CORRELATION VIA LINEAR SEQUENTIAL CIRCUIT APPROXIMATION OF COMBINERS WITH MEMORY

Jovan Dj. Golić

Institute of Applied Mathematics and Electronics, Belgrade
School of Electrical Engineering, University of Belgrade
Bulevar Revolucije 73, 11001 Beograd, Yugoslavia

Abstract: Correlation properties of a general binary combiner with an arbitrary number of memory bits are analyzed. It is shown that there exists a pair of certain linear functions of the output and input, respectively, that produce correlated binary sequences. An efficient procedure, based on a linear sequential circuit approximation, is developed for finding such pairs of linear functions. The result may be a basis for a divide and conquer correlation attack on a stream cipher generator consisting of several linear feedback shift registers combined by a combiner with memory.

I. INTRODUCTION

A common way to combine several linear feedback shift registers (LFSRs) in a pseudorandom sequence generator for cryptographic or spread-spectrum applications is by a memoryless function. Siegenthaler [7] has shown that such structures are not resistant against divide and conquer correlation attacks and has introduced the corresponding concept of correlation immunity of Boolean functions [8]. This concept has been further developed in [9] using the Walsh transform of Boolean functions. According to the Xiao-Massey lemma [9] it follows that the output of a Boolean function is correlated to at least one linear function of its inputs. Given such a linear function, fast cryptanalytic algorithms for the LFSRs initial states reconstruction might also be feasible

This research was supported by the Science Fund of Serbia, grant #0403, through Institute of Mathematics, Serbian Academy of Arts and Sciences.

R.A. Rueppel (Ed.): Advances in Cryptology - EUROCRYPT '92, LNCS 658, pp. 113-123, 1993.

(for basic principles, see [2]). Their efficiency is a monotonic function of the corresponding correlation coefficient. Moreover, it has been shown in [3] that the sum of the squares of the correlation coefficients to all the linear functions of the inputs is equal to one.

In order to increase the resistance against correlation attacks one can use Boolean functions with memory, see [5], [6]. Correlation properties of combiners with one bit memory have been investigated in [4]. The sum of the squares of the correlation coefficients to all the linear functions of input sequences is derived in [4]. It follows that Boolean functions with one bit memory exist whose output is correlated to none of the linear functions of input sequences. However, in this case the sum of two successive outputs is shown to be correlated to at least one linear function of the input sequences.

In this paper, we study the correlation properties a general Boolean function with an arbitrary number, M, of bits of memory. We show that there exists a linear function of at most M+1 successive outputs that is correlated to a linear function of at most M+1 successive inputs. In addition, we propose an efficient procedure for finding such pairs of linear functions, given an arbitrary combiner with memory. The procedure is based on a linear sequential circuit approximation of a nonlinear combiner with memory.

II. GENERAL BINARY COMBINER WITH MEMORY

A general binary combiner with an arbitrary number M of memory bits is a sequential circuit defined by

$$\underset{\sim}{s}_t = G(\underset{\sim}{x}_{t-1}, \underset{\sim}{s}_{t-1}), \quad t \geq 1 \tag{1}$$

$$y_t = f(\underset{\sim}{x}_t, \underset{\sim}{s}_t), \quad t \geq 0 \tag{2}$$

where $G: GF(2)^{N+M} \longrightarrow GF(2)^M$ is a next-state vector Boolean function, $f: GF(2)^{N+M} \longrightarrow GF(2)$ is an output Boolean function, $\underset{\sim}{s}_t=(s_{1t}, \ldots, s_{Mt})$

is a state vector at time t, $\underset{\sim}{s}_0$ is an initial state, $\{\underset{\sim}{x}_t\} = \{(x_{1t}, \dots, x_{Nt})\}$ is an N-dimensional vector input sequence consisting of N binary sequences $\{x_{it}\}$, $1 \leq i \leq N$, and $\{y_t\}$ is the output binary sequence.

In order to study the correlation properties of this combiner we assume that the input sequences are mutually independent, balanced (uniformly distributed), and independent random binary sequences. A necessary condition to be satisfied for cryptographic applications is that the output sequence is also balanced and independent. It is not difficult to see that the output sequence is balanced and independent iff the output function $\underset{\sim}{f}(\underset{\sim}{x}, \underset{\sim}{s})$ is balanced for each $\underset{\sim}{s}$.

