Lecture Four

Basic Efficiency Criteria of LFSR's-Systems

5. Correlation Immunity (CI) Criterion

Remark (5.1): (**correlation attacks**) Suppose that n maximum-length LFSRs $R_1,R_2,...,R_n$ of lengths $r_1,r_2,...,r_n$ are employed in a nonlinear combination generator. If the connection polynomials of the LFSRs and the combining function f are public knowledge, then the number of different keys of the generator is $\prod_{i=1}^{n}(2^{r_i}-1)$.(A key consists of the initial states of the LFSRs) Suppose that there is a correlation between the keystream and the output sequence of R_1, with correlation probability Pr>1/2. If a sufficiently long segment of the keystream is known (e.g., as is possible under a known-plaintext attack on a binary additive stream cipher), the initial state of R_1 can be deduced by counting the number of coincidences between the keystream and all possible shifts of the output sequence of R_1, until this number agrees with the correlation probability Pr. Under these conditions, finding the initial state of R_1 will take at most $2^{r_i}-1$ trials. In the case where there is a correlation between the keystream and the output sequences of each of $R_1,R_2,...,R_n$, the (secret) initial state of each LFSR can be determined independently in a total of about $\prod_{i=1}^{n}(2^{r_i}-1)$ trials; this number is far smaller than the total number of different keys.

In a similar manner, correlations between the output sequences of particular subsets of the LFSRs and the keystream can be exploited.

The combining function f should be carefully selected so that there is no statistical dependence between any small subset of the n LFSR sequences and the keystream. This condition can be satisfied if f is chosen to be m^{th}-order correlation immune.

Definition (5.1) Let $X_1,X_2,...,X_n$ be independent binary variables, each taking on the values 0 or 1 with probability 1/2. A Boolean function $f(x_1,x_2,...,x_n)$ is m^{th}-order correlation immune if for each subset of m random variables $X_{i_1}, X_{i_2},...,X_{i_m}$ with $1\le i_1<i_2<\ldots<i_m\le n$, the random variable $Z=f(X_1,X_2,...,X_n)$ is statistically independent of the random vector $(X_{i_1}, X_{i_2},...,X_{i_m})$.

Example (5.1): the function $f(x_1,x_2,...,x_n)= x_1\oplus x_2\oplus\ldots\oplus x_n$ is $(n-1)^{th}$ order correlation immune.

In light of Remark (5.2), the following shows that there is a tradeoff between achieving high linear complexity and high correlation immunity with a combining function.

Remark (5.2): If a Boolean function $f(x_1,x_2,...,x_n)$ is m^{th}-order correlation immune, where $1\le m<n$, then the nonlinear order of f is at most $n-m$. Moreover, if f is balanced (i.e., exactly half of the output values of f are 0) then the nonlinear order of f is at most $n-m-1$ for $1\le m\le n-2$.

The tradeoff between high linear complexity and high correlation immunity can be avoided by permitting memory in the nonlinear combination function f.

For combination generators, the correlation attack can be prevented by using a combining function f whose output is not correlated to any of its inputs. Such functions are called (n-1)-order correlation-immune.

For the 3-LKG, the CI(S)=n-1=2, but for the 3-PKG, the CI(S)=0, since the non-linear order of product system is 3.

The correlation immunity order can be calculated from logical truth table for CF depending on calculating correlation probability, notice that correlation immunity depends on combining function unit only and there is little effect of LFSR unit. Therefore, the condition to obtain efficient KG's **"**Choosing CF with maximum order correlation immune", this condition is not essential since the correlation (if it exist) can prevented by using some ways. Moreover, Staffelbach, mentioned that to prevent correlation attack, long LFSR's must be used with maximum number of tapping, so the other condition to obtain efficient KG's is "Using long LFSR's with maximum tapping number of connection polynomial".

Table (4) shows the correlation probability p_r for the thee systems using three LFSR's.

Table (4) Correlation probability for the thee systems.

System	Input			Output		
	x_1	x_2	x_3	F_L	F_P	F_B
	0	0	0	0	0	0
	0	0	1	1	0	0
	0	1	0	1	0	0
	0	1	1	0	0	1
	1	0	0	1	0	0
	1	0	1	0	0	1
	1	1	0	0	0	1
	1	1	1	1	1	1
Linear	0.5	0.5	0.5	**Correlation Probability**		
Product	0.625	0.625	0.625			
Brüer	0.75	0.75	0.75			