Period Analysis of Pseudorandom Vector Sequences With Dynamical Polynomial Systems

Pınar Balıkçıoğlu, and Melek D. Yücel

***Abstract*—We examine the method of pseudorandom vector sequence generation proposed by Ostafe and Shparlinski, for the three choices of polynomials given by the same authors. We obtain the period distribution of the generated vector sequences for prime fields up to 13 elements and vector sizes of 2 or 3; and observe that the probability of attaining a maximum-period sequence is extremely low.**

***Index Terms*—Nonlinear pseudorandom number generators, pseudorandom vector sequences, triangular polynomial systems, period distribution**

I. Introduction

A CLASSICAL and still very popular method for the generation of uniform pseudorandom number is the linear congruential method introduced by Lehmer [1] in 1951. Unfortunately, in cryptographic settings, such linear generators have been successfully attacked. Ostafe and Shparlinski have been inspired by Lehmer's recursion in their proposal [2], for generating pseudorandom vector sequences. They study a class of dynamical systems generated by iterations of multivariate polynomials and estimate the degree growth of these iterations in [2], [3], [4] and [6]. In [5], they use this construction and design a new class of hash functions. In [7], they study the period of vector sequences generated by triangular polynomial systems and propose a method of generating maximum-period vector sequences. Our aim in this work is to analyse the distribution of the periods of the vector sequences generated by Ostafe and Shparlinski's method, which is described as follows: Let p be a prime and $F_1, ..., F_m \in \mathrm{F}_p[X_1, ..., X_m]$ be m polynomials in m variables over a finite field of p elements. For each $i = 1, ..., m$, the k-th iteration of the polynomial F_i is defined by the recurrence relation,

$$f_i^{(k+1)} = F_i(f_1^{(k)}, ..., f_m^{(k)}), \forall\ k$$

where

$$f_i^{(0)} = X_i. \tag{1}$$

To simplify the notation, one can define a vector $\mathbf{f}^{(k)} = (f_1^{(k)}, ..., f_m^{(k)}) \in \mathrm{F}_p$, $\mathbf{F} = (F_1, ..., F_m) \in \mathrm{F}_p{}^m\ [X_1, ..., X_m]$ and the recurrence relation given by (1) becomes

$$\mathbf{f}^{(k+1)} = \mathbf{F}(\mathbf{f}^{(k)})\ \forall\ k. \tag{2}$$

In particular, for any $k, n \geq 0$ and $i = 0, 1, ..., m$

$$f_i^{(k+n)} = F_i^{(k)}(\mathbf{f}^{(n)}) = F_i^{(k+n)}(\mathbf{f}^{(0)})$$

and

$$f^{(k+n)} = F^{(k)}(\mathbf{f}^{(n)}) = F^{(k+n)}(\mathbf{f}^{(0)}).$$

It is clear that the above vector sequence of vectors $f^{(k)}$ is eventually periodic with some period $\tau \leq p^m$ since we work over a finite field of p elements. On the other hand, one can assume that the vector sequence is purely periodic with period τ, that is,

$$f^{(k+\tau)} = f^{(k)}, \forall\ k.$$

In the series of papers [2]-[7] multivariate polynomial systems $F_1, ..., F_m$ of m polynomials in m variables over a finite field F_p are described in terms of the first iteration of (2), where the initial condition vector $\mathbf{f}^{(0)}$ is chosen as $\mathbf{X} = (X_1, ..., X_m)$ and the first iteration results $\mathbf{f}^{(1)} = (F_1(\mathbf{X}), ..., F_m(\mathbf{X}))$ are found as

$$F_1(\mathbf{X}) = X_1G_1(X_2, ..., X_m) + H_1(X_2, ..., X_m),$$
$$\vdots$$
$$F_{m-1}(\mathbf{X}) = X_{m-1}G_{m-1}(X_m) + H_{m-1}(X_m),$$
$$F_m(\mathbf{X}) = g_mX_m + h_m,$$

with,

$$G_i, H_i \in F_p[X_{i+1}, ..., X_m],\ i = 1, ..., m-1$$

Pınar Balıkçıoğlu is with the Cryptography Department, Institute of Applied Mathematics, Middle East Technical University, Ankara, Turkey. (pinargurkan82@gmail.com).

