# SUBSTANTIAL NUMBER OF CRYPTOGRAPHIC KEYS AND ITS APPLICATION TO ENCRYPTION DESIGNS

Eiji OKAMOTO

C&C Information Research Laboratories
NEC corporation
4-1-1, Miyazaki, Miyamae-ku
Kawasaki, 213 JAPAN

## ABSTRACT

A new concept of the substantial number of cryptographic keys (SNK) in key spaces is proposed and is applied to encryption designs. SNK is defined as the number of keys which is far from each other. It must be greater than $2^{56}$, for instance, to have essentially the same number of keys in DES. This SNK condition restricts design parameters of encryption systems. In this paper, SNK is strictly defined in key spaces, followed by illustrations of SNK's in fundamental encryption algorithms and product ciphers. Then SNK is applied to the design of encryption systems to decide the design parameters. It is usefui for designing product cipher in particular. SNK should be considered as one of the criteria of encipherment strength.

## I . INTRODUCTION

In encryption designs, the technique of combining two or more fundamental encryption algorithms is very useful, because it produces a complicated encryption scheme and a lot of keys. The product of the numbers of keys in the fundamental encryption algorithms is usually regarded as the number of keys in the combined encryption scheme. Some product ciphers, however, do not have so many keys. In Fig.1, for example, the total number of keys in the product cipher of an $n$ bit block substitution cipher and an $n$ bit block transposition cipher is not substantially equal to $n!\,2^n!$, but $2^n!$. This shows all encryption scheme must have the property of key independence from each other. In other words, the deciphered message with a wrong key must be totally different from the original message.

C.G. Guenther (Ed.): Advances in Cryptology - EUROCRYPT '88, LNCS 330, pp. 361-373, 1988.

There are two methods of designing encryption schemes to overcome the mutual dependence of keys. The first method is based on the key selection such that the keys to select are separated from each other in the key space, called *'sphere packing cipher'*. Nakamura[1] showed this kind of self-synchronizing stream cipher scheme using error correcting codes. The design of transposition ciphers using Reed-Solomon codes in [2] is also based on the same idea.

The second method is based on the design scheme such that the probability of any key lying in the neighborhood of any other key is to be made as small as $2^{-56}$, for instance. This method does not require special selection of keys as in the first method. Users can select any key in the key space.

The second method leads to a new concept of the number of keys, *Substantial Number of Keys (SNK)*. Roughly speaking, SNK is the number of keys which are different from each other in the sense that the close keys are regarded as one key. In this paper, the difference of two keys in the key space is defined precisely and SNK is discussed in this space. The design parameters of any encryption scheme are restricted by the condition that the encryption scheme should have enough SNK to avoid exhaustive key attacks. The sphere packing cipher is also reviewed from the point of SNK. The SNK should be considered as one of the criteria of encipherment strength.

## II . SUBSTANTIAL NUMBER OF KEYS (SNK)

### 1. *Definition of SNK*

A key space consists of a set of all keys, probabilities of selecting any key and differences between any two keys. The key set of transposition cipher, for example, contains all transpositions including the through one of input data. Let $Q_d(K)$ be the probability of selecting a key lying in the sphere of radius $d$ from key $K$. Then, the substantial number of keys, $SNK_d$, regarding any two keys within difference $d$ of each other as same, is defined as

$$SNK_d = \frac{1}{A[Q_d(K)]}, \tag{1}$$

where $A[\ ]$ means the average with respect to the probability of selecting keys. This definition is justified by the following example: the total number $N$ of stones is given by $1/Q$ when the probability of selecting any one stone from all stones is $Q$, because $Q = 1/N$.