III. CORRELATION PROPERTY OF A GENERAL COMBINER WITH MEMORY

Consider an arbitrary vector Boolean function $\underset{\sim}{f}$: $GF(2)^n \longrightarrow GF(2)^m$. The vector function $\underset{\sim}{f}$ consists of m component Boolean functions, that is, $\underset{\sim}{f}=(f_1, \dots, f_m)$. Let $\underset{\sim}{z}=\underset{\sim}{f}(\underset{\sim}{x}, \underset{\sim}{y})$, $\underset{\sim}{x} \in GF(2)^{n_1}$, $\underset{\sim}{y} \in GF(2)^{n_2}$, $n=n_1+n_2$. Assume that $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are statistically independent balanced random variables. Throughout this paper, we for simplicity assume the same notation for a random variable and its values. We study the statistical dependence between $\underset{\sim}{z}$ and $\underset{\sim}{x}$. In accordance with the Xiao-Massey lemma [9], one can establish the following two properties.

Property 1: A vector Boolean function $\underset{\sim}{f}(\underset{\sim}{x})$ is balanced iff all the nonzero linear combinations of its component functions are balanced.

Property 2: A vector Boolean function $\underset{\sim}{f}(\underset{\sim}{x}, \underset{\sim}{y})$ is statistically independent of $\underset{\sim}{x}$ (respectively, is balanced for each $\underset{\sim}{x}$) iff each nonzero linear combination of the component functions of $\underset{\sim}{f}$ is statistically independent of each nonzero (is balanced for each) linear function of $\underset{\sim}{x}$.

Property 2 can be further developed by using the following simple property.

Property 3: A Boolean function $f(\underset{\sim}{x})$ and a balanced Boolean function (for example, a nonzero linear function) $g(\underset{\sim}{x})$ are statistically independent iff their sum is balanced.

Thus we obtain the following.

Property 4: A vector Boolean function $\underset{\sim}{f}(\underset{\sim}{x}, \underset{\sim}{y})$ is statistically independent of $\underset{\sim}{x}$ (respectively, balanced for each $\underset{\sim}{x}$) iff none of the nonzero linear combinations of the component functions of $\underset{\sim}{f}$ can be expressed as the sum of a nonzero linear function (linear function) of $\underset{\sim}{x}$ and a nonbalanced Boolean function of $(\underset{\sim}{x}, \underset{\sim}{y})$.

Now, consider a general binary combiner with M bits of memory described by (1) and (2). From (1) and (2) we have

$$(y_t, y_{t-1}, \ldots, y_{t-M}) = F(\underset{\sim}{x}_t, \underset{\sim}{x}_{t-1}, \ldots, \underset{\sim}{x}_{t-M}, \underset{\sim}{s}_{t-M}), \quad t \geq M \qquad (3)$$

where F is the corresponding vector Boolean function $GF(2)^{N(M+1)+M} \longrightarrow GF(2)^{M+1}$, which can be expressed as a composition of f and G. Input sequence $\{\underset{\sim}{x}_t\}_{t=0}^{\infty}$ is assumed to be balanced and independent random sequence, whereas $\underset{\sim}{s}_0$ is a given initial state. It is easy to see that the output sequence $\{\underset{\sim}{y}_t\}_{t=0}^{\infty}$ is also balanced and independent iff F is balanced for each $\underset{\sim}{s}_{t-M}$, that is, if $f(\underset{\sim}{x}, \underset{\sim}{s})$ is balanced for each $\underset{\sim}{s}$.

Using properties 3 and 4 one can prove the following result.

Theorem: Let the output function $\underset{\sim}{f}(\underset{\sim}{x}, \underset{\sim}{s})$ of a general binary combiner with M bits of memory be balanced for each $\underset{\sim}{s}$. Then, there exists a linear function $L_{\underset{\sim}{\omega}}(y_t, y_{t-1}, \ldots, y_{t-M})$ effectively depending on y_t that can be expressed as the sum of a linear function $L_{\underset{\sim}{w}}(\underset{\sim}{x}_t, \underset{\sim}{x}_{t-1}, \ldots, \underset{\sim}{x}_{t-M})$ effectively depending on $\underset{\sim}{x}_t$ and a nonbalanced Boolean function $\epsilon(\underset{\sim}{x}_t, \underset{\sim}{x}_{t-1}, \ldots, \underset{\sim}{x}_{t-M}, \underset{\sim}{s}_{t-M})$ statistically independent of $\underset{\sim}{s}_{t-M}$, for all $t \geq M$.