Melek D. Yücel is with the Cryptography Department, Institute of Applied Mathematics, Middle East Technical University, Ankara, Turkey. (melekdy@metu.edu.tr).

and

$$g_i, h_i \in F_p, g_m \neq 0, \quad (3)$$

and the following iteration proceeds by substituting the obtained vector $\mathbf{f}^{(1)} = \mathbf{F}(\mathbf{X})$ instead of $\mathbf{X}$ in (3). The structure of the polynomial description given by (3) is called the 'triangular form', because it defines the first polynomial F_1 as a function of all the elements of $\mathbf{X}$, whereas the last polynomial F_m depends on a single element X_m of $\mathbf{X}$. In order to obtain very fast pseudorandom generators, some choices for the polynomials G_i and H_i are proposed by Ostafe and Shparlinski in [2], [3] and [6], which are summarized in Table I.

TABLE I
POLYNOMIAL CHOICES PROPOSED BY OSTAFE AND SHAPARLINSKI

Choice 1 in [2]: $G_i(X_{i+1},...,X_m) = X_{i+1}$ and $H_i(X_{i+1},...,X_m) = h_i$
Choice 2 in [3]: $G_i(X_{i+1},...,X_m) = X_{i+1}^2 - a_i$ for some quadratic nonresidues a_i and $H_i(X_{i+1},...,X_m) = h_i$
Choice 3 in [6]: $G_i(X_{i+1},...,X_m) = g_i,\ g_i \notin \{0,1\}$ and $H_i(X_{i+1},...,X_m) = h_i$

TABLE II
SIZE OF PARAMETER SETS

Choice	1	2	3
$\mathbf{X_i}$	p^m	p^m	p^m
$\mathbf{g_i}$	–	–	$(p-2)^{m-1}$
$\mathbf{h_i}$	p^{m-1}	–	p^{m-1}
$\mathbf{a_i}$	–	$\bar{Q}_p^{m-1}$	–
$\mathbf{g_m}$	$p-1$	$p-1$	$p-1$
$\mathbf{h_m}$	p	p	p

II. PERIOD ANALYSIS OF GENERATED VECTOR SEQUENCES

A. Exhaustive Search for 8 Cases

Our aim is to investigate the period distribution of the vector sequences generated by (3) for the three choices of the polynomials given in Table I.

TABLE III
NUMBER OF POSSIBLE VECTOR SEQUENCES

Choice	Number of Total Vector Sequences
1	$(p-1)(p^{2m})$
2	$\bar{Q}_p^{m-1}(p-1)(p^{m+1})$
3	$(p-2)^{(m-1)}(p-1)(p^{2m})$

In order to perform an exhaustive search over all possible vector sequences, one needs to know the size of the parameter
sets. Table II shows the size of the paramater sets and Table III depicts the number of possible vector sequences. We consider 8 cases, corresponding to relatively small values of the field size p, and the vector size m (p = 3, 5, 7, 11, 13 with m = 2, and p = 3, 5, 7 with m = 3). Description of these 8 cases is given in Table IV and total numbers of possible vector sequences are shown in Table V.

B. Period Distributions

We generate all vector sequences for the 8 cases given in Table IV by Ostafe and Shparlinski's method [2], using the three polynomial choices shown in Table I. Corresponding period distributions are sketched in Fig. 1 to Fig. 8 for the three choices. In all figures, we indicate the period distribution corresponding to Choice 1 by blue, Choice 2 by red and Choice 3 by green bars. As can be observed from Fig. 1 to Fig.
8, Choice 1 and Choice 3 do not produce any maximum-period ($T = p^m$) sequence for the considered cases (p = 3, 5, 7 with m = 2, 3 and p = 11, 13 with m = 2).

TABLE IV
PARAMETERS OF THE EIGHT CASES USED IN THIS WORK

Case	p	m
1	3	2
2	3	3
3	5	2
4	5	3
5	7	2
6	7	3
7	11	2
8	13	2

TABLE V
NUMBER OF POSSIBLE VECTOR SEQUENCES FOR THE EIGHT CASES

Case	Choice 1	Choice 2	Choice 3
1	162	162	162
2	1458	1458	1458
3	2500	5000	7500
4	62500	25000	562500
5	14406	43218	72030
6	705894	6353046	17647350
7	146410	732050	1317690
8	342732	2056392	377052

