

# Multiples of Primitive Polynomials and Their Products over $GF(2)$

A dissertation submitted in partial fulfillment  
of the requirements of M.Tech.(Computer Science)  
degree of Indian Statistical Institute, Kolkata

by

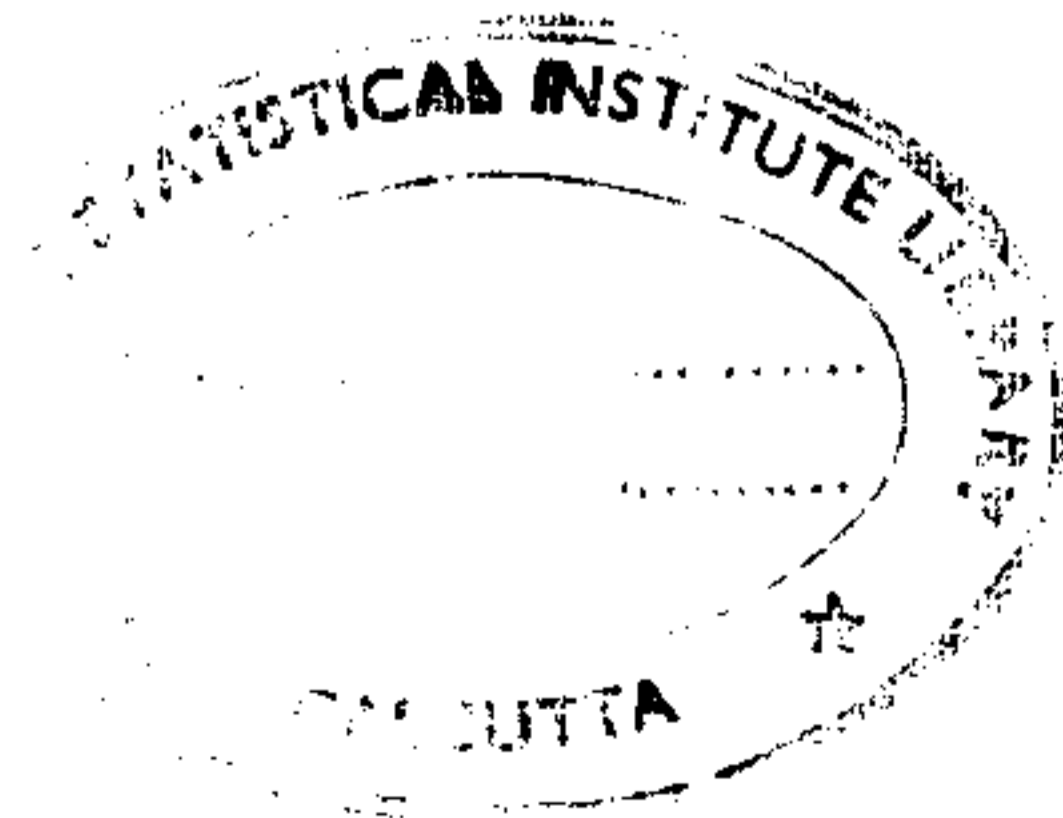
Ayineedi Venkateswarlu

under the supervision of

Dr. Subhamoy Maitra  
Applied Statistics Unit

Indian Statistical Institute  
Kolkata-700 108.

July 12, 2002



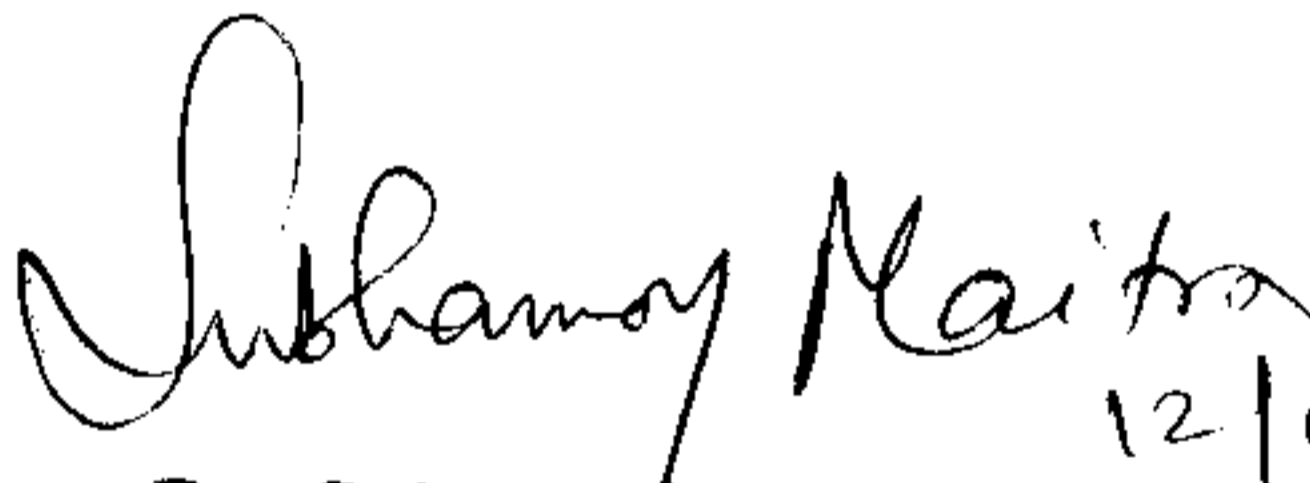
# Indian Statistical Institute

203, Barrackpore Trunk Road,

Kolkata-700 108.

## Certificate of Approval

This is to certify that this thesis titled "Multiples of Primitive Polynomials and Their Products over  $GF(2)$ " submitted by Ayineedi Venkateswarlu towards partial fulfillment of requirements for the degree of M.Tech in Computer Science at Indian Statistical Institute, Kolkata embodies the work done under my supervision.

  
12/07/2002

Dr. Subhamoy Maitra,  
Applied Statistics Unit,  
Indian Statistical Institute,  
Kolkata-700 108.



12/07/02  
( External Expert )

Dr. Sugata Gangopadhyay  
Lecturer, Mathematics Group.

BITS Pilani. 333031

(Rajasthan) .

## Acknowledgments

I take pleasure in thanking Dr. Subhomoy Maitra for his friendly guidance throughout the dissertation period. His pleasant and encouraging words have always kept my spirits up.

I would also like to express my sincere gratitude to Mr. Kishan Chand Gupta and Mr. Sandeepan Chowdhury for agreeing to discussions and help in getting material. I would like to thank members of Cryptology Research Centre, ISI-Kolkata.

Finally I take the opportunity to thank my classmates, friends and family members for their encouragement to finish this work.

Ayineedi Venkateswarlu

## Abstract

A standard model of nonlinear combiner generator for stream cipher system combines the outputs of several independent Linear Feedback Shift Register ( LFSR ) sequences using a Nonlinear Boolean Function to produce the key stream. Given such a model, cryptanalytic attacks have been proposed by finding the sparse multiples of the connection polynomials corresponding to the LFSRs. Analysis of sparse multiples of a primitive polynomial or product of primitive polynomials helps in identifying the the robustness of the steam ciphers based on nonlinear combiner model. In this direction, recently few works are published on  $t$ -nomial multiples of primitive polynomials and degree distribution of these multiples. We present new enumeration results for multiples of product of primitive polynomials and provide some results on degree distribution of these multiples. Further we provide a randomized algorithm for finding sparse multiples of primitive polynomials and their products.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Preliminaries</b>	<b>4</b>
2.1	Definitions . . . . .	4
<b>3</b>	<b><math>t</math>-nomial multiples of Product of Primitive Polynomials</b>	<b>6</b>
<b>4</b>	<b>Some Results on Degree Distribution of <math>t</math>-nomial multiples</b>	<b>10</b>
4.1	Square of degrees for trinomial multiples of primitive polynomials . . . . .	10
4.2	Degrees and square of degrees for $t$ -nomial multiples of products of primitive polynomials . . . . .	11
4.3	Reciprocal Polynomials . . . . .	14
<b>5</b>	<b>Algorithm to get Sparse Multiples</b>	<b>15</b>
5.1	Sparse Multiples of Primitive Polynomials . . . . .	15
5.2	Sparse Multiples of any Polynomial . . . . .	17
<b>6</b>	<b>Concluding Remarks :</b>	<b>18</b>

# Chapter 1

## Introduction

Linear feedback shift registers (LFSRs) are the basic components of most keystream generators since they are appropriate to hardware implementations, produce sequences with good statistical properties and can be easily analyzed.

Linear Feedback Shift Register (LFSR) is a system which generates a pseudo-random bit-sequence using a binary recurrence-relation of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \dots + c_{d-1} a_{n-d+1} + c_d a_{n-d} \quad (1.1)$$

