_t^*$ such that

$$L_r(K, h) \geq Kh/t.$$ 

We will also use Lemma 2 from [5].

**Lemma 2** Consider a finite sequence $(f(x))_{x=0}^{N-1}$ in a field $K$, with linear complexity $L$. Then, for any integers $M \geq 1$, $h \geq 1$, and $0 \leq e_0, \ldots, e_L \leq h$ there are some elements $c_0, \ldots, c_L \in K$ (not all zero) such that

$$L \sum_{j=0}^L c_j f(Mb + e_j) = 0$$

for any integer $b$ with $0 \leq Mk + h \leq N - 1$.

We prove now a result about the number of elements in a generic Naor-Reingold sequence. A more general result can be found in [7].

**Proposition 3** For any integers $n \geq j > 0$ and for all except at most $(3^j - 1)(t-1)^{n-1}/2$ vectors $a \in (\mathbb{F}_t^*)^n$, the Naor-Reingold sequence contains at least $2^j$ distinct elements.

**Proof.** If the sequence $f_a(x)$, $x = 0, \ldots, 2^n - 1$ contains fewer than $2^j$ values, there must be one repetition among the first $2^j$. Suppose that $f_a(x) = f_a(y)$, with

$$x = \sum_{i=0}^{j-1} x_i 2^i, \quad y = \sum_{i=0}^{j-1} y_i 2^i.$$ 

Then,

$$a_0^{x_0} \cdots a_j^{x_j} = a_0^{y_0} \cdots a_j^{y_j}. \quad (1)$$

Let $i$ be the most significant position such that $x_i \neq y_i$. Without loss of generality, we can suppose that $x_i = 1$, $y_i = 0$. Equation (1) gives

$$a_i = a_0^{y_0-x_0} \cdots a_{i-1}^{y_{i-1}-x_{i-1}}.$$ 

Once fixed the values $a_0, \ldots, a_{i-1}$ and the exponents $y_0 - x_0, \ldots, y_{i-1} - x_{i-1}$; then $a_i$ is determined as well. Therefore, there are at most $3^l(t-1)^{n-1}$ possibilities for the parameter vector $a$. Summing up all these values from the possible indices $i$, we obtain the result. ■
3 Linear complexity bound

We are ready to prove the main result of the article. The combination of the technique developed in [5, 7] with Lemma 1 yields a nontrivial result even in the case $n \sim \log t$. Furthermore, this bound improves the one provided in [5] when $n \ll (1 + \log 3) \log t$.

**Theorem 4** Let $\gamma > 0$ and $0 < \varepsilon < 1$ such that

$$n > \log t + \gamma - 4.$$  

The linear complexity $L_a$ of the sequence $(f_a(x))_{x=0}^{2^n-1}$ satisfies:

$$L_a \geq \min(2^\gamma, t^{(1-\varepsilon)/\log 3})$$

for all but at most $O(t^{n-\varepsilon})$ vectors $a \in (\mathbb{F}_t^*)^n$.

**Proof.** Take $a = (a_0, \ldots, a_{n-1})$ and split it into the following two:

$$a^- = (a_0, \ldots, a_{s-1}),$$

$$a^+ = (a_s, \ldots, a_{n-1}),$$

where

$$s = \min \left( \left\lfloor \frac{n}{2} \right\rfloor, \frac{1-\varepsilon}{\log 3} \log t \right).$$

Let $\mathcal{A}$ be the set of vectors such that each of the Naor-Reingold sequences defined by vectors $a^-$ and $a^+$ generate at least $2^s$ distinct elements. Using Proposition 3, we have that $|\mathcal{A}| = (t-1)^n + O(t^{n-\varepsilon})$.

We show that the bound holds for any vector $a \in \mathcal{A}$. Let us consider the set

$$\mathcal{K} := \{a_0^0 \cdots a_{s-1}^x \mid x_0, \ldots, x_{s-1} \in \{0, 1\} = \varphi_a[0, 2^s - 1].$$

The cardinality of this set is at least $2^s$, for $a \in \mathcal{A}$. By Lemma 1, there exists $r \in \mathbb{F}_t^*$ such that $L_r(\mathcal{K}, 2^{s-1}) \geq 2^{2s-1}/t$.

If $L_r(\mathcal{K}, 2^{s-1}) > L_a$, we choose $d_1, \ldots, d_L \in \mathcal{K}$ and such that $0 \leq y_1 := d_ir < 2^{s-1}$. Let $0 \leq e_0 < \cdots < e_{L_a} < 2^s$ be the integers such that $\varphi_a(e_i) = d_i$.

Using Lemma 2, we derive

$$\sum_{i=0}^{L_a} c_i f_a(2^s b + e_i) = 0$$

for $0 \leq b < 2^{n-s} - 1$. Let $m := r^{-1} \in \mathbb{F}_t^*$. We have that

$$f_a(2^s b + e_i) = g^{m \varphi_a(2^s) r \varphi_a(e_i)} = \left(g^{m \varphi_a(2^s b)}ight)^{y_i},$$

where $y_i$ is the integer such that $\varphi_a(e_i) = d_i$. Therefore, we can rewrite the above equation as

$$\sum_{i=0}^{L_a} c_i g^{m \varphi_a(2^s) r \varphi_a(e_i)} = 0.$$
where \( b = \sum_{j=0}^{n-s-1} b_j 2^j \). Now, as \( a \in A \), the polynomial
\[
F(X) = c_0 X^{y_0} + c_1 X^{y_1} + \ldots + c_{L_a} X^{y_{L_a}}
\]
has at least \( 2^s - 1 \) roots. This is impossible because \( \deg F \leq 2^s - 1 \). Therefore, \( L_a \geq L_r(K, 2^{s-1}) \) and the result follows.

Acknowledgments This article was partially supported by the Spanish Ministry of Science, project MTM2007-67088.

References