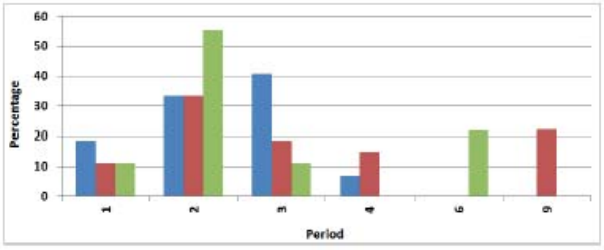

Fig. 1. Period distributions of Choice 1 (blue), Choice 2 (red), Choice 3 (green) (p=3, m=2)

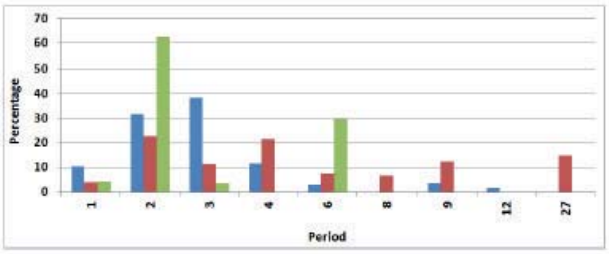

Fig. 2. Period distributions of Choice 1 (blue), Choice 2 (red), Choice 3 (green) (p=3, m=3)

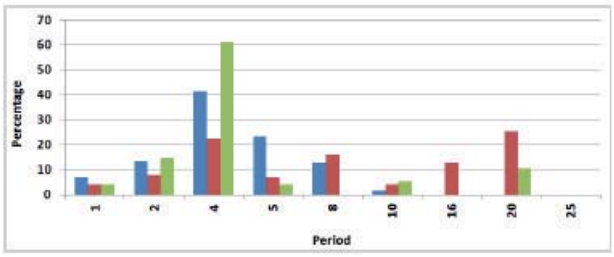

Fig. 3. Period distributions of Choice 1 (blue), Choice 2 (red), Choice 3 (green) (p=5, m=2)

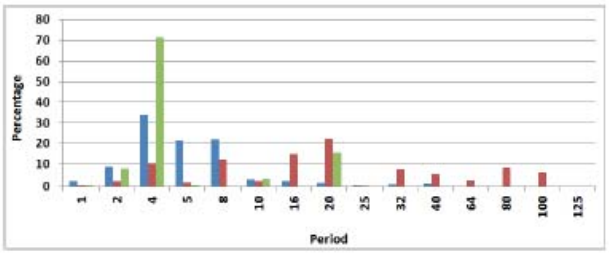

Fig. 4. Period distributions of Choice 1 (blue), Choice 2 (red), Choice 3 (green) (p=5, m=3)

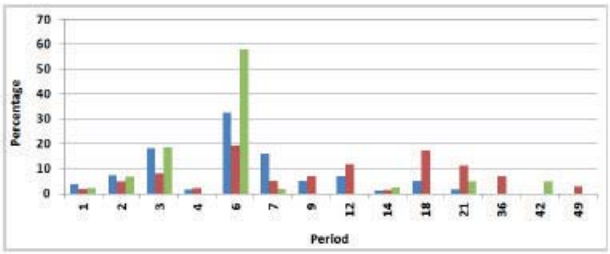

Fig. 5. Period distributions of Choice 1 (blue), Choice 2 (red), Choice 3 (green) (p=7, m=2)

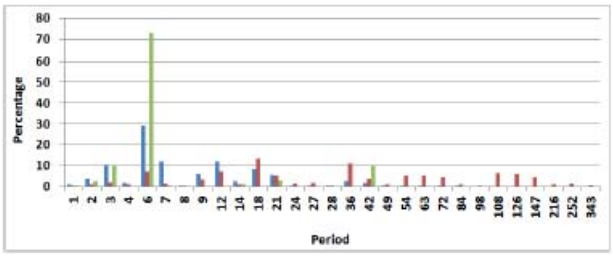

Fig. 6. Period distributions of Choice 1 (blue), Choice 2 (red), Choice 3 (green) (p=7, m=3)

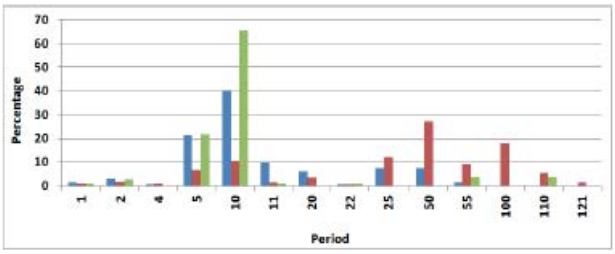

Fig. 7. Period distributions of Choice 1 (blue), Choice 2 (red), Choice 3 (green) (p=11, m=2)