Although the difference of two keys in the key space could be defined variously, this paper employs *reversed-bits rate*[1], $r(K1, K2)$, to define it.

$$r(K1, K2) = A[\frac{h(M, D_{K2}(E_{K1}(M)))}{L(M)}] \tag{2}$$

Here, $M$ is any message, and $E_K(\ )$, $D_K(\ )$ are encryption and decryption with key $K$, respectively. Key $K2$ is not necessarily the corresponding decryption key of $K1$. Function $h(\ ,\ )$ shows *Hamming distance*, and $L(\ )$ shows *length*. In the Eq.(2), $A[\ ]$ is the average when message $M$ is randomly selected from the message space which contains all messages. Then, the difference $\rho(K1, K2)$ between two keys $K1$ and $K2$ is defined as

$$\rho(K1, K2) = \frac{1}{2} - \left| \frac{1}{2} - r(K1, K2) \right| . \tag{3}$$

The difference $\rho$ is the reversed-bits rate $r$ when $r \leq 1/2$, or $1 - r$ when $r > 1/2$. In other words, it means the minimum difference between the reversed-bits rate and 0 or 1. The measure is useful especially for voice data.

*2. Examples of SNK*

This section illustrates SNK's of four block ciphers in Fig.2. In the figure, (a),(b) and (c) are examples of fundamental ciphers and (d) is an example of a product cipher. Every key is selected with equal probability. The integer $n$ means block length of ciphers.

a) Exclusive-or cipher

An exclusive-or cipher has vector $P$ as a key. The key space is an $n$ dimensional space which contains $2^n$ keys in all. If the Hamming distance between the encrypting key $P1$ and the decrypting key $P2$ is $h$, the reversed-bits rate $r$ is given by

$$r = \frac{h}{n} . \tag{4}$$

If $P2$ is a uniform random variable, the distance $h$ is a binomial random variable. Then the probability of $Q_d = A[Q_d(P1)]$ is

$$Q_d = \sum_{\rho<d} \frac{\binom{n}{h}}{2^n} = \sum_{\substack{r<d \\ r>1-d}} \frac{\binom{n}{h}}{2^n} = \sum_{\substack{h<dn \\ h>(1-d)n}} \frac{\binom{n}{h}}{2^n} . \tag{5}$$

Since binomial distribution is approximated by Gaussian:

$$\sum_{i=0}^{k} \binom{n}{i} p^i q^{n-i} \simeq 1 - erf\left( \frac{k - np}{\sqrt{npq}} \right) \tag{6}$$

$$p + q = 1 \tag{7}$$

$$erf(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt\,, \tag{8}$$

The probability $Q_d$ is nearly equal to

$$Q_d \simeq 2\ erf\big((1-2d)\sqrt{n}\big)\,. \tag{9}$$

Therefore $SNK_d$ is

$$SNK_d \simeq \frac{1}{2\ erf\big((1-2d)\sqrt{n}\big)}\,. \tag{10}$$

Figure 3 (a) shows the SNK curve of exclusive-or ciphers with respect to $n$, where $d$ is regarded as a parameter. The number $k$ is a length of SNK:

$$k = \log_2 SNK. \tag{11}$$

The data block length $n$ should be more than 500, if $SNK > 2^{56}$ and the reversed-bits rate lies between 0.3 and 0.7.

b) Substitution cipher

A substitution cipher of $n$-bit block is a permutation of $n$-bit patterns, hence the total number of keys is $2^n!$. Let $K1$,$K2$ denote keys of encryption and decryption transformations, respectively, and $D_{K2}E_{K1}$ be the composite transformation of the two transformations. The reversed-bits rate between any input bit to $D_{K2}E_{K1}$ and any output bit from it is equal to that of between the MSB's (most significant bit) of the input and the output. Figure 4 illustrates an example of substitution ciphers when $n = 3$. When Hamming distance between column $I1$ (MSB in the input bits) and $O1$ (MSB in the output bits) is $2h$, which is always even, the reversed-bits rate is

$$r = \frac{2h}{N},$$

and the total number of substitution ciphers is given by:

$$\binom{\frac{N}{2}}{h}^2 \frac{N}{2}!^2.$$

Therefore, a probability of $r < d$ or $r > 1 - d$ is

$$Q_d = \sum_{\substack{h<dM \\ h>(1-d)M}} \frac{\left(\binom{M}{h} M!\right)^2}{N!} = \frac{2}{\binom{N}{M}} \sum_{h<dM} \binom{M}{h}^2, \tag{12}$$

where $M = N/2 = 2^{n-1}$. As the binomial distribution is approximated by

$$\binom{n}{i} p^i q^{n-i} \simeq \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(i-np)^2}{2npq}}, \tag{13}$$

the probability $Q_d$ is approximately equal to

$$Q_d \simeq 2\ erf\left(\frac{(1-2d)\,N/4}{\sqrt{N/16}}\right) = 2\ erf\left((1-2d)\sqrt{2^n}\right). \tag{14}$$

The equation (14) is the same as Eq.(9), if the integer $n$ in Eq.(9) is replaced with $2^n$. This means substitution ciphers might be exponentially stronger than exclusive-or ciphers. Hence, SNK of substitution ciphers is equal to:

$$SNK_d = \frac{1}{2\ erf\left((1-2d)\sqrt{2^n}\right)}. \tag{15}$$

Figure 3 (b) shows the SNK curve of exclusive-or ciphers with respect to $n$, where $d$ is regarded as a parameter. The data block length $n$ should be more than 8, if $SNK > 2^{56}$ and the reversed-bits rate lie between 0.3 and 0.7.

c) Transposition cipher

There are $n!$ transposition ciphers of $n$-bit block in all. Since an inverse of a transposition cipher and a composite transformation of two transposition ciphers are transposition ciphers, the transformation $D_{K2}E_{K1}$ is another transposition cipher. An example of $D_{K2}E_{K1}$ is illustrated by Fig.5. In the figure, the integer $h$ is the number of bits permutated actually in the product transposition. The reversed-bits rate of the product transposition cipher is

$$r = \frac{h}{2n} < \frac{1}{2}. \tag{16}$$

The total number of transposition ciphers whose $h$ bits are actually permuted is

$$\binom{n}{h} D_h .$$

The symbol $D_h$ means

$$D_h = \sum_{j=o}^{h} (-1)^j\ {}_hP_{h-j} . \tag{17}$$

In other words, $D_h$ is the number of transpositions $(a_1, a_2, \ldots, a_h)$ of $(1, 2, ..., h)$ such that $a_1 \neq 1, a_2 \neq 2, \ldots, a_h \neq h$. When $h$ is large enough, $D_h$ is approximately equal to $h!/e$:

$$\frac{D_h}{h!} \longrightarrow e^{-1} \quad (h \longrightarrow \infty). \tag{18}$$

When $h > 5$, $D_h/h!$ coincides with $e^{-1}$ more than 2 digits. The probability of $r < d$ is obtained by

$$Q_d = \frac{1}{n!} \sum_{h=0}^{2dn} \binom{n}{h} D_h. \tag{19}$$

Though $D_h$ is not equal to $h!/e$ if $h$ is small, we can ignore it in Eq.(19), because then both $\binom{n}{h}$ and $D_h$ are much smaller than that of other terms and so is $h!/e$. Therefore,

$$Q_d \simeq \sum_{h=0}^{2dn} \frac{e^{-1}}{(n-h)!} = F\big((1-2d)n, 1\big), \tag{20}$$

where $F$ is Poisson distribution:

$$F(x, \lambda) = \sum_{k>x} \frac{\lambda^k e^{-\lambda}}{k!}. \tag{21}$$

Since Poisson distribution can be approximated by Gaussian distribution, SNK is approximately

$$SNK_d \approx e \lfloor (1-2d)n \rfloor !. \tag{22}$$

The symbol $\lfloor x \rfloor$ denotes the maximum integer not greater than $x$. Figure 3 (c) shows the SNK curve of exclusive-or ciphers with respect to $n$, regarding $d$ as a parameter. The data block length $n$ should be more than 45, if $SNK > 2^{56}$ and the reversed-bits rate lie between 0.3 and 0.7.

