FEEDFORWARD FUNCTIONS DEFINED BY DE BRUIJN SEQUENCES

Z.D.Dai[1)] K.C.Zeng[2)]

[1)] Universität Karlsruhe, Institut für Algorithmen und Kognitive Systeme
on leave from the Graduate School of USTC, Academia Sinica
[2)] Data and Communications Security Research Center
Graduate School of USTC, Academia Sinica

Abstract

In this paper, we show that feedforward functions defined by de Bruijn sequences, called de Bruijn functions, satisfy some basic cryptographic requirements. It is shown how the family of de Bruijn feedforward functions could be parametrized by a key space. De Bruijn feedforward functions are balanced and complete. A lower bound of the degree of de Bruijn functions is given. A certain correlational weakness of a class of de Bruijn functions is analyzed and an algebraic method to meliorate the weakness is also given and it will not cause any substantial drawbacks withregard to other requirements. The lower bound given in this paper is by no means discouraging, yet there is hope for substantial improvements. So, improving the given lower bound is proposed as an open problem at the end of this paper.

In cryptosystems based on nonlinear feedforward functions of linear feedback shift-register sequences, one applies non-linear functions of the form

$$(a_i, a_{i+1}, \ldots, a_{i+n-1}) \longmapsto c_i = f_k(a_i, a_{i+1}, \ldots, a_{i+n-1})$$

to an LSR-sequence

$$\alpha_k = (a_0, a_1, \ldots, a_i, \ldots,), \quad a_i \in GF(2)$$

to produce the corresponding output sequences

$$\gamma_k = (c_0, c_1, \ldots, c_i, \ldots), \quad c_i \in GF(2).$$

J.J. Quisquater and J. Vandewalle (Eds.): Advances in Cryptology - EUROCRYPT '89, LNCS 434, pp. 544-548, 1990.

The function f_k, as is well known, can be expressed in an unique way as a polynomial in the indeterminates x_i, $0 \leq i \leq n-1$, linear in each of them seperately. f_k should satisfy certain cryptographic requirements. The following is a list of the simplest ones among them.

1. The function f_k with value range $\{0,1\}$ should be balanced, i.e. it will assume the value zero exactly 2^{n-1} times over the space $V_n(GF(2))$ of binary sequences of a certain length n, and complete, i.e. it will contain each of the n indeterminates explicitly.
2. It should have a reasonably high total degree, so as to guarantee a high enough linear complexity for the output sequence.
3. It should be free from certain correlational weaknesses which may influence negatively the cryptograghic strength of the output sequence.
4. The family of feedforward functions used should be parametrized by a space $V_l(GF(2))$, in the sense that to every vector k in the space there will correspond a function f_k in the family, and different k's correspond to different functions. The sequences in the space $V_l(GF(2))$ are called the keys and l the key size.

The feedforward function f_k can be regarded as a one time data base installed in the cryptosystem in accordance with the key k we select. Such a data base can be easily realized by storing the value table of the function f_k in a $2^n \times 1$ RAM.

Since de Bruijn sequences rf. [3] are binary sequences, which have periods equal to powers of 2 and can be produced in large numbers, rf. [1,2,4], it is natural to think of defining the value table of the feedforward function f by means of a de Bruijn sequence

$$\beta = (b_0, b_1, \ldots, b_i, \ldots, b_{2^n-1}; \ldots),$$

by putting

$$f(i_0, i_1, \ldots, i_{n-1}) = b_i, \quad i = \sum_{j=0}^{n-1} i_j 2^j, \qquad i_j \in \{0,1\}$$

and this function f can be called a de Bruijn feedforward function (or de Bruijn function).

First we give an approach for realizing the key-conditioned installation and alteration of the de Bruijn feedforward functions as described in requirement 4. In doing this, we make use of the terminology and notations

introduced in [1], and start with the factor Σ_f of a suitably chosen non-singular n-stage shift-regester $SR(x_n = f)$, together with a fixed reference vertex a belonging to the de Bruijn Good graph G_n. Determine for each cycle σ (not containing the reference vertex a, i.e., $a \notin \sigma$) in Σ_f the edge set $E_a(\sigma)$, write

$$r(\sigma) = \lfloor \log_2 |E_a(\sigma)| \rfloor ,$$

and fix in $E_a(\sigma)$ an arbitrary subset $E'_a(\sigma)$, consisting of $2^{r(\sigma)}$ elements. After that it will be easy to set up an one to one correspondence between binary sequences s of length

$$N = \sum_{\sigma \in \Sigma_f, a \notin \sigma} r(\sigma)$$

and spanning trees $T(s)$ of Γ_f, composed by edges belonging to the sets $E'_a(\sigma)$, exclusively.