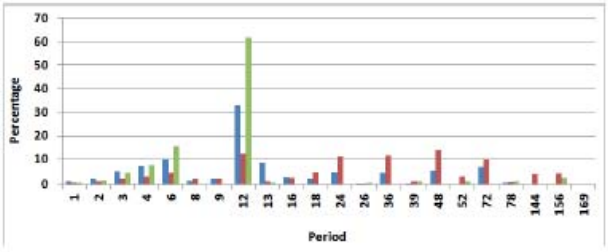

Fig. 8. Period distributions of Choice 1 (blue), Choice 2 (red), Choice 3 (green) (p=13, m=2)

Generation of sequences with maximum periods seems to be possible only with Choice 2; however, they are still not easy to find as shown in Table VI. Another general observation is that 'the sequences with period p or less' seem to be more probable than high-period sequences, especially for Choice 1 and Choice 3. Table VII shows that all three polynomial choices yield sequences with periods less than the field size p, with a high probability.

TABLE VI
PROBABILITY OF GENERATING A SEQUENCE WITH MAXIMUM PERIOD p^m

Case	Choice 1	Choice 2	Choice 3
1	0.00	0.22	0.00
2	0.00	0.15	0.00
3	0.00	0.00	0.00
4	0.00	0.00	0.00
5	0.00	0.03	0.00
6	0.00	0.01	0.00
7	0.00	0.02	0.00
8	0.00	0.00	0.00

TABLE VII
PROBABILITY OF GENERATING A SEQUENCE WITH PERIOD LESS THAN OR EQUAL TO p

Case	Choice 1	Choice 2	Choice 3
1	0.93	0.63	0.78
2	0.80	0.37	0.70
3	0.86	0.42	0.84
4	0.67	0.16	0.81
5	0.80	0.42	0.88
6	0.58	0.13	0.86
7	0.77	0.22	0.92
8	0.71	0.30	0.93

C. *Factors of Periods*

Although the generation of high-period sequences by these three polynomial choices seem to be less probable than low-period sequences, it is still of interest to examine the set of possible periods that can be generated. So, one can start by tabulating all possible periods generated in the 8 considered cases; corresponding to $m = 2$ with $p = 3, 5, 7, 11, 13$ as in Table VIII and $m = 3$ with $p = 3, 5, 7$ as in Table IX. We observe that the possible period values are found by multiplying the factors of $p-1$ with each other, or by p for all three choices of polynomials. More specifically, each period $T = ab$ shown in Table VIII is the product of two integers, where a and b can be equal to p, or to a factor of $p-1$. Choice 2 generates the largest set of period values, which almost always contains Choice 1 sets and some extra values shown as overlined. Choice 1 and Choice 3 sets differ from each other by underlined elements. Choice 3 sets are the smallest of the three; however, the bold underlined values of Choice 3 sets are unique and do not exist in other choices.

It is also of interest to find out the effect of increasing the number of polynomials (i.e., the vector size m) from 2 to 3, which can be done by examining Table IX. Each period shown in Table IX is the product of three integers, $T = abc$, where a, b and c can be equal to p, or to a factor of $p-1$. As understandable from its simple description, Choice 3 gains nothing in terms of the period, if m is raised from 2 to 3; whereas the other two choices do have larger periods as opposed to $m = 2$ case.

TABLE VIII
PERIODS OBTAINED WITH TWO POLYNOMIALS
(m=2)