The main point in the proof is that the function F as a balanced (M+1)-dimensional vector Boolean function can not be balanced for each $(\underset{\sim}{x}_t, \underset{\sim}{x}_{t-1}, \ldots, \underset{\sim}{x}_{t-M})$, since $\underset{\sim}{s}_{t-M}$ has dimension only M. The theorem essentially states that for a general binary combiner with M memory bits there exists a nonzero linear function of at most M+1 successive outputs that is correlated to a nonzero linear function of at most M+1 successive inputs. Given these two linear functions, it is possible to apply the standard cryptanalytic methods [7], [2], for example, developed for memoryless combiners, in order to reconstruct the initial states of all those LFSRs involved by the linear function of the inputs. Note that for M=1, the theorem states that either the output or the sum of two successive outputs is correlated to a nonzero linear function of at most two successive inputs. This is in accordance with the results obtained in [4].

IV. LINEAR APPROXIMATION OF A GENERAL BINARY COMBINER WITH MEMORY

The correlation theorem for a general binary combiner with memory, given in the previous section, asserts the existence of a pair of certain linear functions of the output and the input that are statistically dependent, which is a basis for divide and conquer correlation attacks. However, the problem remains how to find such a pair. Exhaustive search would require balance checkings of $2^{M(N+1)}(2^N-1)$ Boolean functions of $N(M+1)$ variables, which is intractable for large M or N, even if we use the Walsh transform which requires $O((NM+N)2^{NM+M+N})$ time complexity.

We propose an efficient procedure having $O((M+1)(M+N)2^{M+N})$ time complexity at worst, which very likely leads to the desired solution. The procedure is based on the linear approximation of the output and the next-state functions of a binary combiner with memory, see Fig.1.

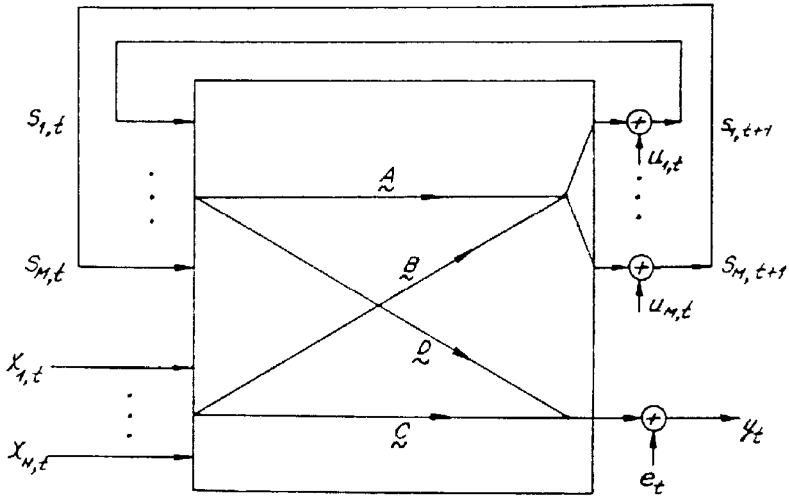

Fig.1. Linear sequential circuit approximation.

First, decompose the output function f and each of the component functions of the next-state function $G = (g_1, \ldots, g_M)$ into the sum of a linear function and a nonbalanced Boolean function. This decomposition is always possible due to the correlation properties of a memoryless combiner. If the function being decomposed is balanced, then the linear function is nonzero and, according to property 3, statistically independent of the nonbalanced function. If the function being decomposed is nonbalanced, then one can choose the linear function to be zero.

After the decomposition, the basic equations (1) and (2) using the matrix notation become

$$\underset{\sim}{s}_t = \underset{\sim}{A}\,\underset{\sim}{s}_{t-1} + \underset{\sim}{B}\,\underset{\sim}{x}_{t-1} + \underset{\sim}{u}(\underset{\sim}{x}_{t-1}, \underset{\sim}{s}_{t-1}), \quad t \geq 1 \tag{4}$$

$$y_t = \underset{\sim}{C}\,\underset{\sim}{x}_t + \underset{\sim}{D}\,\underset{\sim}{s}_t + e(\underset{\sim}{x}_t, \underset{\sim}{s}_t), \quad t \geq 0 \tag{5}$$

where the vectors are regarded as one-column matrices, $\underset{\sim}{A} = [a_{ij}]_{M\times M}$, $\underset{\sim}{B} = [b_{ij}]_{M\times N}$, $\underset{\sim}{C} = [c_{ij}]_{1\times N}$, and $\underset{\sim}{D} = [d_{ij}]_{1\times M}$ are binary matrices, and e and each of the component functions of $\underset{\sim}{u} = (u_1, \ldots, u_M)$ are nonbalanced Boolean functions.

