On the Linear Consistency Test (LCT) in Cryptanalysis with Applications *

Kencheng Zeng[1], C.H. Yang[2], and T.R.N. Rao[2]

[1]Graduate School of USTC
Academia Sinica
P.O. Box 3908
Beijing
People's Republic of China

[2]The Center for Advanced Computer Studies
University of Southwestern Louisiana
P.O. Box 44330
Lafayette, LA 70504-4330

Abstract. *In this paper, we give at first a precise estimation for the consistency probability of a system of linear algebraic equations $Ax = b$ with random $m \times n$ coefficient matrix A, $m > n$, and fixed non-zero right side b. A new test in cryptanalysis is then formulated on the basis of the estimation and applied to attack the multiplexing generator of Jennings (1980) and the multiple-speed generator of Massey-Rueppel (1984). Some security remarks concerning the perfect linear cipher of the latter authors are also made.*

Linearity is the curse of the cryptographer

— J. Massey —

I. Introduction

Cryptanalysis is in the last run a matter of searching [1]. In cracking a more or less seriously designed cryptosystem, exhaustive searching at some level is inevitable. The problem is in the range of which objectives to make the searching tests so as to minimize the amount of work needed, and according to which criteria to signalize discovery of the objectives in search so as to maximize the probability of successful key identification.

* This research is supported by Board of Regents of Louisiana Grant #86-USL(2)-127-03

G. Brassard (Ed.): Advances in Cryptology - CRYPTO '89, LNCS 435, pp. 164-174, 1990.

If the entire key of secrecy $\mathbf{K}$ in a system can be revealed by exhaustive searching concentrated on a certain subkey $\mathbf{K}_1$, then this will mean that only the $|\mathbf{K}_1|$ key bits in search are responsible for the cryptographic strength of the system, and the remaining $|\mathbf{K}| - |\mathbf{K}_1|$ bits of key information are redundant. The ratio $\rho = \dfrac{|\mathbf{K}| - |\mathbf{K}_1|}{|\mathbf{K}|}$ can be called the *key information redundancy* rate of the system.

Systems which can be cracked by pure analytic attack, such as those discussed in [2], have key information redundancy rate $\rho = 1$, but similar cases rarely happen in practice.

The problem now is how to discover the redundancy. The rubric of J. Massey quoted above gives us an important hint that in certain cases such redundancy may be found by making use of the linearity latent in the systems under consideration.

Following this idea, we prove in the present paper a theorem on the consistency probability of a system of linear algebraic equations $A\,x = b$ with random $m \times n$ coefficient matrix A, $m > n$, and fixed non-zero right side vector b. On the basis of this theorem, we set up a new cryptanalytic test, called the linear consistency test (LCT), and apply it to disclose the key information redundancy in several random bit generators published in the open literature.

II. The Consistency Probability of $A\,x = b$

We start with proving the following two simple algebraic propositions.

Lemma 1. Let $A = \left(a(i,j)\right)$ be an $m \times n$ random binary matrix with entries satisfying, independently from each other, the distribution $Prob\left(a(i,j) = 0\right) = \dfrac{1}{2}$. Then for any integer r, $0 < r \leqslant n$, the probability for A to have rank r is

$$prob\left(rank(A) = r\right) = C_n^r\, 2^{-m(n-r)} \prod_{i=m-r+1}^{m} \left(1 - \frac{1}{2^i}\right). \tag{1}$$

Proof. Consider the direct product $G = GL(m, F_2) \times S_n$ of the m-dimensional general linear group $GL(m, F_2)$ and the symmetric group S_n of degree n, acting on the object set $\Omega = \{A\}$ of all possible $m \times n$ matrices over F_2, in such a way that for any $A \in \Omega$ and $g = (P, Q)$, $P \in GL(m, F_2)$, $Q \in S_n$, we have $\pi_g(A) = PAQ$. It is well known [5], that the subset of all $m \times n$ matrices of rank r form a G-orbit with representative

$$I_{m,n,r} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus we see the number of all $m \times n$ matrices of rank r over F_2 is equal to

$$N_{m,n,r} = \left| \Omega_{I_{m,n,r}} \right| = \frac{\left| GL(m, F_2) \right| \times \left| S_n \right|}{\left| Stab_G(I_{m,n,r}) \right|} .$$