Now the key-conditioned installation of feedforward functions can be realized in the following way. Choose the key size to be

$$l = N + n,$$

and decompose every key k of length l into a juxtaposition $k = k_1 \# k_2$ of two segments of lengths N and n respectively. Use the former to determine the feedback logic and the latter as the initial state of the de Bruijn sequence to be constructed, and store the sequence thus obtained in a RAM in the way described above, to serve as the value table of the feedforward function f_k. It is easy to see that the mapping $k \longmapsto f_k$ is one to one. It is obvious that de Bruijn functions are balanced. It is interesting that they also turn out to be complete.

Theorem 1 *If $n \geq 2$, then the function f defined by a de Bruijn sequence of degree n is complete.*

Being a balanced function in n indeterminates, the total degree of the function f defined by a de Bruijn sequence cannot exceed $n - 1$. What worries us is how low the degree of f may happen to be. The lower bound obtained in Theorem 2 below shows that the situation here is by no means discouraging.

Theorem 2 *If the function f is defined by a de Bruijn sequence of degree n, then we have*

$$\deg(f) \geq \lfloor \log_2 n \rfloor .$$

But one must keep clearly in mind that functions defined by certain classes of de Bruijn sequences may suffer from systematic correlational weaknesses, which may influence the hardness of the output sequences. We illustrate this point by analyzing the case, where the de Bruijn sequences are obtained by starting with the shift-register $SR(x_n = x_0)$. It is well known that the number of cycles in the factor Σ_{x_0} is

$$Z(n) = \sum_{d|n} \phi(n/d)2^d/n.$$

We write

$$\epsilon = 1/2 - (Z(n) - 1)/2^{n-1}$$

and state the mentioned correlational weakness in the form of the following

Theorem 3 *If the de Bruijn sequence $\beta = (b_0, b_1, \ldots, b_i, \ldots)$ is defined by a spanning tree T of the skeleton Γ_{x_0}, rf. [1], then for any integer $t > 0$ we have*

$$\text{Prob}(b_{i+tn} = b_i) = 1/2 + (2\epsilon)^{t-1}\epsilon.$$

This situation, however, is by no means hopeless. Besides choosing more appropriate de Bruijn sequences β_k, one may, for example, make use of sequences

$$\gamma_k = d(L)\beta_k$$

derived from the sequences β_k mentioned in Theorem 3 by applying a fixed operator of the form

$$d(L) = \sum_{i=0}^{p} d_i L^i, \quad d_0 = d_p = 1, \quad p < n,$$

with an odd number of nonzero coefficients d_i. This will meliorate the correlation characteristics of the feedforward functions, without causing substantial drawbacks with regard to the other requirements listed above. For we have

Theorem 4 *The functions f_k defined by the sequences γ_k are balanced and the mapping $k \longmapsto f_k$ is injective. Moreover, if*

$$2^r + p \leq n,$$

then we shall have

$$\deg(f_k) \geq \lfloor \log_2(n - p) \rfloor.$$

We conclude the present paper by making the following remarks.

1. The lower bounds to $\deg(f_k)$ given in Theorem 2 and Theorem 4 are too modest as compared with concrete results computed for a large number of de Bruijn sequences. New ideas are needed for improving these bounds.

2. Since the functions f_k have been shown to be nonlinear, it is natural to ask, to what degree they can be approximated by linear functions $l(X) = l(x_0, x_1, \ldots, x_{n-1})$ over the space $V_n(GF(2))$. To state the problem more precisely, we require to make an estimation of the quantity

$$d = \max_k \max_{l(X)} \{|\text{Prob}(f_k(a_0, \ldots, a_{n-1}) = l(a_0, \ldots, a_{n-1})) - 1/2|\}.$$

This is in fact a problem concerning the I/O correlation immunity of the family of de Bruijn functions.

References

[1] Z.D.Dai, *On the Construction and Cryptographic Application of de Bruijn Sequences*, Submitted to Journal of Cryptology.

[2] Division of Algebra in the Institute of Mathematics and Department of USTC, *On Methods of Constructing Feedback Functions of M-Sequences*, ACTA MATHEMATICAE APPLICATAE SINICA, Nov.1977 (in Chinese).

[3] H. Fredricksen, *A Survey of Full Lenth Nonlinear Shift Register Cycle Algorithms*, SIAM REVIEW, Vol.24, No.2. 1982.

[4] R.H.Xiong, *The Methods of Feedback Functions of M-Sequences I*, ACTA MATHEMATICAE APPLICATAE SINICA, Vol.9, No.2. Apr.,1986.

[5] K.C.Zeng, *On the Key Entropy Leakage Phenomena in Cryptosystem*, Report to Symposium on Cryptography, Beijing, 1986.