### d) Transposition & Exclusive-or ciphers

The substantial number of keys in a product cipher of a transposition cipher and an exclusive-or cipher is calculated as an example of SNK in product ciphers. The product cipher has $2^n\, n!$ keys in all. Although this product cipher is simple, it is rather important in radio transmission, for instance, because it is the general form with no error propagation[3]. That is, the decryption process does not expand errors occurred in transmission, and the cipher with no error propagation is only the transposition and exclusive-or product cipher.

The composite transformation of the encryption with key $K1$ and the decryption with key $K2$ is another transposition and exclusive-or transformation. Figure

6 shows an example of the product cipher $D_{K2}E_{K1}$. In the figure, the reversed-bits rate is

$$r = \frac{a + \frac{h}{2}}{n}. \tag{23}$$

The integer $h$ is the number of actually permutated bits, and $a$ is the number of 1's in $P$ that are not permutated. The total number of transposition and exclusive-or ciphers which have $h$ bits permutated actually and $a$ bits of 1's in $P$ as just described, is given by:

$$\binom{n}{h} D_h \cdot \binom{n-h}{a} 2^h .$$

Using Eq.(18), the total number equals to

$$\frac{2^h \, n!}{e\,(n-h-a)!\,a!} .$$

Hence, the probability of $r < d$ or $r > 1 - d$ is

$$\begin{aligned} Q_d &\simeq 2e^{-1} \sum_{a+\frac{h}{2}<dn} \frac{1}{(n-h-a)!\,a!\,2^{n-h}} \\ &= 2e^{-1} \sum_{l=(1-2d)n}^{n} \sum_{\substack{j=l \\ j-l:even}}^{n} \frac{1}{\frac{j+l}{2}!\,\frac{j-l}{2}!\,2^j} \\ &\approx 2e^{-1} \sum_{l=(1-2d)n}^{n} \frac{1}{2^l\,l!} \\ &\approx \frac{2}{e\,2^{(1-2d)n}\lfloor(1-2d)n\rfloor!} . \end{aligned} \tag{24}$$

Here, the second and third $\approx$ hold because the terms corresponding with $j = l$ and $l = (1 - 2d)n$ are much larger than other terms. Therefore, SNK of the transposition and exclusive-or cipher is obtained by

$$SNK_d \approx e\,2^{(1-2d)n-1}\lfloor(1-2d)n\rfloor! . \tag{25}$$

The length of SNK of the transposition and exclusive-or ciphers, $k_{T\&E}$, is nearly equal to

$$k_{T\&E} \approx k_T + (1-2d)n - 1 , \tag{26}$$

where $k_T$ indicates the length of SNK of the transposition cipher. This shows the SNK length of the transposition cipher increases owing to exclusive-or of bit pattern $P$. Figure 3 d) illustrates the SNK. The data block length of the transposition and exclusive-or ciphers should be more than 37, when SNK is more than $2^{56}$ and the reversed-bits rate lies between 0.3 and 0.7.

## III . BOUNDARY OF SNK

The substantial number of keys are closely related with sphere packing. In this section, boundary of SNK is given with the number of spheres packed in key spaces. Though the difference defined by Eq.(3) does not necessarily constitute *distance* in key spaces, the key spaces are assumed to be metric spaces in this section. The differences in exclusive-or ciphers or nonlinear feedback shift register stream ciphers[1], for instance, are provèd to be *distance*.