But we know

$$\left| GL(m, F_2) \right| = 2^{m^2} \prod_{i=1}^{m} (1 - \frac{1}{2^i}), \quad \left| S_n \right| = n! ,$$

so for the purpose of computing $N_{m,n,r}$, we need only determine the order of the stabilizer of $I_{m,n,r}$ in G. In doing this, we partition the square matrices P, Q^{-1} into block forms compatible with $I_{m,n,r}$

$$P = \begin{pmatrix} P(1,1) & P(2,1) \\ P(1,2) & P(2,2) \end{pmatrix}, \quad Q^{-1} = \begin{pmatrix} Q(1,1) & Q(2,1) \\ Q(1,2) & Q(2,2) \end{pmatrix} .$$

Then it follows from $PI_{m,n,r} = I_{m,n,r}Q^{-1}$ that

$$P(1,1) = Q(1,1), \quad P(2,1) = 0, \quad Q(1,2) = 0 ,$$

where

$$P(2,2) \in GL(m-r, F_2), \quad Q(1,1) \in S_r, \quad Q(2,2) \in S_{m-r} .$$

Moreover, we have $Q(2,1) = 0$, since Q^{-1} is a permutation matrix, and $P(1,2)$ can be an arbitrary matrix. Therefore, we have

$$\left| Stab_G(I_{m,n,r}) \right| = r!\,(n-r)!\,2^{m(m-r)} \prod_{i=1}^{m-r} (1 - \frac{1}{2^i}) ,$$

and

$$N_{m,n,r} = C_n^r 2^{mr} \prod_{i=m-r+1}^{m} (1 - \frac{1}{2^i}) .$$

But $\left| \Omega \right| = 2^{mn}$, so we get (1).

Lemma 2. Let b be any given non-zero vector in the m-dimensional vector space $V_m(F_2)$, and r any non-negative integer not greater than m. If the r-dimensional subspaces of $V_m(F_2)$ can be generated equiprobabilistically, then the probability for a randomly generated r-dimensional subspace W to contain b is

$$Prob\left(b \in W\right) = \frac{2^r - 1}{2^m - 1} . \tag{2}$$

Proof. Every r-dimensional subspace of $V_m(F_2)$ containing b can be spanned by a basis of the form $(b, w_1, w_2, \ldots, w_{r-1})$. There are altogether

$$B_r = 2^{m(r-1)} \prod_{i=m-r+1}^{m-1} (1 - \frac{1}{2^i})$$

similar vector sets in $V_m(F_2)$, and the vector set $(b, w'_1, w'_2, \ldots, w'_{r-1})$ will span the same subspace as $(b, w_1, w_2, \ldots, w_{r-1})$ iff

$$(b, w'_1, w'_2, \ldots, w'_{r-1}) = (b, w_1, w_2, \ldots, w_{r-1}) \begin{pmatrix} 1 & c \\ 0 & Q \end{pmatrix},$$

where $Q \in GL(r-1, F_2)$ and c is an arbitrary $(r-1)$-tuple. So we see the number of r-dimensional subspace in $V_m(F_2)$, which contain b, is

$$N_{r,b} = \frac{B_r}{2^{r-1} \left| GL(r-1, F_2) \right|}.$$

On the other hand, the number of arbitrary r-dimensional subspaces of $V_m(F_2)$ can be derived in the same way to be

$$N_r = \frac{2^{mr} \prod_{i=m-r+1}^{m} (1 - \frac{1}{2^i})}{\left| GL(r, F_2) \right|}.$$

So we have

$$Prob\left\{b \in W\right\} = \frac{N_{r,b}}{N_r} = \frac{2^r - 1}{2^m - 1}.$$

Theorem 1. Let A and b be as described in the lemmas and $m > n$, then the probability for the linear system $A x = b$ to be consistent is

$$Prob\left\{A x = b \text{ is consist.}\right\} < \frac{1}{2^{m-n}} (1 + \frac{1}{2^{m+1}})^n. \qquad (3)$$