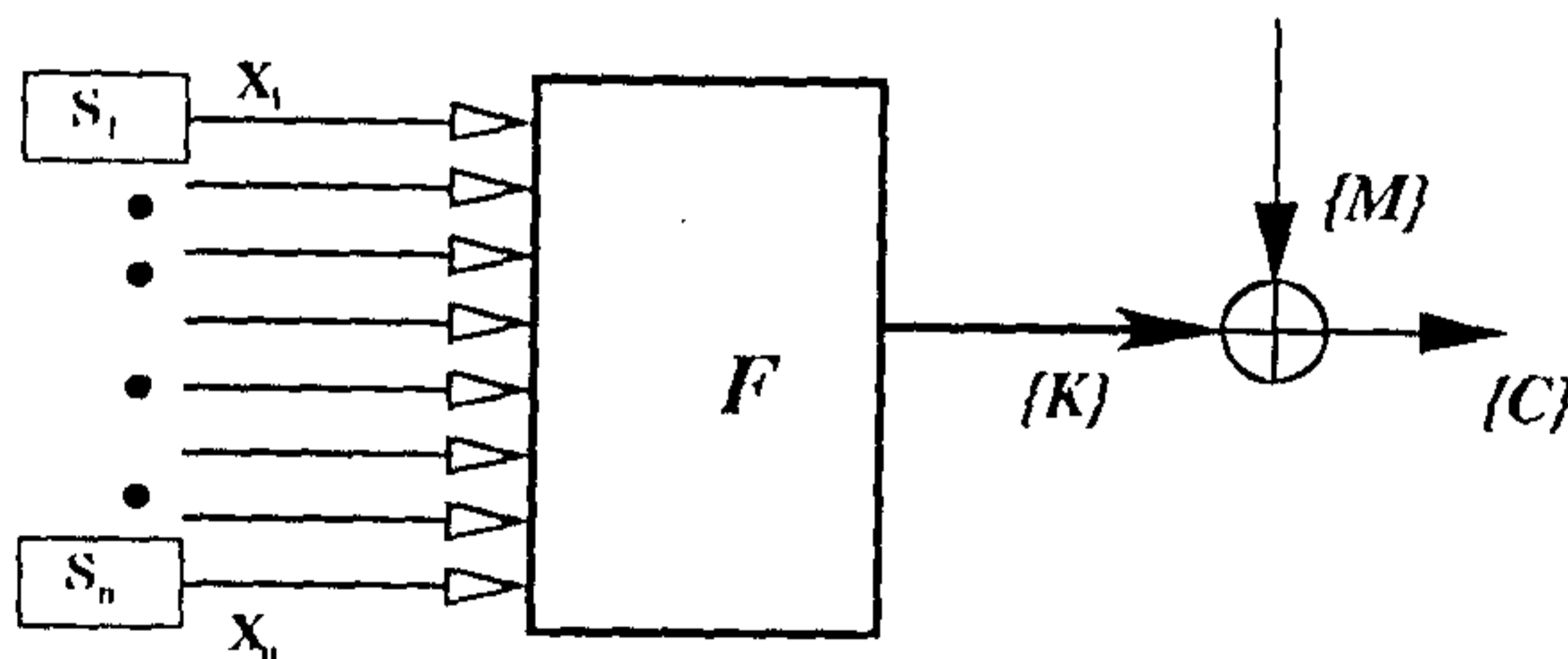
where  $c_d = 1$  and for  $1 \leq i < d$ ,  $c_i \in \{0, 1\}$ . The length of LFSR correspond to the order  $d$  of the linear-recurrence-relation used. The successive bits of the LFSR are emitted using the chosen recurrence relation after initialising the seed  $(a_0, a_1, a_2, \dots, a_{d-1})$  of LFSR.

The characteristic polynomial of 1.1 over  $GF(2)$  is

$$C(x) = 1 + c_1 x + c_2 x^2 + \dots + c_d x^d \quad (1.2)$$

This polynomial is called the *Connection Polynomial* of the LFSR. The taps of an LFSR are at the positions corresponding to  $c_i = 1$ , for  $0 \leq i < d$ .

In nonlinear combiner model of stream cipher systems,  $n$  bits from  $n$  different LFSRs ( $S_i$ ) are generated at each clock. These  $n$  bits are the input to the Boolean function  $F(X_1, X_2, X_3, \dots, X_n)$ . The output of the Boolean function  $F$  is the key-stream  $K$ . The cipher stream  $C$  is generated by XORing the key stream  $K$  and the message stream  $M$ , i.e.,  $C = K \oplus M$ . The decryption machinery is identical to the encryption machinery (see Figure below).



One very standard attack on this nonlinear combiner model is the *correlation attack*. This attack utilises the correlation between the sequence generated by the combining LFSR(s) and the running key ( $K$ ) or the cipher bits and the statistical nature of the message source. Correlation attack was first proposed by Siegenthaler [16]. This attack can be resisted by using the correlation immune Boolean function. The *fast correlation attack* proposed subsequently in [14, 1, 2, 9] reduces the complexity of the attack (by avoiding exhaustive search) and also taken care of the correlation immunity of the Boolean function. The basic outline of these *fast correlation attacks* is as follows. Let us consider  $F(X_1, \dots, X_n)$  is an  $n$ -variable,  $m$ -resilient Boolean function combining the output sequences of  $n$  LFSRs  $S_i$  having feedback polynomials  $c_i(x)$ . The Walsh transform of the Boolean function  $F$  gives,  $W_F(\bar{w}) \neq 0$  for some  $\bar{w}$  with  $wt(\bar{w}) = m + 1$ . Thus the Boolean function  $F$  and the linear function  $\bigoplus_{i=1}^n \omega_i X_i$  are correlated. Let  $\omega_{i_1} = \dots = \omega_{i_{m+1}} = 1$ . Now consider the composite LFSR  $S$  which produces the same sequence as the XOR of the sequences of the LFSRs  $S_{i_1}, \dots, S_{i_{m+1}}$ . The connection polynomial of the composite LFSR will be  $\prod_{j=1}^{m+1} c_{i_j}(x)$ . Since  $F$  and  $\bigoplus_{i=1}^n \omega_i X_i$  are correlated, the attacks target to estimate the stream generated from the composite LFSR  $S$  having the connection polynomial  $\psi(x)$ . Towards resisting this sort of correlation attacks, the connection polynomial and the Boolean function needs to satisfy certain cryptographic properties. For details about the cryptographic properties of the Boolean functions mentioned above, see [1]. Here we concentrate on the properties of the connection polynomial.

The connection polynomial is generally taken as primitive over  $GF(2)$  in order to maximise the periodicity of the sequence generated by the LFSR. The weight of (i.e.,  $\#c_i \neq 0$ ) the connection polynomial also needs to be high [14, 7]. The sequence generated by an LFSR i.e., by its linear recurrence relation (LRR), also satisfies the LRR corresponding to the multiple of the connection polynomial of the LFSR under consideration. The fast correlation attacks proposed in [14, 1] utilises this basic property. Thus it is very important to find out sparse multiples of the connection polynomial of the LFSR in order to carry out these fast correlation attacks. In terms of the practical nonlinear combiner model, we have discussed earlier, it is important to find out the sparse multiples of the connection polynomial (i.e.,  $\psi(x)$ ) of the composite LFSR  $S$ . This connection polynomial is actually product of a number of primitive polynomials. So instead of finding the multiples of a single primitive polynomial, it is more important to find out the multiples of the product of a number of primitive polynomials. Also from designer's point of consideration, the connection polynomials of the combining LFSRs of a nonlinear combiner model, should not have sparse multiples [14, 1] (see also [9] and the references). With this motivation, finding sparse multiples of primitive polynomials has received a lot of attention recently [7, 5, 6]. In this dissertation thesis we concentrate on studying different properties of the multiples of primitive polynomials and their products.

In [6] it has been shown that any primitive polynomial of degree  $d$ , has exactly  $N_{d,t} = \frac{\binom{2^d-2}{t-2} - N_{d,t-1} - \frac{t-1}{t-2}(2^d-t+1)N_{d,t-2}}{t-1}$  many  $t$ -nomial multiples (having constant term 1) with initial conditions  $N_{d,2} = N_{d,1} = 0$ . Generally the degree of the primitive polynomials are taken to be coprime for generation of key stream having better cryptographic properties [10, Page 224]. Consider  $k$  distinct primitive polynomials  $f_1(x), f_2(x), \dots, f_k(x)$  having degree  $d_1, d_2, \dots, d_k$  respectively, where  $d_1, d_2, \dots, d_k$  are pairwise coprime. Then the number of  $t$ -nomial multiples with



degree  $< (2^{d_1} - 1)(2^{d_2} - 1) \dots (2^{d_k} - 1)$  of  $f_1(x)f_2(x) \dots f_k(x)$  is at least  $((t-1)!)^{k-1} \prod_{r=1}^k N_{d_r, t}$ , where  $N_{d_r, t}$  is as defined above (see also [6]). In fact, we present a more general result, which works for product of polynomials (may not be primitive) and then as a special case, we deduce the result for products of primitive polynomials. We discuss these issues in Chapter 3.

In Chapter 4, we prove some important results related to degree of the multiples. Earlier these results were observed for small examples.

1. Consider any primitive polynomial  $f(x)$  of degree  $d$ . Consider that the degree of the trinomial multiples (having degree  $\leq 2^d - 2$ ) of  $f(x)$  are  $\hat{d}_1, \hat{d}_2, \dots, \hat{d}_{N_{d,3}}$ . Then  $\sum_{s=1}^{N_{d,3}} \hat{d}_s^2 = \frac{2}{3}(2^d - 1)(3 \cdot 2^{d-2} - 1)N_{d,3}$ .
2. Consider  $k$  distinct primitive polynomials  $f_1(x), f_2(x), \dots, f_k(x)$  with degrees  $d_1, d_2, \dots, d_k$  respectively, where  $d_1, d_2, \dots, d_k$  are pairwise coprime. It is observed [13] that the distribution of the degrees of  $t$ -nomial multiples (having constant term 1) of product of primitive polynomials is very close with the distribution of maximum of the tuples having size  $(t-1)$ . In [13] the following observations were made based on experiments.
  - (a) The average of degree of the  $t$ -nomial multiples of  $\prod_{r=1}^k f_r(x)$  is fixed and is equal to  $\frac{t-1}{t}\delta$ , where  $\delta$  is the exponent of  $\prod_{r=1}^k f_r(x)$ .
  - (b) The average of the square of degree of the trinomial multiples of  $\prod_{r=1}^k f_r(x)$  is fixed but not exactly equal to the estimated value  $\frac{2}{3}\delta(\frac{3(\delta+1)}{4} - 1)$ .

With the work of [13] and the results proved in Chapter 4, it is established more strongly that the *degree distribution of  $t$ -nomial multiples of primitive polynomials, degree distribution of  $t$ -nomial multiples of products of primitive polynomials and the distribution of maximum of  $(t-1)$  tuples* are almost indistinguishable.

In Chapter 5, we discuss a randomized algorithm to get  $t$ -nomial multiples of primitive polynomials and their products.



# Chapter 2

## Preliminaries

### 2.1 Definitions

In this section, the definitions of basic terms and with some basic results which are used in this document are provided. Most of these definitions are taken from [10]. We denote the field of prime order  $p$  by  $GF(p)$  and the extension field of dimension  $d$  over  $GF(p)$  by  $GF(p^d)$  or simply by  $GF(q)$ , where  $q = p^d$ . In the rest of the document the base field is  $GF(2)$ .