Second, consider (4) and (5) as the basic equations of a general binary linear combiner with memory, that is, of a binary linear sequential circuit, LSC, (see [1], for example), formally assuming that $\{\underset{\sim}{u}_t\}_{t=0}^{\infty} = \{\underset{\sim}{u}(\underset{\sim}{x}_t, \underset{\sim}{s}_t)\}_{t=0}^{\infty}$ and $\{e_t\}_{t=0}^{\infty} = \{\underset{\sim}{e}(\underset{\sim}{x}_t, \underset{\sim}{s}_t)\}_{t=0}^{\infty}$ are input sequences to this LSC. Let us call this binary LSC a linear approximation of a general binary combiner with memory. Then find a solution to (4) and (5) using the generating function representation (often called the formal power series or the D-transform) of binary sequences [1]. Namely, let $\underset{\sim}{S} = \Sigma_{t=0}^{\infty} \underset{\sim}{s}_t z^t$, $\underset{\sim}{X} = \Sigma_{t=0}^{\infty} \underset{\sim}{x}_t z^t$, $\underset{\sim}{U} = \Sigma_{t=0}^{\infty} \underset{\sim}{u}_t z^t$, $E = \Sigma_{t=0}^{\infty} e_t z^t$, and $Y = \Sigma_{t=0}^{\infty} y_t z^t$ denote the corresponding generating functions. Then (4) and (5) result in

$$\underset{\sim}{S} = z\,\underset{\sim}{A}\,\underset{\sim}{S} + z\,\underset{\sim}{B}\,\underset{\sim}{X} + z\,\underset{\sim}{U} + \underset{\sim}{s}_0 \tag{6}$$

$$Y = \underset{\sim}{C}\,\underset{\sim}{X} + \underset{\sim}{D}\,\underset{\sim}{S} + E. \tag{7}$$

The solution to (6) and (7) is

$$Y = \left(\underset{\sim}{C} - z\,\frac{\underset{\sim}{D}\,(z\underset{\sim}{A} - \underset{\sim}{I})^{\mathrm{adj}}\,\underset{\sim}{B}}{\det\,(z\underset{\sim}{A} - \underset{\sim}{I})}\right)\underset{\sim}{X} - \frac{\underset{\sim}{D}\,(z\underset{\sim}{A} - \underset{\sim}{I})^{\mathrm{adj}}}{\det\,(z\underset{\sim}{A} - \underset{\sim}{I})}\,(z\underset{\sim}{U} + \underset{\sim}{s}_0) + E \tag{8}$$

where det $(z\underset{\sim}{A}-\underset{\sim}{I}) = \varphi(z)$ is a nonzero polynomial in z of degree at most rang $\underset{\sim}{A} \leq M$, which is the reciprocal of the characteristic polynomial of $\underset{\sim}{A}$, and the elements of the matrix $(z\underset{\sim}{A}-\underset{\sim}{I})^{adj}$ are polynomials in z of degree at most M-1. Accordingly, (8) can be put in the form

$$Y = \frac{1}{\varphi(z)} \sum_{i=1}^{N} h_i(z)X_i + \frac{1}{\varphi(z)} \sum_{j=1}^{M} p_j(z)(zU_j+s_{j0}) + E, \qquad (9)$$

with the polynomials in z satisfying deg $\varphi(z) \leq$ rang $\underset{\sim}{A} \leq M$, $\varphi(0)=1$, deg $h_i(z) \leq M$, $1 \leq i \leq N$, and deg $p_j(z) \leq M-1$, $1 \leq j \leq M$, which is equivalent to

$$\varphi(z)Y = \sum_{i=1}^{N} h_i(z)X_i + \sum_{j=1}^{M} p_j(z)(zU_j+s_{j0}) + \varphi(z)E. \qquad (10)$$