Proof. Denote by $L(A)$ the subspace of $V_m(F_2)$ spanned by the n column vectors of A, then the system is consistent iff $b \in L(A)$. Therefore we have

$$Prob\left\{A x = b \text{ is consistent}\right\} = Prob\left\{b \in L(A)\right\}$$

$$= \sum_{r=0}^{n} Prob\left[rank\ A = r\right] Prob\left[b \in L(A) \mid dim\ L(A) = r\right]$$

$$= \sum_{r=0}^{n} C_n^r\, 2^{-m(n-r)} \prod_{i=m-r+1}^{m} (1 - \frac{1}{2^i}) \cdot \frac{2^r - 1}{2^m - 1}$$

$$= \frac{1}{2^m} \sum_{r=0}^{n} C_n^r\, 2^{-m(n-r)} \prod_{i=m-r+1}^{m-1} (1 - \frac{1}{2^i}) \cdot (2^r - 1)$$

$$< \frac{1}{2^m} \sum_{r=0}^{n} C_n^r\, 2^{-m(n-r)} \cdot (2^r - 1)$$

$$= \frac{1}{2^m} \sum_{r=0}^{n} C_n^r\, 2^{-m(n-r)} \cdot 2^r - \frac{1}{2^m} \sum_{r=0}^{n} C_n^r\, 2^{-m(n-r)}$$

$$= \frac{1}{2^m}(2 + \frac{1}{2^m})^n - \frac{1}{2^m}(1 + \frac{1}{2^m})^n$$

$$< \frac{1}{2^{m-n}}(1 + \frac{1}{2^{m+1}})^n .$$

III. The Linear Consistency Test (LCT)

In considering a keystream generator, it is sometimes possible to single out a certain subkey $\mathbf{K}_1$ from the entire key of secrecy $\mathbf{K}$ and write out a system of linear equations of the form

$$A(\mathbf{K}_1)\, x = b\ , \tag{4}$$

where the coefficient matrix $A(\mathbf{K}_1)$ is determined by the bit-generating algorithm and is parametrized by $\mathbf{K}_1$, while b is determined by the captured segment of the output sequence. The solution vector x, in general, can be used to determine the remaining part of $\mathbf{K}$.

If the parameter $\mathbf{K}_1$ coincides with the subkey used in generating the captured segment under consideration, then (4) certainly will be consistent. On the other hand, if the parameter $\mathbf{K}_1$ is not the subkey used, then by theorem 1 the consistency probability of the system will be very small when the captured segment is long enough.

Thus, for the purpose of finding the right subkey $\mathbf{K}_1$, we need only test the consistency of (4) with respect to all possible choices of the parameter $\mathbf{K}_1$, and signalize discovery of the subkey in search whenever the system is found to be consistent. The number of cases to be tested is $2^{|\mathbf{K}_1|}$, and the work factor needed for each test is that of the Gauss elimination algorithm applied to the augmented matrix $(A(\mathbf{K}_1), b)$.

In order to make the number of false consistency alarms as small as possible, the number of equations in (4) should exceed $|x| + |\mathbf{K}_1|$ significantly. This being the case, another consequence is that the solution x of a consistent system (4) will be, *with probability nearly 1*, unique. In certain situations, for example, in the problems to be considered below, this means no further large scale searches will be needed for revealing the entire key.

The following *pop melody* in stream cipher cryptography is in many cases helpful in forming up the linear system (4) needed in applying the LCT.

Lemma 3. If the linear recursive sequence $\mathbf{c} = \{c(t) \mid t \geqslant 0\}$ has a feedback polynomial $f(x)$ of degree n, and

$$x^t = r(x) = r_{t,0} + r_{t,1}x + \cdots + r_{t,n-1}x^{n-1} \quad mod\ f(x)$$

then

$$c(t) = r_{t,0}c(0) + r_{t,1}c(1) + \cdots + r_{t,n-1}c(n-1)\,. \tag{5}$$

Proof. Write $x^t = q(x)f(x) + r(x)$, then we have

$$\mathbf{d} \triangleq (x^t + r(x))\mathbf{c} = (x^t + r(x))\mathbf{c} + q(x)f(x)\mathbf{c} = (x^t + q(x)f(x) + r(x))\mathbf{c} = 0\,,$$

and (5) follows from examining the expression for the signal $d(0)$.

IV. Cracking the Generators of Jennings and Massey-Rueppel

The generators proposed by Jennings and Massey-Rueppel both use two LFSR sequences with primitive feedback polynomials $f(x)$ and $g(x)$, of degrees l and n respectively, as source sequences, but combine them by different key-controlled algorithms. The LCT will show that both of them suffer from a fairly large key information redundancy.