#### Galois Field of order $p^d$ :

Let  $p$  be a prime and let  $d$  be any positive integer. Then there exists a field (It is unique up to Isomorphism) of order  $p^d$ . This field is called *Galois Field of order  $p^d$*  and it is denoted by  $GF(p^d)$ .

The set  $GF(2^d)^*$  of non zero elements of  $GF(2^d)$  is a cyclic group under multiplication with a generator  $\alpha$  and  $\alpha^{p^d-1} = 1$ . Generator  $\alpha$  is called *Group Primitive Element* of  $GF(2^d)$ .  $GF(2^d) = \{0, 1, \alpha, \alpha^2, \alpha^3, \dots, \alpha^{2^d-2}\}$ .

#### Polynomial over $GF(q)$ :

A polynomial of degree  $d$  over  $GF(q)$  is of the form  $c_0 + c_1x + c_2x^2 + \dots + c_dx^d$ , where  $c_d \neq 0$  and  $c_i \in GF(q)$ , for  $0 \leq i \leq d$ .

A degree  $d$  polynomial over  $GF(2)$  is of the form  $c_0 + c_1x + c_2x^2 + \dots + c_dx^d$ , where  $c_d = 1$  and  $c_i \in \{0, 1\}$ , for  $0 \leq i < d$ .

#### Irreducible Polynomial over $GF(q)$ :

A polynomial of degree  $d$  is called an *Irreducible Polynomial over  $GF(q)$*  if it is not a product of two polynomials of degree  $< d$  over the field  $GF(q)$ .

#### Primitive Polynomial over $GF(q)$ :

Let  $f(x)$  be an irreducible polynomial of degree  $d$  over  $GF(q)$ . Then  $f(x)$  is said to be a *Primitive Polynomial of degree  $d$*  if the roots of  $f(x)$  are the generators of the field  $GF(q^d)$ . The roots of primitive polynomials having degree  $d$  are the primitive elements in  $GF(q^d)^*$ .

The number of primitive polynomials of degree  $d$  over  $GF(2)$  is  $\frac{\phi(2^d-1)}{d}$ , where  $\phi$  is an *Euler phi-function*.

**Exponent of Polynomial :**

Let  $f \in GF(q)[x]$  be a nonzero polynomial. If  $f(0) \neq 0$ , then the least positive integer  $e$  for which  $f(x)$  divides  $x^e - 1$  is called the *exponent of  $f(x)$* . If  $f(0) = 0$ , then  $f(x) = x^h g(x)$ , where  $h \in N$  and  $g \in F_q[x]$  with  $g(0) \neq 0$ , then exponent of  $f$  is defined to be exponent of  $g$ .

Exponent of a degree  $d$  polynomial over  $GF(2)$  is at most  $2^d - 1$ . Exponent of a degree  $d$  primitive polynomial over  $GF(2)$  is  $2^d - 1$ .

 **$t$ -nomial over  $GF(2)$  :**

A polynomial with  $t$  non zero terms, one of them being the constant term is called  $t$ -nomial, or in other words a polynomial of weight  $t$  with nonzero constant term.

The polynomials of the form  $1 + x^{i_1} + x^{i_2} + \dots + x^{i_{t-1}}$ , where  $1 \leq i_1 < i_2 < \dots < i_{t-1}$ , are  $t$ -nomials over  $GF(2)$ .

 **$t$ -nomial multiple of a polynomial over  $GF(2)$ :**

Let  $f \in GF(2)[x]$  be a nonzero polynomial. A  $t$ -nomial multiple of  $f$  is a  $t$ -nomial over  $GF(2)$  and is divisible by  $f$ .

We denote the number of  $t$ -nomial multiples with degree  $< e$  of  $f$  by  $N_t^f$ , where  $e$  is the exponent of  $f$ .

# Chapter 3

## $t$ -nomial multiples of Product of Primitive Polynomials

We have already discussed (see Chapter 1) the importance of finding  $t$ -nomial multiples of product of primitive polynomials instead of  $t$ -nomial multiples of just a single primitive polynomial. The attack presented in [1], uses  $t$ -nomial multiples  $t = 3, 4, 5$ . In this chapter we discuss a method of enumerating  $t$ -nomial multiples of product of primitive polynomials. In design of this model of stream cipher, generally the degree of the primitive polynomials are taken to be coprime to each other [10, Page 224] to achieve better cryptographic properties. We have taken care of this case also.

Note that in [1, Page 581], it has been assumed that the approximate count of multiples of primitive polynomials and multiples of products of primitive polynomials are close. However, this is not always true. In fact, it is possible to find products of primitive polynomials having same degree which do not have any  $t$ -nomial multiple for some  $t$ . The construction of BCH code [11] uses this idea. If the degree of the primitive polynomials are pairwise coprime, then we will show that it is always guaranteed to get  $t$ -nomial multiples of their product. Considering this, we will show that the approximate count of the multiples of a degree  $d$  primitive polynomial and a degree  $d$  polynomial which is product of some primitive polynomials each having degree  $d_r$ , i.e.,  $\sum d_r = d$  are close. So the assumption of [1, Page 581] is a good approximation. Let us now discuss the following theorem.

**Theorem 3.1** Consider  $k$  many polynomials  $f_1(x), f_2(x), \dots, f_k(x)$  over  $GF(2)$  having degrees  $d_1, d_2, \dots, d_k$  and exponents  $e_1, e_2, \dots, e_k$  respectively, with the following conditions :

1.  $e_1 \neq e_2 \neq \dots \neq e_k$  are pairwise coprime,
2.  $f_1(0) = f_2(0) = \dots = f_k(0) = 1$ ,
3.  $\gcd(f_r(x), f_s(x)) = 1$  for  $1 \leq r \neq s \leq k$ ,
4. number of  $t$ -nomial multiples (with degree  $< e_r$ ) of  $f_r(x)$  is  $n_r$ .

Then the number of  $t$ -nomial multiples with degree  $< e_1 e_2 \dots e_k$  of the product  $f_1(x) f_2(x) \dots f_k(x)$  is at least  $((t-1)!)^{k-1} n_1 n_2 \dots n_k$ .

**Proof:** Consider that any polynomial  $f_r(x)$  has a  $t$ -nomial multiple  $x^{i_{1,r}} + x^{i_{2,r}} + \dots + x^{i_{t-1,r}} + 1$  of degree  $< e_r$ . Now we try to get a  $t$ -nomial multiple of  $f_1(x)f_2(x)\dots f_k(x)$  having degree  $< e_1e_2\dots e_k$ .

Consider the set of equations  $I_1 = i_{1,r} \pmod{e_r}$ ,  $r = 1, \dots, k$ . Since  $e_1, \dots, e_k$  are pairwise coprime, we will have a unique solution of  $I_1 \pmod{e_1e_2\dots e_k}$  by Chinese remainder theorem [8, Page 53]. Similarly, consider  $I_j = i_{j,r} \pmod{e_r}$  for  $r = 1, \dots, k$  and  $j = 1, \dots, t-1$ . By Chinese remainder theorem, we get a unique solution of  $I_j \pmod{e_1e_2\dots e_k}$ .

First we like to show that  $f_r(x)$  (for  $r = 1, \dots, k$ ) divides  $x^{I_1} + x^{I_2} + \dots + x^{I_{t-1}} + 1$ . The exponent of  $f_r(x)$  is  $e_r$ . So we need to show that  $f_r(x)$  divides  $x^{I_1 \pmod{e_r}} + x^{I_2 \pmod{e_r}} + \dots + x^{I_{t-1} \pmod{e_r}} + 1$ . We have  $i_{j,r} = I_j \pmod{e_r}$  for  $r = 1, \dots, k$ ,  $j = 1, \dots, t-1$ . Thus,  $x^{I_1 \pmod{e_r}} + x^{I_2 \pmod{e_r}} + \dots + x^{I_{t-1} \pmod{e_r}} + 1$  is nothing but  $x^{i_{1,r}} + x^{i_{2,r}} + \dots + x^{i_{t-1,r}} + 1$ . Hence  $f_r(x)$  (for  $r = 1, \dots, k$ ) divides  $x^{I_1} + x^{I_2} + \dots + x^{I_{t-1}} + 1$ .

Here we need to show that  $x^{I_1} + x^{I_2} + \dots + x^{I_{t-1}} + 1$  is indeed a  $t$ -nomial, i.e.,  $I_j \not\equiv I_l \pmod{e_1\dots e_k}$  for  $j \neq l$ . If  $I_j = I_l$ , then it is easy to see that  $i_{j,r} = i_{l,r} \pmod{e_r}$  and hence,  $x^{i_{1,r}} + x^{i_{2,r}} + \dots + x^{i_{t-1,r}} + 1$  itself is not a  $t$ -nomial for any  $r$ , which is a contradiction.

Moreover, we have  $\gcd(f_r(x), f_s(x)) = 1$  for  $r \neq s$ . Thus,  $f_1(x)f_2(x)\dots f_k(x)$  divides  $x^{I_1} + x^{I_2} + \dots + x^{I_{t-1}} + 1$ . Also it is clear that degree of  $x^{I_1} + x^{I_2} + \dots + x^{I_{t-1}} + 1$  is less than  $e_1e_2\dots e_k$ .

Corresponding to the  $t$ -nomial multiple of  $f_1(x)$ , i.e.,  $x^{i_{1,1}} + x^{i_{2,1}} + \dots + x^{i_{t-1,1}} + 1$ , we fix the elements in the order  $i_{1,1}, i_{2,1}, \dots, i_{t-1,1}$ . Let us name them  $p_{1,1}, p_{2,1}, \dots, p_{t-1,1}$ .

For  $r = 2, \dots, k$ , the case is as follows. Corresponding to the  $t$ -nomial multiple  $x^{i_{1,r}} + x^{i_{2,r}} + \dots + x^{i_{t-1,r}} + 1$  of  $f_r(x)$ , we use any possible permutation of the elements  $i_{1,r}, i_{2,r}, \dots, i_{t-1,r}$  as  $p_{1,r}, p_{2,r}, \dots, p_{t-1,r}$ . Thus we will use any of the  $(t-1)!$  permutations for each  $t$ -nomial multiple of  $f_r(x)$  for  $r = 2, \dots, k$ .