Sphere packing is to pack as many spheres in thp key space as possible. The maximum number of spheres of diameter $d$, that is the number of keys of the sphere packing cipher $N_d$, is less than or equal to $SNK_{\frac{d}{2}}$:

$$N_d \le SNK_{\frac{d}{2}} \,. \tag{27}$$

This inequality may be considered as nearly equal. However, $N_d$ is much larger than $SNK_d$ in general, because the radius of the sphere is $d/2$:

$$N_d \gg SNK_d \,. \tag{28}$$

Hence, $SNK_d$ is bounded by:

$$N_{2d} \le SNK_d \ll N_d \,. \tag{29}$$

## IV . APPLICATION TO ENCRYPTION DESIGN

In encryption designs, both substantial number of keys SNK and difference $d$ (or reversed-bits rate $r$) are given as design parameters. When $SNK = 2^{56}$ and the reversed-bits rate is more than 0.3 and less than 0.7 ($d = 0.3$), for example, Fig.3 shows the block size $n$ should be

$$n_E \ge 500$$

$$n_S \ge 9$$

$$n_T \ge 46$$

$$n_{T\&E} \geq 38\,.$$

Under these SNK conditions, one can pick up any key in the key space as an encryption key. One does not have to select special keys. An arbitrary $n$-bit pattern $P$ can be used as a key in the exclusive-or cipher. You don't have to worry about an eavesdropper happening to pick up a decipher key close to the right key, because the probability is less than $SNK^{-1} = 2^{-56}$.

The sphere packing ciphers have to satisfy the SNK condition too. Though $N_d$ is the number of keys of the ciphers, the condition $N_d \geq 2^{56}$ is not enough. The ciphers must also satisfy $SNK_d \geq 2^{56}$. Otherwise, the key picked up by an eavesdropper, which is not necessarily the key of this scheme, is close to the right key with probability greater than $2^{-56}$. This shows the condition $N_d \geq 2^{56}$ is meaningless. Eq.(28) shows $SNK_d$, not $N_d$, is critical.

DES probably satisfies the SNK condition, because SNK of DES is much larger than $2^{56}$. SNK of DES is approximately given by $2\,e^{(2^{17}(1-2d)^2)}/\pi$ using EQ.(15), if DES is treated as a huge substitution cipher. When DES is considered as a product cipher, SNK would be less than that, but much larger than $2^{56}$, though actual calculation is very complicated.

The SNK condition is useful when one wishes to construct a rather simple encryption scheme by the combination of fundamental ciphers.

## V . CONCLUSION

The substantial number of keys, SNK, is defined and illustrated with examples of fundamental ciphers and a product cipher. SNK is one of the encipherment strength criteria. In encryption designs, SNK is used to condition design parameters. The SNK is useful for designs of product cipher in particular.

I would like to thank Mr. Nakamura and Ms. Tanaka for lots of helpful discussions.

## REFERENCES

[1] K. Nakamura,"On Self-synchronization Encryption Systems," *24th Allerton Conf.*, pp.1057-1063, 1986, (also in Proc. of SITA'80, pp.371-377, in Japanese).

[2] E. Okamoto and K. Nakamura,"Permutation Ciphers Based on Reed-Solomon Codes," *1983 IECE Conf.*, pp.1.463-1.464 (in Japanese).

[3] E. Okamoto and K. Nakamura,"Relation between Error Propagation and Nonlinearity in Cryptosystems," *1985 IECE Conf.*, p.6.27 (in Japanese).