(A) The Multiplexing Generator of Jennings

According to Chambers [6], similar schemes have been recommended by the European Broadcasting Union as standards for scrambling television broadcasts. The generator produces the output signals $c(t)$, $t \geqslant 0$, in the following way: Fix a positive integer $h \leqslant \min(l, \lfloor \log_2 n \rfloor)$ and a tap pattern $0 \leqslant i_0 < i_1 < \cdots < i_{h-1} \leqslant l-1$ on LFSR-1. For every moment $t \geqslant 0$ form the number

$$u(t) = a(t + i_0) + a(t + i_1)2 + \cdots + a(t + i_{h-1})2^{h-1}$$

and transform it into

$$\theta(u(t)) = s_0(t) + s_1(t)2 + \cdots + s_{k-1}(t)2^{k-1}\,, \qquad k = \lceil \log_2 n \rceil$$

by an injective mapping θ: $\{0, 1, ..., 2^b - 1\} \to \{0, 1, ..., n - 1\}$, which together with the initial states of LFSR-1 and LFSR-2 form the key of secrecy of the system. The output signal is $c(t) = b\left(t + \theta(u(t))\right)$.

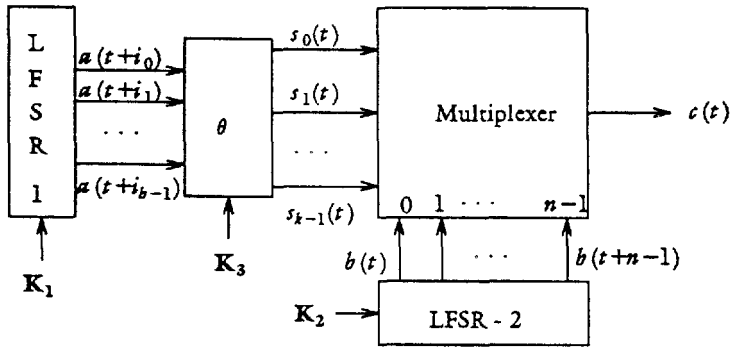

Figure 1. The Jennings Generator

It has been shown [3], that if (l, n) = 1, then the output sequence has period $(2^l - 1)(2^n - 1)$ and linear complexity $LC \leqslant n(1 + \sum_{i=1}^{b} C_i^l)$, with equality if the tap positions are spaced at equal intervals. Thus, the strive for the highest linear complexity will greatly limit the choice of the tap patterns. Therefore, without losing any generality of our analysis, we can assume the tap pattern is fixed and known.

Theorem 2. If the feedback polynomials $f(x)$ and $g(x)$ are known to the cryptanalyst, then the Jennings generator can be broken on an output segment of length $N \geqslant l + n2^b$ by 2^{l+b} linear consistency tests.

Proof. The cracking procedure starts with applying the LCT to determine the initial state $\mathbf{a}_0$ of LFSR-1 corresponding to the captured segment $\mathbf{c}$ of length N.

Step 1. For each $0 \leqslant t \leqslant N-1$, divide x^t by $g(x)$ to obtain the remainder

$$r_t(x) = r_{t,0} + r_{t,1}x + \cdots + r_{t,n-1}x^{n-1}$$

and store the vector $\mathbf{r}(t) = (r_{t,0}, r_{t,1}, ..., r_{t,n-1})$.

Step 2. For every non-zero $\mathbf{a} \in V_l(F_2)$, set LFSR -1 to work with $\mathbf{a}$ as initial state and form 2^b linear systems

$$S_k: A_k x = \mathbf{c}_k, \quad 0 \leqslant k \leqslant 2^b - 1, \tag{6}$$

by putting the equation $(\mathbf{r}(t), x) = \mathbf{c}(t)$ into the system S_k whenever

$u(t) = k$.