Now we use Chinese remainder theorem to get  $I_j$  having value  $< e_1e_2\dots e_k$  from  $p_{j,r}$ 's for  $r = 1, \dots, k$ . Each  $p_{j,r}$  is less than  $e_r$ . Here  $p_{1,r}, p_{2,r}, \dots, p_{t-1,r}$  (related to  $f_r(x)$ ) can be permuted in  $(t-1)!$  ways and we consider the permutation related to all the  $t$ -nomials except the first one.

Corresponding to  $k$  many  $t$ -nomial multiples (one each for  $f_1(x), \dots, f_k(x)$ ), we get  $((t-1)!)^{k-1}$  many  $t$ -nomial multiples (degree  $< e_1e_2\dots e_k$ ) of the product  $f_1(x)f_2(x)\dots f_k(x)$ . Using Chinese remainder theorem, it is routine to check that all these  $((t-1)!)^{k-1}$  multiples are distinct.

Since, each  $f_r(x)$  has  $n_r$  distinct  $t$ -nomial multiples of degree  $< e_r$ , the total number of  $t$ -nomial multiples of the product  $f_1(x)f_2(x)\dots f_k(x)$  having degree  $< e_1e_2\dots e_k$  is  $((t-1)!)^{k-1}n_1n_2\dots n_k$ .

To accept the above count is a lower bound, one needs to show that the  $t$ -nomials generated by this method are all distinct. Consider two collections of  $t$ -nomial multiples  $x^{a_{1,r}} + x^{a_{2,r}} + \dots + x^{a_{t-1,r}} + 1$  and  $x^{b_{1,r}} + x^{b_{2,r}} + \dots + x^{b_{t-1,r}} + 1$  of  $f_r(x)$  for  $r = 1, \dots, k$ . There exists at least one  $s$  in the range  $1, \dots, k$  such that  $x^{a_{1,s}} + x^{a_{2,s}} + \dots + x^{a_{t-1,s}} + 1$  and  $x^{b_{1,s}} + x^{b_{2,s}} + \dots + x^{b_{t-1,s}} + 1$  are distinct. Let us consider that one of the common multiples from these two sets of  $t$ -nomials are same, say  $x^{A_{1,v}} + x^{A_{2,v}} + \dots + x^{A_{t-1,v}} + 1$  (from the set  $x^{a_{1,r}} + x^{a_{2,r}} + \dots + x^{a_{t-1,r}} + 1$ ) and  $x^{B_{1,v}} + x^{B_{2,v}} + \dots + x^{B_{t-1,v}} + 1$  (from the set  $x^{b_{1,r}} + x^{b_{2,r}} + \dots + x^{b_{t-1,r}} + 1$ ).

Without loss of generality we consider  $A_{1,v} > A_{2,v} > \dots > A_{t-1,v}$  and  $B_{1,v} > B_{2,v} > \dots > B_{t-1,v}$ . Since these two  $t$ -nomials are same, we have  $A_{j,v} = B_{j,v} \pmod{e_1e_2\dots e_k}$ . This immediately says that  $A_{j,v} = B_{j,v} \pmod{e_r}$ , which implies  $a_{j,r} = b_{j,r} \pmod{e_r}$  for each  $j$  in  $1, \dots, t-1$  and each  $r$  in  $1, \dots, k$ . This contradicts to the statement that  $x^{a_{1,s}} + x^{a_{2,s}} + \dots + x^{a_{t-1,s}} + 1$  and  $x^{b_{1,s}} + x^{b_{2,s}} + \dots + x^{b_{t-1,s}} + 1$  are distinct.



From the above point it is clear that the number of  $t$ -nomial multiples with degree  $< e_1 e_2 \dots e_k$  of  $f_1(x) f_2(x) \dots f_k(x)$  is at least  $((t-1)!)^{k-1} n_1 n_2 \dots n_k$ . ■

**Corollary 3.1** Consider  $k$  many primitive polynomials  $f_1(x), f_2(x), \dots, f_k(x)$  having degree  $d_1, d_2, \dots, d_k$  respectively, where  $d_1, d_2, \dots, d_k$  are pairwise coprime. Then the number of  $t$ -nomial multiples with degree  $< (2^{d_1} - 1)(2^{d_2} - 1) \dots (2^{d_k} - 1)$  of  $f_1(x) f_2(x) \dots f_k(x)$  is at least  $((t-1)!)^{k-1} \prod_{r=1}^k N_{d_r, t}$ , where  $N_{d_r, t}$  is as defined in introduction.

**Proof :** Since we are considering the primitive polynomials, the exponent  $e_r = 2^{d_r} - 1$ . Also, given  $d_1, d_2, \dots, d_k$  are mutually coprime,  $e_1, e_2, \dots, e_k$  are also mutually coprime. Moreover, There is no common divisor of any two primitive polynomials. The proof then follows from Theorem 3.1 putting  $n_r = N_{d_r, t}$ . ■

**Corollary 3.2** In Theorem 3.1, for  $t = 3$ , the number of trinomial multiples with degree  $< e_1 e_2 \dots e_k$  of  $f_1(x) f_2(x) \dots f_k(x)$  is exactly equal to  $2^{k-1} n_1 n_2 \dots n_k$ .

**Proof :** Consider a trinomial multiple  $x^{I_1} + x^{I_2} + 1$  having degree  $< e_1 e_2 \dots e_k$  of the product  $f_1(x) f_2(x) \dots f_k(x)$ . Since, the product  $f_1(x) f_2(x) \dots f_k(x)$  divides  $x^{I_1} + x^{I_2} + 1$ , it is clear that  $f_r(x)$  divides  $x^{I_1} + x^{I_2} + 1$ . Hence,  $f_r(x)$  divides  $x^{I_1 \bmod e_r} + x^{I_2 \bmod e_r} + 1$  having degree  $< e_r$ . Now take,  $i_{1,r} = I_1 \bmod e_r$  and  $i_{2,r} = I_2 \bmod e_r$ , for  $r = 1, \dots, k$ . It is clear that  $I_1 \not\equiv I_2 \pmod{e_r}$  (i.e.,  $i_{1,r} \neq i_{2,r}$ ), otherwise  $f_r(x)$  divides 1, which is not possible.

Also note that either  $i_{1,r}$  or  $i_{2,r}$  can not be zero, otherwise  $f_r(x)$  divides either  $x^{i_{2,r}}$  or  $x^{i_{1,r}}$ , which is not possible. Thus,  $f_r(x)$  divides  $x^{i_{1,r}} + x^{i_{2,r}} + 1$ . Then using the construction method in the proof of Theorem 3.1, one can get back  $x^{I_1} + x^{I_2} + 1$  as the multiple of  $f_1(x) f_2(x) \dots f_k(x)$  which is already considered in the count  $2^{k-1} n_1 n_2 \dots n_k$  as described in the proof of Theorem 3.1. Hence this count is exact. ■

**Corollary 3.3** Consider  $k$  many primitive polynomials  $f_1(x), f_2(x), \dots, f_k(x)$  having degree  $d_1, d_2, \dots, d_k$  respectively, where  $d_1, d_2, \dots, d_k$  are pairwise coprime. Then the number of trinomial multiples with degree  $< (2^{d_1} - 1)(2^{d_2} - 1) \dots (2^{d_k} - 1)$  of  $f_1(x) f_2(x) \dots f_k(x)$  is exactly equal to  $2^{k-1} \prod_{r=1}^k N_{d_r, 3}$ , where  $N_{d_r, 3}$  is as defined in introduction.

**Proof :** The proof follows from Corollary 3.1 and Corollary 3.2. ■

Corollary 3.2 shows that number of trinomial multiples of  $f_1(x) f_2(x) \dots f_k(x)$  is exactly  $2^{k-1} n_1 n_2 \dots n_k$ . However, it is important to mention that for  $t \geq 4$ ,  $((t-1)!)^{k-1} n_1 n_2 \dots n_k$  is indeed a lower bound and not an exact count. The reason is as follows.

Consider  $f_r(x)$  has a multiple  $x^{a_{1,r}} + x^{a_{2,r}} + \dots + x^{a_{t-1,r}} + 1$ . Note that for  $t \geq 5$ , we get  $(t-2)$ -nomial multiples of  $f_r(x)$  having degree  $< e_r$ . Consider the  $(t-2)$ -nomial multiple as  $x^{a_{1,r}} + x^{a_{2,r}} + \dots + x^{a_{t-3,r}} + 1$ . Now, from the  $(t-2)$ -nomial multiple we construct a multiple  $x^{a_{1,r}} + x^{a_{2,r}} + \dots + x^{a_{t-1,r}} + 1$ , where,  $a_{t-2,r} = a_{t-1,r} = w$ , where,  $w < e_r$ . Then if we apply Chinese remainder theorem as in Theorem 3.1, that will very well produce a  $t$ -nomial multiple of  $f_1(x) f_2(x) \dots f_k(x)$  which is not counted in Theorem 3.1. Thus the count is not exact and only a lower bound. For the case of  $t = 4$ , we can consider the multiples of the form  $x^{i_r} + x^{i_r} + 1 + 1$  of  $f_r(x)$ . These type of multiples of  $f_r(x)$ 's will contribute additional multiples of the product  $f_1(x) f_2(x) \dots f_k(x)$  which are not counted in Theorem 3.1.

**Corollary 3.4** In Theorem 3.1, for  $t \geq 4$ , the number of  $t$ -nomial multiples with degree  $< e_1 e_2 \dots e_k$  of the product  $f_1(x) f_2(x) \dots f_k(x)$  is strictly greater than  $((t-1)!)^{k-1} n_1 n_2 \dots n_k$ .

**Proof :** Proof fallows from the above discussion.

Let us consider the product of two primitive polynomials of degree 3, 4, degree 3, 5 and degree 4, 5 separately. Table 3.1 compares the lower bound given in Theorem 3.1 and the exact count by running computer program. Note that it is clear that for  $t = 3$ , the count is exact as mentioned in Corollary 3.3. On the other hand, for  $t \geq 4$ , the count is a lower bound (strictly greater than the exact count) as mentioned in Corollary 3.4. In Table 3.1, for a few cases the lower bound is zero, since  $N_{3,5} = N_{3,6} = 0$ .