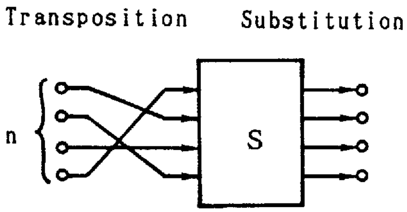
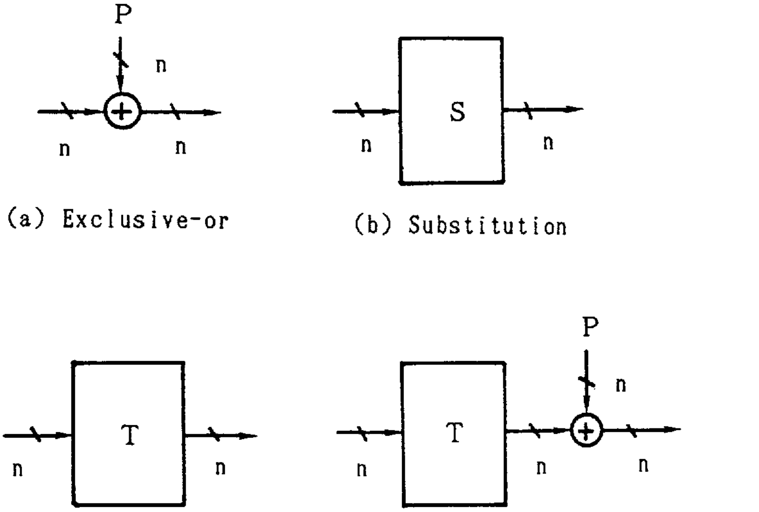