For $0 \leqslant k \leqslant 2^b - 1$, test the consistency of S_k. Discard $\mathbf{a}$ whenever inconsistency is alarmed. The vectors $\mathbf{a}$, for which all the systems S_k turn to be consistent, will be reserved as candidates for $\mathbf{a}_0$. The true initial vector $\mathbf{a}_0$ will certainly be reserved and the probability p for an arbitrary $\mathbf{a}$ to be reserved can be estimated in the following way: Let m_k be the number of equations in S_k and assume that $m_k > n$ for $k < q$ and $m_k \leqslant n$ for $k \geqslant q$. By theorem 1, the probability for S_k to be consistent is

$$p_k < \frac{1}{2^{m_k - n}}\left(1 + \frac{1}{2^{m_k+1}}\right)^n, \quad 0 \leqslant k \leqslant q-1,$$

so we see

$$p < 2^{qn - \sum_{k=0}^{q-1} m_k} \cdot \prod_{k=0}^{q-1}\left(1 + \frac{1}{2^{m_k+1}}\right)^n$$

$$\leqslant 2^{2^b n - N} \cdot \left(1 + \frac{1}{2^{n+2}}\right)^{qn} < 2^{-l}\exp\left(\frac{n^2}{2^{n+2}}\right).$$

Step 3. Let $\mathbf{a}$ be any candidate vector. Consider any system in (6), for which the coefficient matrix A_k has the largest rank, and denote the set of the solutions as vectors in $V_n(F_2)$ by V. According to lemma 1, we have $|V| = 1$ with probability nearly 1. Choose an arbitrary $v_0 \in V$ and consider the vectors

$$v_{-n+1}, \cdots, v_{-1}, v_0, v_1, \cdots v_{n-1} \tag{7}$$

which can be generated successively by LFSR-2 starting from v_{-n+1}. Check whether there is in (7) a subset of 2^b vectors

$$v_{i_0}, v_{i_1}, \cdots, v_{i_{2^b-1}}, \tag{8}$$

satisfying the following two conditions:

$$\text{(a) } A_k v_{i_k}{}^T = \mathbf{c}_k, \quad 0 \leqslant k \leqslant 2^b - 1; \quad \text{(b) } \max\{i_k\} - \min\{i_k\} < n. \tag{9}$$

Discard v_0 if such a subset in (7) does not exist. Discard $\mathbf{a}$ if all vectors in V are discarded.

Since the probability for an arbitrary set of 2^b vectors in $V_n(F_2)$ to have the property that LFSR-2, starting from a certain one of them, will generate the remaining vectors within $n-1$ steps is of the order of magnitude $O\left(2^{-n(2^b-1)}\right)$, all the candidate vectors $\mathbf{a}$, except $\mathbf{a}_0$, will be discarded and $v_0 \in V$ will also be uniquely determined.

Step 4. Write $\sigma \triangleq \min i_k$, $\tau \triangleq \max i_k$, $\rho \triangleq n + \sigma - \tau$, and suppose LFSR-2, starting from v_σ, arrives at the vector v_{i_k} of (8) in n_k steps, conclude

$$\theta(k) = n_k + \nu , \quad 0 \leqslant k \leqslant 2^b - 1 ,$$

where ν may be any non-negative integer not exceeding ρ, and the corresponding initial state of LFSR-2 will be the vector generated by LFSR-2 after ν steps of work with v_σ as start.

(B) The Multiple Speed Generator and the Perfect Linear Cipher

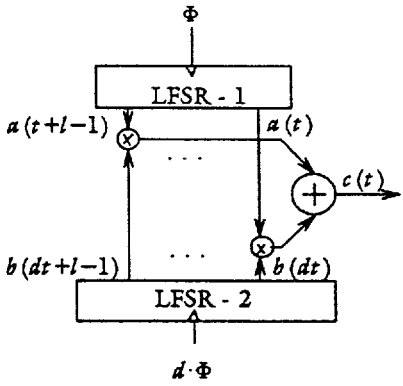

Figure 2. Multiple speed generator

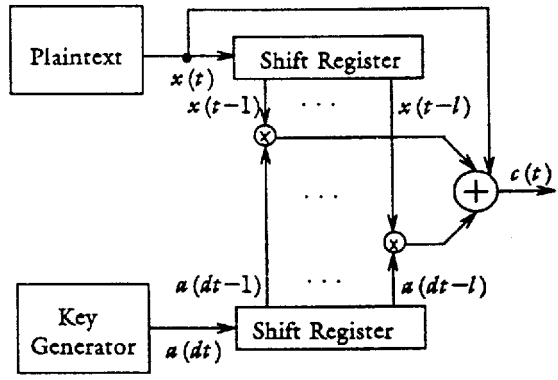

Figure 3. Perfect linear cipher

In the multiple speed generator of Massey-Rueppel, the shift register LFSR-2 is clocked at a speed $d \geqslant 2$ times as fast as the LFSR-1 and the output signal $c(t)$ is produced according to

$$c(t) = \sum_{i=0}^{l-1} a(t+i)b(dt+i) .$$

The speed factor d is variable and is used as a part of the key of secrecy. As a result of various technical limitations, d cannot be too large, and one can estimate an upper bound $d_{\max}$ to it. Therefore, if the feedback polynomials $f(x)$ and $g(x)$ are known, then the cryptanalyst can determine d and the initial state of LFSR-1, and hence break the system by the LCT applied to an output segment of length $N \geqslant l + n + \log_2 d_{\max}$.