Table 3.1: Count for  $t$ -nomial multiples of product of primitive polynomials.

$t$	3	4	5	6	7
Lower bound	42	672	0	0	146160
Exact count	42	1460	35945	717556	11863632

Product of degree 3, 4

$t$	3	4	5
Lower bound	90	3360	0
Exact count	90	6564	344828

Product of degree 3, 5

$t$	3	4	5
Lower bound	210	23620	1128060
Exact count	210	32508	3723686

Product of degree 4, 5

We already know that the lower bound result presented in Corollary 3.1 is invariant on the choice of the primitive polynomials. We observe that this is also true for the exact count found by computer search. As example, if one chooses any primitive polynomial of degree 3 and any one of degree 4, the exact count does not depend on the choice of the primitive polynomials.

Thus we make the following experimental observation. Consider  $k$  many primitive polynomials  $f_1(x), f_2(x), \dots, f_k(x)$  having degree  $d_1, d_2, \dots, d_k$  respectively, where  $d_1, d_2, \dots, d_k$  are pairwise coprime. Then the exact number of  $t$ -nomial multiples with degree  $< (2^{d_1} - 1)(2^{d_2} - 1) \dots (2^{d_k} - 1)$  of  $f_1(x) f_2(x) \dots f_k(x)$  is same irrespective of the choice of primitive polynomial  $f_r(x)$  of degree  $d_r$ .

And also experimental results shown that the probability of  $f_r(x)$ 's having a common  $t$ -nomial multiple with degree  $< 2^{d_r} - 1$ , for  $1 \leq r \leq k$ , is very low. So, in many cases, the degree of the lowest degree  $t$ -nomial multiple of the product  $f_1(x) f_2(x) \dots f_k(x)$  is greater than  $2^{d_r}$ , where  $d_r$  is minimum of  $\{d_1, d_2, \dots, d_k\}$ .

# Chapter 4

## Some Results on Degree Distribution of $t$ -nomial multiples

In this chapter we will discuss important results on the degrees of  $t$ -nomial multiples of primitive polynomials and their products. First we concentrate on the degrees of multiples of primitive polynomials. After that we will discuss multiples of products of primitive polynomials.

### 4.1 Square of degrees for trinomial multiples of primitive polynomials

In [6], the distribution of the degrees for the  $t$ -nomial multiples (having constant term 1) of primitive polynomials has been discussed. Given any primitive polynomial  $f(x)$  of degree  $d$ , it is clear that  $f(x)$  has  $N_{d,t}$  number of  $t$ -nomial multiples having degree  $\leq 2^d - 2$ . From cryptanalytic point of view, it is an important question that how many  $t$ -nomial multiples are there having degree less than or equal to some  $c$ . Since, this result is not settled, in [6], an estimation has been used. In [6], any  $t$ -nomial multiple  $1 + x^{i_1} + x^{i_2} + \dots + x^{i_{t-2}} + x^{i_{t-1}}$  has been interpreted as the  $(t-1)$ -tuple  $\langle i_1, i_2, \dots, i_{t-2}, i_{t-1} \rangle$ . It was also empirically justified using experimental results [6] that by fixing  $f(x)$ , if one enumerates all the  $N_{d,t}$  different  $(t-1)$  tuples, then the distribution of the tuples seems random. Moreover, the distribution of the degrees of the  $t$ -nomial multiples seems very close with the distribution of maximum value of each of the ordered tuples  $\langle i_1, i_2, \dots, i_{t-2}, i_{t-1} \rangle$  with  $1 \leq i_1 < i_2 < \dots < i_{t-2} < i_{t-1} \leq 2^d - 2$ .

To analyse the degree distribution of these  $t$ -nomial multiples, the random variate  $X^{d,t}$  is considered in [6], which is  $\max(i_1, i_2, \dots, i_{t-2}, i_{t-1})$ , where  $1 + x^{i_1} + x^{i_2} + \dots + x^{i_{t-2}} + x^{i_{t-1}}$  is a  $t$ -nomial multiple of  $f(x)$ . There are  $N_{d,t}$  such multiples. The mean value [6] of the distribution of  $X^{d,t}$  is  $\frac{t-1}{t}(2^d - 1)N_{d,t}$  divided by  $N_{d,t}$ , i.e.,  $\bar{X}^{d,t} = \frac{t-1}{t}(2^d - 1)$ . On the other hand, consider all the  $(t-1)$ -tuples  $\langle i_1, i_2, \dots, i_{t-2}, i_{t-1} \rangle$  in the range 1 to  $2^d - 2$ . There are  $\binom{2^d - 2}{t-1}$  such tuples. Each tuple is in ordered form such that  $1 \leq i_1 < i_2 < \dots < i_{t-2} < i_{t-1} \leq 2^d - 2$ . Consider the random variate  $Y^{d,t}$  which is  $\max(i_1, i_2, \dots, i_{t-2}, i_{t-1})$ . It has been shown in [6] that the mean of this distribution is  $\bar{Y}^{d,t} = \frac{t-1}{t}(2^d - 1)$ .

Thus, given any primitive polynomial  $f(x)$  of degree  $d$ , the average degree of its  $t$ -nomial multiples with degree  $\leq 2^d - 2$  is equal to the average of maximum of all the distinct  $(t-1)$  tuples form 1 to  $2^d - 2$ . With this result and experimental observations, the work of [6]



assumes that the distributions  $X^{d,t}, Y^{d,t}$  are very close. Further experimental results have been presented in [13] to strengthen the claim of [6] that the distributions  $X^{d,t}, Y^{d,t}$  are very close. In this direction, it has been shown in [13] that in terms of average of squares, the distributions  $X^{d,t}, Y^{d,t}$  are very close. The average of squares of the values in  $Y^{d,t}$  have been calculated in [13] as  $\frac{t-1}{t}(2^d - 1)(\frac{t2^d}{t+1} - 1)$  and it has been shown experimentally that the average the squares of values in  $X^{d,t}$  are very close to that of  $Y^{d,t}$ . In [13], it has been observed that for  $t = 3$ , the average of the squares of the elements of distribution  $Y^{d,3}$  and the average of the squares of the degrees of trinomial multiples (i.e., for  $X^{d,3}$ ) are same for all the experiments, which is  $\frac{2}{3}(2^d - 1)(3 \cdot 2^{d-2} - 1)$ . We theoretically prove the result here.

**Theorem 4.1** Consider any primitive polynomial  $f(x)$  of degree  $d$ . Consider that the degree of the trinomial multiples (having degree  $\leq 2^d - 2$ ) of  $f(x)$  are  $\hat{d}_1, \hat{d}_2, \dots, \hat{d}_{N_{d,3}}$ . Then  $\sum_{s=1}^{N_{d,3}} \hat{d}_s^2 = \frac{2}{3}(2^d - 1)(3 \cdot 2^{d-2} - 1)N_{d,3}$ .

**Proof :** Consider a trinomial multiple of  $f(x)$  of the form  $x^i + x^j + 1$ , where  $i > j$ . Let  $e = 2^d - 1$ . Let  $i \neq \frac{2(2^d-1)}{3}, j \neq \frac{2^d-1}{3}$ . Then  $x^{(e-i)+j} + x^{e-i} + 1$  and  $x^{e-j} + x^{i-j} + 1$  are two more distinct trinomial multiples of  $f(x)$  (multiplying  $x^i + x^j + 1$  by  $x^{e-i}$  and  $x^{e-j}$  respectively). Now, consider the difference  $(i^2 - j^2) + ((e - i + j)^2 - (e - i)^2) + ((e - j)^2 - (i - j)^2)$ , which is equal to  $e^2$ .

Further take the case  $i = \frac{2(2^d-1)}{3}, j = \frac{2^d-1}{3}$ , when  $d$  is even. In that case all the three trinomials generated in the above manner are same. Thus we will only consider one difference,  $(\frac{2(2^d-1)}{3})^2 - (\frac{2^d-1}{3})^2$ , which is equal to  $\frac{e^2}{3}$ .

Let the trinomial multiples (having degree  $< e$ ) of  $f(x)$  be  $x^{i_s} + x^{j_s} + 1$ , where  $i_s > j_s$ , for  $s = 1, \dots, N_{d,3}$ . That is  $\hat{d}_s = i_s$ . We will consider  $\sum_{s=1}^{N_{d,3}} (i_s^2 - j_s^2)$ . If  $d$  is odd we will get  $\frac{N_{d,3}}{3}$  different groups each contributing  $e^2$  in this sum. If  $d$  is even, we will get  $\frac{N_{d,3}-1}{3}$  different groups each contributing  $e^2$  in this sum except one trinomial which contributes  $\frac{e^2}{3}$  when  $i_s = \frac{2(2^d-1)}{3}, j_s = \frac{2^d-1}{3}$ .

Thus,  $\sum_{s=1}^{N_{d,3}} (i_s^2 - j_s^2) = \frac{N_{d,3}}{3}e^2$ . Now add  $\sum_{s=1}^{N_{d,3}} (i_s^2 + j_s^2)$  in both sides. Then  $2 \sum_{s=1}^{N_{d,3}} i_s^2 = \frac{N_{d,3}}{3}e^2 + \sum_{s=1}^{N_{d,3}} (i_s^2 + j_s^2)$ .

Note that, considering the values of  $i_s, j_s$  for all  $s$  we basically get all the integers in the range 1 to  $e - 1$ . Thus,  $\sum_{s=1}^{N_{d,3}} (i_s^2 + j_s^2) = 1^2 + 2^2 + \dots + (e - 1)^2$ . We already know [5] that  $N_{d,3} = 2^{d-1} - 1$ . Simplifying, we get  $\sum_{s=1}^{N_{d,3}} i_s^2 = \frac{2}{3}(2^d - 1)(3 \cdot 2^{d-2} - 1)N_{d,3}$ . ■

Theorem 4.1 proves the observation of [13]. This is now theoretically proved that for  $t = 3$ , the average of squares of the values in  $Y^{d,3}$ , i.e.,  $\frac{2}{3}(2^d - 1)(\frac{3 \cdot 2^d}{4} - 1)$  is exactly equal to the average of square of the values in  $X^{d,3}$ .