Letting $\varphi(z)= \sum_{k=0}^{M} \varphi_k z^k$, $h_i(z)= \sum_{k=0}^{M} h_{ik} z^k$, $1 \leq i \leq N$, $p_j(z)= \sum_{k=0}^{M-1} p_{jk} z^k$, $1 \leq j \leq M$, (10) in time domain reduces to

$$\sum_{k=0}^{M} \varphi_k \, y_{t-k} = \sum_{i=1}^{N} \sum_{k=0}^{M} h_{ik} \, x_{i,t-k} + \sum_{j=1}^{M} \sum_{k=0}^{M-1} p_{jk} \, u_j(\underset{\sim}{x}_{t-1-k}, \underset{\sim}{s}_{t-1-k}) + $$
$$+ \sum_{k=0}^{M} \varphi_k \, e(\underset{\sim}{x}_{t-k}, \underset{\sim}{s}_{t-k}), \quad t \geq M. \qquad (11)$$

Now, if by using the next-state function, $\underset{\sim}{s}_{t-k}$ is expressed in terms of $(\underset{\sim}{x}_{t-k-1}, \dots, \underset{\sim}{x}_{t-M}, \underset{\sim}{s}_{t-M})$, $0 \leq k \leq M$, (11) reduces to

$$\sum_{k=0}^{M} \varphi_k \, y_{t-k} = \sum_{i=1}^{N} \sum_{k=0}^{M} h_{ik} \, x_{i,t-k} + \varepsilon(\underset{\sim}{x}_t, \dots, \underset{\sim}{x}_{t-M}, \underset{\sim}{s}_{t-M}), \quad t \geq M, \qquad (12)$$

which is the desired composition. However, it seems impossible to prove in general that the noise function ε is a nonbalanced function. Namely, $u_j(\underset{\sim}{x}_{t-1-k}, \underset{\sim}{s}_{t-1-k})$, $1 \leq j \leq M$, $0 \leq k \leq M-1$, and

$e(\underset{\sim}{x}_{t-k}, \underset{\sim}{s}_{t-k})$, $0 \leq k \leq M$, are nonbalanced provided that $\underset{\sim}{s}_{t-k}$ is balanced, $0 \leq k \leq M$. If $\underset{\sim}{s}_{t-k}$ is expressed in terms of $(\underset{\sim}{x}_{t-k-1},\dots,\underset{\sim}{x}_{t-M}, \underset{\sim}{s}_{t-M})$, $0 \leq k \leq M$, it might not be balanced any more, since we then assume that $\underset{\sim}{s}_{t-M}$ is balanced. Therefore, it is not impossible that u_j as a function of $(\underset{\sim}{x}_{t-1-k},\dots, \underset{\sim}{x}_{t-M}, \underset{\sim}{s}_{t-M})$, $1 \leq j \leq M$, $0 \leq k \leq M-1$, or e as a function of $(\underset{\sim}{x}_{t-k},\dots, \underset{\sim}{x}_{t-M}, \underset{\sim}{s}_{t-M})$, $0 \leq k \leq M$, become balanced. On the other hand, it is also not impossible that $\epsilon(\underset{\sim}{x}_{t},\dots, \underset{\sim}{x}_{t-M}, \underset{\sim}{s}_{t-M})$ as the sum of nonbalanced, not necessarily statistically independent, functions becomes balanced. Nevertheless, it appears highly unlikely that $\epsilon(\underset{\sim}{x}_{t},\dots, \underset{\sim}{x}_{t-M}, \underset{\sim}{s}_{t-M})$ is balanced.

Note that (9) is a basic equation which can be further modified. For example, one can remove from $\varphi(z)$ all the factors that are common to all $p_j(z)$, $1 \leq j \leq M$, and $h_i(z)$, $1 \leq i \leq N$, provided they exist, and thus obtain and use $\varphi'(z)$.

V. CORRELATION ATTACK

We now analyze the discrimination potential of (11), that is, (12) regarding the statistical reconstruction of the input sequences, provided that the noise function ϵ is nonbalanced. Suppose that each of the input sequences $\{x_{it}\}$ is generated by a LFSR with feedback polynomial $F_i(z)$, $1 \leq i \leq N$. Also suppose that $F_i(z)$, $1 \leq i \leq N$, are pairwise coprime. Then, using (12) one can statistically reconstruct all the input sequences $\{x_{it}\}$, $1 \leq i \leq N$, such that $h_{ik} \neq 0$ for at least one k, $0 \leq k \leq M$, and that $F_i(z)$ does not divide $h_i(z)$. It is important to note that for the statistical procedure one need not know how much ϵ is nonbalanced, although the knowledge of the overall correlation coefficient, that is, the correlation coefficient between ϵ and the zero function would enable the calculation of the necessary length of the observed output segment. However, one may also take a modified approach that leads to the reconstruction of individual input sequences. Namely, if we multiply both sides of (10) by the product $F_i'(z)$, of degree L_i', of $F_j(z)$ for all $j\neq i$ that effectively appear in (12), then (10) reduces to