Fig. 1 Product Cipher



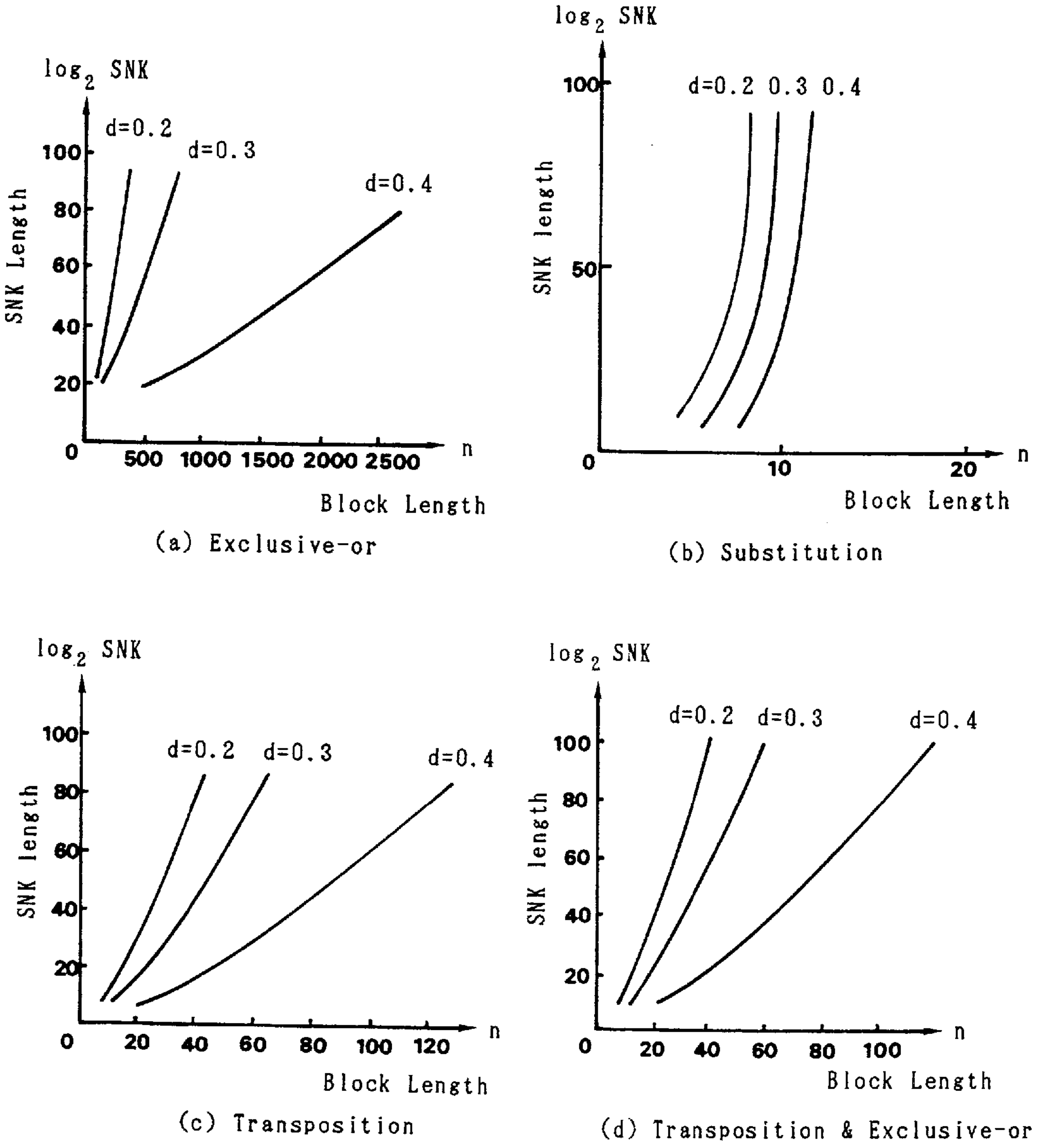
(a) Exclusive-or

(b) Substitution

(c) Transposition

(d) Transposition & Exclusive-or

Fig. 2 Examples of Cipher



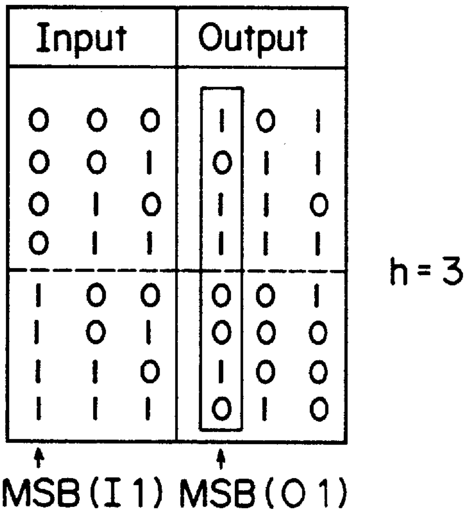
(a) Exclusive-or

(b) Substitution

(c) Transposition

(d) Transposition & Exclusive-or

Fig.3 Examples of SNK



| Input | | | Output | | |
|---|---|---|---|---|---|
| 0 | 0 | 0 | 1 | 0 | 1 |
| 0 | 0 | 1 | 0 | 1 | 1 |
| 0 | 1 | 0 | 1 | 1 | 0 |
| 0 | 1 | 1 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 | 0 | 1 |
| 1 | 0 | 1 | 0 | 0 | 0 |
| 1 | 1 | 0 | 1 | 0 | 0 |
| 1 | 1 | 1 | 0 | 1 | 0 |

h = 3
MSB(I 1)
MSB(O 1)

Fig. 4 Substitution Cipher

$D^{S}_{K2}\ E^{S}_{K1}$

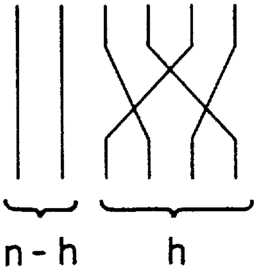


Fig. 5 Transposition Cipher

$D^{T}_{K2}\ E^{T}_{K1}$

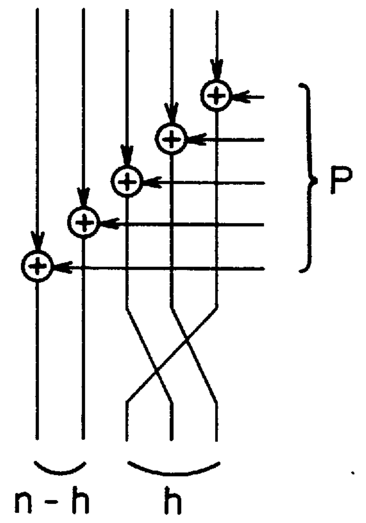


Fig. 6  Product Cipher $D_{K2}^{T\&E}\ E_{K1}^{T\&E}$