## 4.2 Degrees and square of degrees for $t$ -nomial multiples of products of primitive polynomials

Consider  $k$  many primitive polynomials  $f_1(x), f_2(x), \dots, f_k(x)$  having degrees  $d_1, d_2, \dots, d_k$  respectively. Further, the degrees are pairwise coprime. We here follow the notations of [13]. To analyse the degree distribution of these  $t$ -nomial multiples of the products of primitive polynomials, let us consider the random variate  $X^{(d_1, \dots, d_k), t}$ , which is  $\max(I_1, \dots, I_{t-1})$ , where

$x^{I_1} + x^{I_2} + \dots + x^{I_{t-1}} + 1$  is a  $t$ -nomial multiple of  $f(x) = f_1(x)f_2(x)\dots f_k(x)$ . Let  $\delta = (2^{d_1} - 1)(2^{d_2} - 1)\dots(2^{d_k} - 1)$ , the exponent of  $f(x)$ . On the other hand, consider all the  $(t-1)$ -tuples  $\langle I_1, \dots, I_{t-1} \rangle$ , in the range 1 to  $\delta - 1$ . There are  $\binom{\delta-1}{t-1}$  such tuples. Consider the random variate  $Y^{(d_1, \dots, d_k), t}$ , which is  $\max(I_1, \dots, I_{t-1})$ , where  $\langle I_1, \dots, I_{t-1} \rangle$  is any ordered  $t$ -tuple from the values 1 to  $\delta - 1$ . With some experimental results, in [13], it was mentioned that the distributions  $X^{(d_1, \dots, d_k), t}$ ,  $Y^{(d_1, \dots, d_k), t}$  are very close. Based on experimental results, the following two observations were made in [13].

1. The average of degree of the  $t$ -nomial multiples of  $\prod_{r=1}^k f_r(x)$  is fixed and is equal to  $\frac{t-1}{t}\delta$ , where  $\delta$  is the exponent of  $\prod_{r=1}^k f_r(x)$ .
2. The average of the square of degree of the trinomial multiples of  $\prod_{r=1}^k f_r(x)$  is fixed but not exactly equal to the estimated value  $\frac{2}{3}\delta(\frac{3(\delta+1)}{4} - 1)$ .

We here prove these theoretically. First we present a technical result.

**Lemma 4.1** *Let  $f(x)$  be a polynomial over  $GF(2)$  having degree  $d$  and exponent  $e$  and  $1+x$  does not divide  $f(x)$ . Let the number of  $t$ -nomial multiples (with degree  $< e$  and constant term 1) of  $f(x)$  be  $N_t^f$ . Then  $\frac{N_t^f}{t} = \frac{N_{e-t}^f}{e-t}$ , where  $2 < t < e - 2$ .*

**Proof :** Note that  $f(x)$  divides  $1+x^e$ . Since  $1+x$  does not divide  $f(x)$ ,  $f(x)$  divides  $\frac{1+x^e}{1+x}$ , i.e.,  $f(x)$  divides  $1+x+x^2+\dots+x^{e-1}$ . This is the  $e$ -nomial multiple of  $f(x)$ . Whenever  $x^{i_1} + x^{i_2} + \dots + x^{i_t}$  (constant term 0) is a multiple of  $f(x)$  (here  $1 \leq i_1 < i_2 < \dots < i_t < e$ ), adding with  $1+x+x^2+\dots+x^{e-1}$ , we will get an  $(e-t)$ -nomial multiple  $1 + \sum_{i=1, i \neq i_1, i_2, \dots, i_t}^{e-1} x^i$  (having constant term 1) of  $f(x)$ .

We will count the number of such multiples of  $f(x)$ , which is equal to the number of  $(e-t)$ -nomials. Consider a  $t$ -nomial multiple  $x^{j_1} + x^{j_2} + \dots + x^{j_{t-1}} + 1$  of  $f(x)$ . Multiplying it by  $x^j$  for  $0 \leq j < e$ , we will get  $t$  many  $t$ -nomial multiples having constant term 1 and  $(e-t)$  many multiples of the form  $x^{i_1} + x^{i_2} + \dots + x^{i_t}$ , (having constant term 0) where  $1 \leq i_1 < i_2 < \dots < i_t < e$ . Considering any one of these  $t$  many  $t$ -nomials (having constant term 1) will produce the same set of  $(e-t)$  many  $(e-t)$ -nomial multiples. So,  $t$  many  $t$ -nomials giving  $(e-t)$  many  $(e-t)$ -nomials and vice versa. Hence, we get  $\frac{N_t^f}{t} = \frac{N_{e-t}^f}{e-t}$ . ■

Let us now present the following theorem.

**Theorem 4.2** *Consider a degree  $d$  polynomial  $f(x)$  over  $GF(2)$  with exponent  $e$  such that  $(1+x)^2$  does not divide  $f(x)$ . Let the number of  $t$ -nomial multiples (with degree  $< e$  and constant term 1) of  $f$  be  $N_t^f$ . Then the sum of the degrees of all its  $t$ -nomial multiples with degree  $< e$  is  $\frac{t-1}{t}eN_t^f$ .*

**Proof :** Consider the case  $1+x$  does not divide  $f(x)$ . Consider each  $t$ -nomial multiple of degree  $\hat{d}_s$ , where  $1 \leq s \leq N_t^f$ . Now multiply each  $t$ -nomial by  $x^i$ , for  $1 \leq i \leq (e - \hat{d}_s - 1)$ , we will get multiples of the form  $x^{i_1} + x^{i_2} + \dots + x^{i_t}$ , where  $1 \leq i_1 < i_2 < \dots < i_t < e$ . Thus each  $t$ -nomial will provide  $(e - \hat{d}_s - 1)$  many multiples of the above form and observe that these are distinct. Similar to proof of Lemma 4.1,  $\sum_{s=1}^{N_t^f} (e - \hat{d}_s - 1)$  gives the count of  $(e-t)$ -nomial multiples. Moreover, from the proof of Lemma 4.1, we will get  $N_{e-t}^f = \frac{e-t}{t}N_t^f$ , i.e.,  $\sum_{s=1}^{N_t^f} (e - \hat{d}_s - 1) = \frac{e-t}{t}N_t^f$ . Hence  $\sum_{s=1}^{N_t^f} \hat{d}_s = (e - 1 - \frac{e-t}{t})N_t^f = \frac{t-1}{t}eN_t^f$ .

Now consider the case where  $f(x)$  is divisible by  $(1+x)$ , but not divisible by  $(1+x)^2$ . Since  $f(x)$  is divisible by  $(1+x)$ ,  $f(1) = 0$  and hence  $f(x)$  itself and all its multiples must contain even number of terms. That means,  $f(x)$  does not have any  $t$ -nomial multiple for odd  $t$ . Let  $f(x) = (1+x) \cdot g(x)$ . Thus, the  $t$ -nomial multiples of  $g(x)$  with even  $t$  are the only  $t$ -nomial multiples of  $f(x)$ , since  $(1+x)$  divides  $t$ -nomial multiples ( $t$  even) of  $g(x)$ . Hence, if  $(1+x)$  divides  $f(x)$ , then  $N_t^f = 0$  for  $t$  odd. Moreover,  $N_t^f = N_t^g$ , for  $t$  even. Now note that, the exponent of  $g(x)$  and  $f(x)$  are same. Thus, the sum of degrees of all the  $t$ -nomial multiples of  $f(x)$  is  $\frac{t-1}{t}eN_t^f$ , where,  $e$  is the exponent of  $f(x)$ . ■

**Corollary 4.1** Consider  $k$  many primitive polynomials  $f_1(x), f_2(x), \dots, f_k(x)$  having degrees  $d_1, d_2, \dots, d_k$  respectively (the degrees are pairwise coprime). The average of degree of the  $t$ -nomial multiples (having degree  $< \delta$ ) of  $\prod_{r=1}^k f_r(x)$  is fixed and it is equal to  $\frac{t-1}{t}\delta$ , where  $\delta$  is the exponent of  $\prod_{r=1}^k f_r(x)$ .

**Proof :** Let  $f(x) = \prod_{r=1}^k f_r(x)$ . Since each  $f_r(x)$  is a primitive polynomial of degree  $d_r$ , all the conditions of Theorem 4.2 are satisfied. Thus,  $\frac{\sum_{s=1}^{N_t^f} d_s}{N_t^f} = \frac{t-1}{t}\delta$ . ■

Hence, we prove that the average of the values in distributions  $X^{(d_1, \dots, d_k), t}$ , and  $Y^{(d_1, \dots, d_k), t}$  are same which was presented as an observation in [13]. Next we consider the square of the degrees of trinomial multiples of  $\prod_{r=1}^k f_r(x)$ .

**Theorem 4.3** Consider  $k$  many primitive polynomials  $f_1(x), f_2(x), \dots, f_k(x)$  over  $GF(2)$  having degrees  $d_1, d_2, \dots, d_k$  and exponents  $e_r = 2^{d_r} - 1$ , for  $1 \leq r \leq k$ , which are pairwise coprime. Then sum of squares of degrees of trinomial multiples of  $f(x) = f_1(x)f_2(x) \dots f_k(x)$  with degree  $< e = e_1 e_2 \dots e_k$  is

$$\frac{e^2}{12} \prod_{r=1}^k (2^{d_r} - 2) + \frac{(e-1)e(2e-1)}{12} - \frac{1}{2} \sum_{r=1}^{k-1} \sum_{A_r \subset \{e_1, e_2, \dots, e_k\}} \left[ (-1)^{r+1} \left( \prod_{e_j \in A_r} e_j^2 \right) \left( \sum_{l=1}^{\prod_{e_j \in A_r} e_j - 1} l^2 \right) \right]$$

where  $|A_r| = r$ .

**Proof :** Similar to proof of Theorem 4.1, considering all the trinomials  $x^{i_s} + x^{j_s} + 1$  of  $f(x)$  with  $1 \leq j_s < i_s < e$  for  $1 \leq s \leq N_3^f$ , we have  $2 \sum_{s=1}^{N_3^f} i_s^2 = \frac{N_3^f}{3} e^2 + \sum_{s=1}^{N_3^f} (i_s^2 + j_s^2)$ .