$$F_i'(z)\varphi(z)Y = F_i'(z)h_i(z)X_i + F_i'(z)\sum_{j=1}^{M}p_j(z)(zU_j+s_{j0}) + F_i'(z)\varphi(z)E + \\ + \Delta_i(z). \qquad (13)$$

with $\Delta_i(z)$ being a polynomial of degree at most $M+L_i'-1$, which results in a decomposition equation analogous to (12), with $M_i'=M+L_i')M$ instead of M. If the resulting noise function ϵ' is nonbalanced and $F_i(z)$ does not divide $h_i(z)$, the reconstruction of $\{x_{it}\}$ is possible in principle.

Similarly, one can conjecture that the linear function of the output sequence in the generating function domain defined by $\varphi(z)\ \Pi_{i=1}^{N}F_i(z)\ Y$ is with high probability nonbalanced.

Note that the described decomposition procedure may lead to efficient reconstruction only if the absolute value of the correlation coefficient between the resulting noise function and the zero function is sufficiently large. Since the noise function is expressed as the sum of individual noise functions, the absolute value of the correlation coefficient is in general much smaller than for memoryless combiners. It appears reasonable that in order to maximize the absolute value of the overall correlation coefficient one should approximate the output function and the components of the next-state function by the linear functions with maximum absolute values of correlation coefficients. If we use the Walsh transform, determination of such functions requires $O((M+1)(M+N)2^{M+N})$ time complexity, assuming that most of the functions effectively depend on all $M+N$ input and state variables. If many of them depend on just a subset of the input and state variables, the time complexity can be considerably smaller.

VI. CONCLUSION

In this paper, we study the correlation properties of a general binary combiner with an arbitrary number, H. of memory bits. We show that there exists a pair of correlated linear functions of at most M+l successive output and input bits, respectively, which is a generalization of a result from [4] regarding the binary combiners with one bit of memory. We also develop and analyze an efficient procedure for finding such pairs of linear functions. The procedure is based on a linear sequential circuit approximation of a nonlinear combiner with memory. It is still an open question whether there exist other such procedures. The result may be a basis for a divide and conquer correlation attack on a stream cipher generator consisting of several linear feedback shift registers combined by a combiner with memory. In general, it also applies to an arbitrary synchronous finite-state machine as well.

REFERENCES

[l] A. Gill. Linear Sequential Circuits. McGraw-Hill. 1966.

[2] W. Meier, 0. Staffelbach. "Fast correlation attacks on certain stream ciphers", Journal *of* Cryptology, Vol. 1 (3), pp. 159-176. 1989.

[3] W. Meier. 0. Staffelbach. "Nonlinearity criteria for cryptographic functions". Aduances In Cryptology - - *EUROCRYPT* "89, Proceedings, LflCS. Vol. 434. pp. 549-562. Springer-Verlag. 1990.

[4] W. Meier. 0. Staffelbach. "Correlation properties of combiners with memory in stream ciphers". *Advances Ln* Cryptology - *EUROCRYPT* '90, Proceedings, tWCS, Vol. 473, pp. 204-213. Springer-Verlag, 1991.

[5] R.A. Rueppel, Analysis and Design of Stream Ciphers. Springer-Verlag, 1986.

[6] R.A. Rueppel, "Correlation immunity and the summation generator". Aduances In Cryptology - CRTPTO "85, Proceedings. LNCS. pp. 260-272. Springer-Verlag. 1986.

[7] T. Siegenthaler, "Decrypting a class of stream ciphers using ciphertext only". IE££ Trans. *Conput . ,* Vol. C-34, pp. 81-85. Jan. 1985.

[8] T. Siegenthaler, •'Correlation-immunity of nonlinear combining functions for cryptographic applications", *IEEE Trans. Inform. Theory,* Vol. IT-30, pp. 776-780. Sept. 1984. *

[9] G.Z. Xiao. J.L. Massey. "A spectral characterization of correlation-immune combining functions". IEEE Trans. Inform. Theory. Vol. IT-34. pp. 569-571, Hay 1988.