More interesting, however, is the perfect linear cipher considered in connection with the problem of introducing automatic authentication into crypto-systems. As pointed out in [7, p.419], in the case of a block cipher, as a result of the existence of the diffusion effect, this can be achieved by applying to the plaintext any error control code, linear or nonlinear, before

encrypting. For stream ciphers in general, only non-linear or keyed codes can serve the purpose. In the perfect linear cipher, however, a single error in the plaintext will propagate over a range of length l, so linear codes can be used to realize effective authentication. But the analysis below will show that this will make the system insecure.

Theorem 3. If the key generator in Figure 3 is an LFSR with known primitive feedback polynomial $f(x)$ of degree l, and the plaintext is a string of codewords of a known systematic linear code of information rate $r = k/n$, then the cipher can be broken by a ciphertext-only attack consisting of $n2^l$ LCT applied to a captured segment $\mathbf{c}$ of length

$$N \geqslant \frac{2l + \log_2 n}{1 - r} + n \,. \tag{10}$$

Proof. Assume the speed factor known and let the generator matrix of the linear code be $\mathbf{G} = [\mathbf{I}_k : (p_{ij})]$. The claimed ciphertext-only attack starts with the observation that if for some $0 \leqslant w \leqslant n-1$, we assume $x(w)$ to be the first bit of a codeword, then for any $0 \leqslant q < B = \lfloor \frac{N-w}{n} \rfloor$, we shall have $n-k$ linear relations of the form

$$x(w+qn+k+j) = \sum_{i=0}^{k-1} p_{i,j}\, x(w+qn+i) \,, \qquad 0 \leqslant j \leqslant n-k-1 \tag{11}$$

to express the parity signals in terms of the information signals. After making these substitutions in the relations

$$c(t) = x(t) + \sum_{i=1}^{l} a(dt-i)x(t-i) \,, \qquad w \leqslant t < w + nB \,, \tag{12}$$

we shall have a set of nB linear equations of the form

$$L_t\left(x(w-l),..,x(w-1);..;x(w+qn),x(w+qn+1),..,x(w+qn+k-1);..\right) = c(t) \tag{13}$$

in $l + kB$ unknowns. Thus we can apply the LCT to determine the initial state of the key generator and the number w. Since the consistency probability of (13) is $p < \frac{1}{2^{nB-kB-l}}$, the mathematical expectation of the number of possible consistency alarms will be $E < n2^l p < 1$. False consistency alarms can be effectively discarded by checking the code structure of the recovered plaintext. This proves the theorem.

References

1. Martin E. Hellman, Ehud D. Karnin, Justin Reyneri, *On the Necessity of Cryptanalytic Exhaustive Search*, ACM SIGACT, Vol. 18, No. 2, 1984, pp. 40-44.

2. C.H. Yang, Kencheng Zeng, and T.R.N. Rao, *An Improved Linear Syndrome Algorithm in Cryptanalysis with Applications*, to appear.

3. Sylvia M. Jennings, *A Special Class of Binary Sequences*, University of London, 1980, Ph.D. Thesis.

4. James L. Massey and Rainer A. Rueppel, *Linear Ciphers and Random Sequence Generators with Multiple Clocks*, EUROCRYPT 84, 1984, pp. 74-87.

5. A. Adrian Albert, *Fundamental Concepts of Higher Algebra*, University of Chicago Press, 1956.

6. W.G. Chambers, *Clock-controlled shift registers in binary sequence generators*, Proc. IEE, Pt. E., Jan. 1988, pp. 17-24.

7. Whitfield Diffie, Martin E. Hellman, *Privacy and Authentication: An Introduction to Cryptography*, Proc. IEEE, Mar. 1979, pp. 397-427.