Now we will see the possible values for  $i_s, j_s$  in the range  $[1, e-1]$ . It is important to see that unlike the proof of Theorem 4.1, the set  $i_s, j_s$  for  $1 \leq s \leq N_3^f$  does not cover all the integers in the range  $[1, \dots, e-1]$ .

Note that,  $x^{i \bmod e_r} + x^{j \bmod e_r} + 1$  is a trinomial multiple of  $f_r(x)$ , for  $1 \leq r \leq k$  except the following case. If  $i \bmod e_r = 0$  or  $j \bmod e_r = 0$ , then  $x^{i \bmod e_r} + x^{j \bmod e_r} + 1$  is not a trinomial multiple of  $f_r(x)$ .

On the other hand, consider  $x^i + 1$ , where  $1 \leq i < e$  and  $i \not\equiv 0 \pmod{e_r}$ , for  $1 \leq r \leq k$ . Since  $f_r(x)$  is primitive polynomial, for each  $x^{i \bmod e_r} + 1$ , where  $1 \leq r \leq k$ , we will get  $x^{i \bmod e_r} + 1 \equiv x^{l_r} \pmod{f_r(x)}$ , where  $1 \leq l_r < e_r$ , i.e.,  $x^{i \bmod e_r} + x^{l_r} + 1$  is a trinomial multiple of  $f_r(x)$ . By using Chinese remainder theorem [8, Page 53], we get a unique integer  $l \bmod e$ , where  $l \equiv l_r \pmod{e_r}$ , for  $1 \leq r \leq k$ , as  $e_r$ 's are pairwise coprime.

Hence, we have to discard the cases where,  $1 \leq l < e$  and  $l \equiv 0 \pmod{e_r}$ , for any  $r$ ,  $1 \leq r \leq k$ . Then  $\sum_{s=1}^{N_3^f} (i_s^2 + j_s^2) = \sum_{i=1}^{e-1} i^2 - \sum_{x \in S} x$ , where  $S = \{l^2 : 1 \leq l < e \text{ and } l \equiv 0 \pmod{e_r}, \text{ for any } r, 1 \leq r \leq k\}$ .

Consider the sets  $S_r = \{e_r^2, (2 \cdot e_r)^2, \dots, ((\frac{e}{e_r} - 1) \cdot e_r)^2\}$ , for  $1 \leq r \leq k$ . Observe that  $\cup_{r=1}^k S_r = S$ . We now calculate  $\sum_{x \in S} x$  using inclusion and exclusion principle.

Consider  $r$  distinct integers  $n_1, n_2, \dots, n_r$  in the range  $[1, k]$ . Now we consider  $\cap_{q=1}^r S_{n_q}$ , which contains  $\prod_{q=1}^r e_{n_q}^2, 2^2 \cdot \prod_{q=1}^r e_{n_q}^2, \dots, (e / \prod_{q=1}^r e_{n_q} - 1)^2 \cdot \prod_{q=1}^r e_{n_q}^2$ . Hence,

$$\sum_{x \in \cap_{q=1}^r S_{n_q}} x = \left( \prod_{q=1}^r e_{n_q}^2 \right) \left( \sum_{l=1}^{(e / \prod_{q=1}^r e_{n_q}) - 1} l^2 \right).$$

Finally we get

$$\sum_{x \in S} x = \sum_{x \in \cup_{r=1}^k S_r} x = \sum_{r=1}^{k-1} \sum_{A_r \subset \{e_1, e_2, \dots, e_k\}} [(-1)^{r+1} \left( \prod_{e_j \in A_r} e_j^2 \right) \left( \sum_{l=1}^{(e / \prod_{e_j \in A_r} e_j) - 1} l^2 \right)],$$

where  $|A_r| = r$ . So, we have

$$2 \sum_{s=1}^{N_3^f} i_s^2 = \frac{N_3^f}{3} e^2 + \sum_{s=1}^{N_3^f} (i_s^2 + j_s^2) = \frac{N_3^f}{3} e^2 + \sum_{i=1}^{e-1} i^2 - \sum_{x \in S} x.$$

Hence,

$$\sum_{s=1}^{N_3^f} i_s^2 = \frac{N_3^f}{6} e^2 + \frac{(e-1)e(2e-1)}{12} - \frac{1}{2} \sum_{r=1}^{k-1} \sum_{A_r \subset \{e_1, e_2, \dots, e_k\}} [(-1)^{r+1} \left( \prod_{e_j \in A_r} e_j^2 \right) \left( \sum_{l=1}^{(e / \prod_{e_j \in A_r} e_j) - 1} l^2 \right)],$$

where  $|A_r| = r$ .

From [13], we have the exact formula for the number of trinomial multiples (having degree  $< e$ ) of  $f(x)$ , which is  $\frac{1}{2} \prod_{r=1}^k (2^{d_r} - 2)$  and this is the value of  $N_3^f$ . Hence the proof. ■

### 4.3 Reciprocal Polynomials

Consider two polynomials  $f(x)$  and  $g(x)$  of degree  $d$ , such that they are reciprocal to each other. Note that exponent of  $f(x)$  is equal to exponent of  $g(x)$ . Consider the multiset  $W(f(x), d, t)$ , which contains the degree of all the  $t$ -nomial multiples (having degree  $< e$ ) of polynomial  $f(x)$ . Now we have the following result.

**Lemma 4.2** *Let  $f(x)$  and  $g(x)$  be two polynomials reciprocal to each other with exponent  $e$ . Then  $W(f(x), d, t) = W(g(x), d, t)$ .*

**Proof :** Note that  $f(x)$  divides a  $t$ -nomial  $x^{i_1} + x^{i_2} + \dots + x^{i_{t-2}} + x^{i_{t-1}} + 1$  iff  $g(x)$  divides a  $t$ -nomial  $x^{i_1} + x^{i_1 - i_2} + \dots + x^{i_1 - i_{t-2}} + x^{i_1 - i_{t-1}} + 1$ . Without loss of generality, we consider that  $i_1 > i_2 > \dots > i_{t-2} > i_{t-1}$ . This gives the proof. ■



# Chapter 5

## Algorithm to get Sparse Multiples

In this chapter we try to find sparse multiples of a polynomial ( may not be primitive ) at as low degree as possible. First we discuss an algorithm to get  $t$ -nomial multiples of primitive polynomials. We also discuss implementation issues of the algorithm. Next we generalise this algorithm to get  $t$ -nomial multiples of any polynomial over GF(2).

### 5.1 Sparse Multiples of Primitive Polynomials

Consider a primitive polynomial  $f(x)$  of degree  $d$ . Let  $\alpha$  be a root of  $f(x)$ . Consider that we choose  $(t-2)$  distinct integers  $i_1, \dots, i_{t-2}$  in the range 1 to  $c$  uniformly at random where  $c < 2^d$ . It is clear, that  $1 + \alpha^{i_1} + \dots + \alpha^{i_{t-2}}$  must be equal to some  $\alpha^{i_{t-1}}$  for  $0 \leq i_{t-1} \leq 2^d - 2$ . Thus,  $1 + x^{i_1} + \dots + x^{i_{t-2}} + x^{i_{t-1}}$  will be a multiple of  $f(x)$ . Note that, if  $i_{t-1} \notin \{0, i_1, \dots, i_{t-2}\}$ , then  $1 + x^{i_1} + \dots + x^{i_{t-2}} + x^{i_{t-1}}$  will be a  $t$ -nomial multiple of  $f(x)$ . Moreover, if  $i_{t-1} \leq c$ , then we get a  $t$ -nomial multiple (of  $f(x)$ ) having degree  $\leq c$ .

**General Description of the Algorithm :**

**Algorithm 5.1 Inputs :**

- a primitive polynomial  $f(x)$  of degree  $d$ , and its root  $\alpha$ ,
  - the value  $t$ , for the  $t$ -nomial multiple,
  - an integer  $c \leq 2^d - 2$ , the maximum degree of the  $t$ -nomial multiple.
1. Choose  $(t-2)$  distinct integers  $i_1, \dots, i_{t-2}$  in the range 1 to  $c$  uniformly at random.
  2. Find out  $i_{t-1}$ , where,  $1 + \alpha^{i_1} + \dots + \alpha^{i_{t-2}} = \alpha^{i_{t-1}}$ .
  3. If  $i_{t-1} \notin \{0, i_1, \dots, i_{t-2}\}$  and  $i_{t-1} \leq c$ , then report  $1 + x^{i_1} + \dots + x^{i_{t-2}} + x^{i_{t-1}}$  and terminate<sup>1</sup>.  
Else go to step 1.

---

<sup>1</sup>Note that, if the step 3 in Algorithm 5.1 produces  $i_{t-1} \in \{0, i_1, \dots, i_{t-2}\}$ , then we get a  $(t-2)$ -nomial multiple (having degree  $\leq c$ ) of  $f(x)$ .

In [1, Page 580], similar algorithm has been discussed. In the actual implementation of the algorithm [1, Page 580], an array of length  $2^d$  is required. However, an array of length  $2^d$  is not possible to manage in practical computer systems if  $d \geq 40$ . If  $c$  is as large as  $2^d$ , the the multiple will be of very high degree and the the cryptanalytic attack will not succeed as the degree of the multiple should be of the order of (approximately half) the length of available cipher text. Thus, we need to consider  $c$  much lower than  $2^d$ . We present an algorithm, which gives  $t$ -nomial multiples, even if  $d \geq 40$ , for  $c$  much lower than  $2^d$ .