Case	p	Choice 1	Choice 2	Choice 3
1	3	1, 2, 3, $\underline{4}$	1, 2, 3, 4, $\overline{9}$	1, 2, 3, $\underline{\mathbf{6}}$
3	5	1, 2, 4, 5, $\underline{8}$, 10	1, 2, 4, 5, 8, 10, $\overline{16}$, $\overline{20}$	1, 2, 4, 5, 10, $\underline{20}$
5	7	1, 2, 3, $\underline{4}$, 6, 7, $\underline{9}$, $\underline{12}$, 14, $\underline{18}$, 21	1, 2, 3, 4, 6, 7, 9, 12, 14, 18, 21, $\overline{36}$, $\overline{49}$	1, 2, 3, 6, 7, 14, 21, $\underline{\mathbf{42}}$
7	11	1, 2, $\underline{4}$, 5, 10, 11, $\underline{20}$, 22, $\underline{25}$, $\underline{50}$, 55	1, 2, 4, 5, 10, 11, 20, 22, 25, 50, 55, $\overline{100}$, $\overline{110}$, $\overline{121}$	1, 2, 5, 10, 11, 22, 55, $\underline{110}$
8	13	1, 2, 3, 4, 6, $\underline{8}$, $\underline{9}$, 12, 13, $\underline{16}$, $\underline{18}$, $\underline{24}$, 26, $\underline{36}$, 39, $\underline{48}$, $\underline{72}$, 78	1, 2, 3, 4, 6, 8, 9, 12, 13, 16, 18, 24, 26, 36, 39, 48, $\overline{52}$, 72, 78, $\overline{144}$, $\overline{156}$	1, 2, 3, 4, 6, 12, 13, 26, 39, $\underline{52}$, 78, $\underline{156}$
Maximum period		$(p-1)^2$ or $\frac{p(p-1)}{2}$	$p(p-1)$ or p^2	$p(p-1)$

TABLE IX
PERIODS OBTAINED WITH TWO POLYNOMIALS
(m=3)

Case	p	Choice 1	Choice 2	Choice 3
2	3	1, 2, 3, 4, 6, 9, 12	1, 2, 3, 4, 6, 8, 9, 27	1, 2, 3, 6
4	5	1, 2, 4, 5, 8, 10, 16, 20, 25, 32, 40	1, 2, 4, 5, 8, 10, 16, 20, 25, 32, 40, 64, 80, 100	1, 2, 4, 5, 10, 20
6	7	1, 2, 3, 4, 6, 7, 8, 9, 12, 14, 18, 21, 24, 27, 28, 36, 42, 49, 54, 63, 72, 84	1, 2, 3, 4, 6, 7, 8, 9, 12, 14, 18, 21, 24, 27, 28, 36, 42, 49, 54, 63, 72, 84, 98, 108, 126, 147, 216, 252, 343	1, 2, 3, 6, 7, 14, 21, 42
Maximum period		$p(p-1)^2$ or $\frac{p(p-1)^2}{2}$ or $\frac{p(p-1)^2}{3}$	$p^2(p-1)$ or p^3	$p(p-1)$

III. CONCLUSION

Three polynomial choices for the recursive generation method of pseudorandom vector sequences proposed by Ostafe and

Shparlinski were analysed exhaustively in terms of the period distributions and multiplying factors of the generated periods; for the specific choices of the field size p and the vector size m. Our exhaustive search for p = 3, 5, 7 with m = 2, 3 and p = 11, 13 with m = 2 shows that periods are less than p for more than half of the generated vector sequences; and there is no maximum-period sequence for the 1st and the 3rd polynomial choices given in Table I. Choice 2 is more promising since it has smaller percentage of small-period sequences than other choices; and maximum-period sequences with period p^m do exist, although their existence probability is less than 3% if $p > 3$. As an important clue for further theoretical computations**;** our search also shows that the sequence period T is equal to the product of m terms, each of which can be either equal to p or to a factor of $p-1$. Another future work is the investigation of randomness properties of the vector sequences produced by this method.

REFERENCES

[1] D. H. Lehmer, *Mathematical Methods in Large-Scale Computing Units*, in Proc. 2nd Sympos. on Large-Scale Digital Calculating Machinery, Cambridge MA, 1949, Harvard University Press, Cambridge, MA, 1951, pp. 141-146.

[2] A.Ostafe and I. E. Shparlinski, *On The Degree Growth In Some Polynomial Dynamical Systems And Non-linear Pseudorandom Number Generators*, 2010.

[3] A. Ostafe, I. E. Shparlinski E. Pelican, *On Pseudorandom Numbers From Multivariate Polynomial Systems*, 2010.

[4] A.Ostafe, *Multivariate Permutation Polynomial Systems and Non-linear Pseudorandom Number Generators*, 2010.

[5] A. Ostafe and I. E. Shparlinski, *Pseudorandom Numbers and Hash Functions from Iterations of Multivariate Polynomials*, 2010.

[6] A. Ostafe, *Pseudorandom Vector Sequences Derived from Triangular Polynomial Systems with Constant Multipliers*, 2010.

[7] A.Ostafe, *Pseudorandom Vector Sequences of Maximal Period Generated by Triangular Polynomial Dynamical Systems*, 2010.