### Exact Implementation :

**Algorithm 5.2** *Inputs as in Algorithm 5.1.*

1. **Take** an array of integer  $Arr$  of length  $c + 1$  having indices 0 to  $c$ . Load the values  $\alpha^i$  ( $d$  length bit patterns interpreted as integers) in the location  $Arr[i]$  for  $0 \leq i \leq c$ . Also **take** another array of integer  $Idx$  of same length with  $Idx[i] = i$ . Sort the array  $Arr$  in ascending order and maintain the corresponding order in  $Idx$ . That is, after sorting, if  $Arr[i] = \alpha^j$ , then  $Idx[i] = j$ .
2. Choose  $(t - 2)$  distinct integers  $i_1, \dots, i_{t-2}$  in the range 1 to  $c$  uniformly at random.
3. Calculate  $\beta = 1 + \alpha^{i_1} + \dots + \alpha^{i_{t-2}}$ .
4. Use binary search to see if  $\beta$  belongs to the array  $Arr$ .
5. If  $\beta$  belongs to the array  $Arr$ , say  $Arr[j] = \beta$ , then  $i_{t-1} = Idx[j]$ . Report  $1 + x^{i_1} + \dots + x^{i_{t-2}} + x^{i_{t-1}}$  (it will either be a  $t$ -nomial or a  $(t - 2)$ -nomial) and terminate. Else go to step 2.

Note that the space required for the algorithm is dominated by  $2(c + 1)$  integers needed for the arrays  $Arr, Idx$  (see step 1) in Algorithm 5.2. The time complexity for the sorting in step 1 of Algorithm 5.2 is  $O(c \log_2 c)$ . The expected number of iterations is  $s$ , where each iteration means execution of step 2 to step 5. Among these, step 5 needs  $O(\log_2 c)$  time for binary search. Thus the time complexity is  $O(c \log_2 c + s \log_2 c) = O((c + s) \log_2 c)$ . In [17], it has shown that  $cs = 2^{d+2}$ . The time complexity is minimum when  $c = s = 2^{\frac{d}{2}+1}$ , i.e., the complexity is  $O(d2^{\frac{d}{2}})$ .

It is very clear that we are restricted in terms of available RAM in the computers. Let us explain it with an example. Currently a computer with 256 Megabytes ( $2^{28}$  bytes) is available at nominal cost. Consider that, we try to find sparse multiples of a degree 64 primitive polynomial. Now, storing each integer for the arrays  $Arr, Idx$  in step 1 of Algorithm 5.2 will need 64 bits, i.e., 8 byte space. Thus the maximum value of  $c$  is restricted by  $2 \times c \times 8 = 2^{28}$ , i.e.,  $c = 2^{24}$ . Taking  $c = 2^{24}$ , the value of  $s$  is around  $2^{42}$ , which is also computationally very high.

In fact, considering the memory requirements for the operating system and other parts of the program, the value of  $c$  will decrease further. If  $c < s$ , then the time complexity will be dominated by  $s = 2^{\frac{d}{2}+a}$ , where  $a > 1$ . In that case, the time complexity will be  $O(s \log_2 c)$ , i.e.,  $O(d2^{\frac{d}{2}+a})$ .

As example, we have considered the primitive polynomial  $x^{40} + x^{38} + x^{33} + x^{32} + x^{29} + x^{27} + x^{25} + x^{21} + x^{19} + x^{17} + x^{12} + x^{11} + x^9 + x^5 + x^3 + x^1 + 1$ . We attempted to get a 5-nomial having degree  $\leq 2^{13}$ . After 212504282 ( $< 2^{28}$ ) runs, we got the 5-nomial  $x^{7558} + x^{3297} + x^{1156} + x^{883} + 1$ . We also considered the primitive polynomial  $x^{50} + x^{48} + x^{47} + x^{44} + x^{43} + x^{40} + x^{39} + x^{36} + x^{35} + x^{34} + x^{33} + x^{32} + x^{30} + x^{28} + x^{24} + x^{23} + x^{22} + x^{19} + x^{18} + x^{17} + x^{16} + x^{15} + x^{13} + x^{12} + x^{10} + x^9 + x^8 + x^4 + 1$ . We attempted to get a 5-nomial with degree  $\leq 2^{20}$ . After 412070778 ( $< 2^{29}$ ) runs, we got the 5-nomial  $x^{913548} + x^{483148} + x^{470942} + x^{277080} + 1$ .

## 5.2 Sparse Multiples of any Polynomial

Now consider the case, where the polynomial is not primitive. Consider a polynomial  $f(x)$  over  $GF(2)$  with degree  $d$  and exponent  $e$ . We are interested in finding  $t$ -nomial multiples of  $f(x)$  with degree  $< e$ . Consider the ring of residue classes  $R = \frac{GF(2)[x]}{\langle f(x) \rangle}$  whose elements are  $g(x) + \langle f(x) \rangle$ , denoted by  $[g(x)]$ , with  $g(x) \in GF(2)[x]$ . From [10, Theorem 1.61, Page 25], it is clear that any element of  $R$  is linear combination of  $1, x, x^2, \dots, x^{d-1}$ . The zero element of  $R$  is denoted by  $[0]$ . Consider the set  $G$  consisting of  $[x]^i$ , for  $0 \leq i < e - 1$ . Clearly  $G \subset R$ . As  $e < 2^d - 1$ ,  $G$  does not contain all linear combinations of  $1, x, x^2, \dots, x^{d-1}$ . It is important to note that  $G$  is a cyclic group under multiplication modulo  $f(x)$ . Further  $G$  is not closed with respect to addition modulo  $f(x)$ . That is, it may very well happen that  $[x]^i, [x]^j \in G$ , but  $[x]^i + [x]^j = [x^i + x^j] \notin G$ , for some  $i, j \in \{0, 1, \dots, e - 1\}$ . Consider an expression  $[x]^{i_1} + [x]^{i_2} + \dots + [x]^{i_{t-1}} + 1$ , for  $e > i_1 > i_2 > \dots > i_{t-1} \geq 1$ . If this is equal to  $[0]$ , i.e.,  $[x^{i_1} + x^{i_2} + \dots + x^{i_{t-1}} + 1] = [0]$ , then  $f(x)$  divides  $x^{i_1} + x^{i_2} + \dots + x^{i_{t-1}} + 1$ . So we have a  $t$ -nomial multiple  $x^{i_1} + x^{i_2} + \dots + x^{i_{t-1}} + 1$  of  $f(x)$ .

So, we can apply the above mentioned algorithms, for getting  $t$ -nomial multiples of any arbitrary polynomial having exponent  $e$ , considering the representations of  $[x]^i$ , for  $0 \leq i < e$ , as linear combination of  $1, x, x^2, \dots, x^{d-1}$  in  $\frac{GF(2)[x]}{\langle f(x) \rangle}$ . Next refer to the algorithms presented in the last section. Consider a polynomial  $f(x)$  having degree  $d$  and exponent  $e$ . We choose  $c < e$ . Identify  $\alpha$  as  $[x]$  and  $\beta$  as  $[h(x)]$  in Algorithm 5.2. Here  $Arr$  contains the  $d$  bit representations of  $[x]^i$ . Observe that  $[h(x)]$  may not be in  $G$ , as  $G$  does not satisfy closure property with respect to addition modulo  $f(x)$ . In such case,  $t$ -nomial multiple will not be available and one has to go for next iteration.



## Chapter 6

### Concluding Remarks :

Here we have shown that the degrees of sparse multiples of product of primitive polynomials ( of reasonable degree ), in general are sufficiently large. The results in Chapter 4 shown that the degree distribution is random. Getting lower degree  $t$ -nomial multiples (  $t$  small ) is computationally very high. This conclusively establishes that sparse multiples variant of various correlation attacks on LFSR based stream cipher systems are in general infeasible requiring very long ciphertexts.

# Bibliography

- [1] A. Canteaut and M. Trabbia. Improved fast correlation attacks using parity-check equations of weight 4 and 5. In *Advances in Cryptology - EUROCRYPT 2000*, number 1807 in Lecture Notes in Computer Science, pages 573–588. Springer Verlag, 2000.
- [2] V. V. Chepyzhov, T. Johansson and B. Smeets. A Simple Algorithm for Fast Correlation Attacks on Stream Ciphers. In *Proceedings of FSE 2000*, LNCS volume 1978, 2001.
- [3] C. Ding, G. Xiao, and W. Shan. *The Stability Theory of Stream Ciphers*. Number 561 in Lecture Notes in Computer Science. Springer-Verlag, 1991.
- [4] S. W. Golomb. *Shift Register Sequences*. Aegean Park Press, 1982.
- [5] K. C. Gupta and S. Maitra. Primitive polynomials over  $GF(2)$  – A cryptologic approach. In *ICICS 2001*, number 2229 in LNCS, Pages 23–34, November 2001.
- [6] K. C. Gupta and S. Maitra. Multiples of primitive polynomials over  $GF(2)$ . INDOCRYPT 2001, number 2247 in LNCS, Pages 62–72, December 2001.
- [7] K. Jambunathan. On choice of connection polynomials for LFSR based stream ciphers. INDOCRYPT 2000, number 1977 in LNCS, Pages 9–18, 2000.
- [8] G. A. Jones and J. M. Jones. *Elementary Number Theory*. Springer Verlag London Limited, 1998.
- [9] T. Johansson and F. Jonsson. Fast correlation attacks through reconstruction of linear polynomials. In *Advances in Cryptology - CRYPTO 2000*, number 1880 in Lecture Notes in Computer Science, pages 300–315. Springer Verlag, 2000.
- [10] R. Lidl and H. Niederreiter. *Introduction to finite fields and their applications*. Cambridge University Press, 1994.
- [11] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error Correcting Codes*. North Holland, 1977.
- [12] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. *Handbook of Applied Cryptography*. CRC Press, 1997.
- [13] S. Maitra, K. C. Gupta and A. Venkateswarlu. Multiples of Primitive Polynomials and Their Products over  $GF(2)$ . In *SAC 2002*, August 2002, proceedings to be published in Lecture Notes in Computer Science.

- [14] W. Meier and O. Stafflebach. Fast correlation attacks on certain stream ciphers. *Journal of Cryptology*, 1:159-176, 1989.
- [15] T. Siegenthaler. Correlation-immunity of nonlinear combining functions for cryptographic applications. *IEEE Transactions on Information Theory*, IT-30(5):776-780, September 1984.
- [16] T. Siegenthaler. Decrypting a class of stream ciphers using cipher text only. *IEEE Transactions on Computers*, C-34(1):81-85, January 1985.
- [17] S. Maitra and A. Venkateswarlu. Further on Multiples of Primitive Polynomials and Their Products over GF(2). ( preprint 2002 ).