Cryptanalysis of Substitution-Permutation Networks Using Key-Dependent Degeneracy*

Howard M. Heys

Electrical Engineering, Faculty of Engineering and Applied Science
Memorial University of Newfoundland
St. John's, Newfoundland, Canada A1B 3X5

Stafford E.Tavares

Department of Electrical and Computer Engineering
Queen's University
Kingston, Ontario, Canada K7L 3N6

* This research was supported by the Natural Sciences and Engineering Research Council of Canada and the Telecommunications Research Institute of Ontario and was completed during the first author's doctoral studies at Queen's University.

Cryptanalysis of Substitution-Permutation Networks Using Key-Dependent Degeneracy

Keywords — Cryptanalysis, Substitution-Permutation Network, S-box

Abstract — This paper presents a novel cryptanalysis of Substitution-Permutation Networks using a chosen plaintext approach. The attack is based on the highly probable occurrence of key-dependent degeneracies within the network and is applicable regardless of the method of S-box keying. It is shown that a large number of rounds are required before a network is resistant to the attack. Experimental results have found 64-bit networks to be cryptanalyzable for as many as 8 to 12 rounds depending on the S-box properties.

I. Introduction

The concept of Substitution-Permutation Networks (SPNs) for use in block cryptosystem design originates from the "confusion" and "diffusion" principles introduced by Shannon [1]. The SPN architecture considered in this paper was first suggested by Feistel [2] and consists of rounds of non-linear substitutions (S-boxes) connected by bit permutations. Such a cryptosystem structure, referred

to as LUCIFER[1] by Feistel, is a simple, efficient implementation of Shannon's concepts.

A general N-bit SPN is composed of R rounds of $n \times n$ S-boxes. We shall denote the cryptosystem plaintext input as $\mathbf{P} = P_1 ... P_N$ and the ciphertext output as $\mathbf{C} = C_1 ... C_N$. S-boxes in the network are defined as a mapping $S_{rs} : \mathbf{X}_{rs} \rightarrow \mathbf{Y}_{rs}$ where $\mathbf{X}_{rs} = X_1^{(rs)} ... X_n^{(rs)}$ and $\mathbf{Y}_{rs} = Y_1^{(rs)} ... Y_n^{(rs)}$. The value of s identifies the number of the S-box within round r, $1 \leq r \leq R$, and $1 \leq s \leq N/n$. A simple example of an SPN is illustrated in Figure 1 with $N = 16$, $R = 4$, and $n = 4$.

We shall consider S-boxes that are keyed using one of the following methods[2]:

1. *selection keying:* the key bits may be used to select which S-box mapping from a set of mappings will be used for a particular S-box, and
2. *XOR mask keying:* the key bits may be exclusive-ORed with the network bits prior to entering an S-box.

Recent cryptanalysis techniques have had a notable effect on the perceived security of SPN cryptosystems. For example, in [6] and [7], Biham and Shamir introduce a powerful chosen plaintext cryptanalysis technique referred to as differential cryptanalysis. Utilizing highly probably occurrences of differential sequences,

[1] Another variant of LUCIFER [3] more closely resembles the network structure of DES [4].

[2] Note that method 2 may actually be considered as a special case of method 1. We distinguish between the two methods for clarity. Using method 2 only is a way of ensuring that a mapping for a particular S-box is selected from the same cryptographic equivalence class [5].

probabilities can be assigned to possible key values with the most probable key being selected as correct. As well, in [8], Matsui introduces the known plaintext attack of linear cryptanalysis which makes use of the likely satisfaction of linear equations involving the plaintext, ciphertext, and key bits. The applicability of differential and linear cryptanalysis to SPNs is thoroughly discussed in [9].

The cryptanalysis presented in this paper is an efficient technique for determining the network key bits. It uses a divide-and-conquer approach by examining the ciphertexts corresponding to a number of chosen plaintexts and counting the number of times a particular sub-key is consistent with a key-dependent degeneracy in the observed ciphertext. Depending on the number of rounds in the network, the correct sub-key is consistent with a significantly higher probability than the incorrect sub-keys.

II. Terminology

The following terminology is fundamental to the understanding of the cryptanalysis.

Degenerate Function: An m-input boolean function, $f(\mathbf{X})$, $\mathbf{X} = X_1 ... X_m$, is a *degenerate function* in input X_i if changing X_i only does not change the function output for all possible inputs $\mathbf{X} \in \{0,1\}^m$.

Degenerate Mapping: A $m \times n$ mapping is a *degenerate mapping* in input X_i if changing X_i only does not change the mapping output for all possible inputs

$\mathbf{X} \in \{0,1\}^m$.

Target S-box: A *target S-box* is the S-box under examination within the network. The cryptanalysis targets one S-box at a time in order to find the key bits associated with that S-box.

Target Sub-Key: The key bits associated with the target S-box are referred to as the *target sub-key*.

Ciphertext Sub-Block: A *ciphertext sub-block* is a block of n ciphertext bits which are associated with a particular S-box in the last round of the network. These may or may not be contiguous in the output block depending on whether there is a final permutation after the last round of S-boxes. There are N/n sub-blocks in a ciphertext block.

Sub-Block Mapping: A *sub-block mapping* is generated by considering a mapping from the n output bits of the target S-box to an n-bit ciphertext sub-block. A partial sub-block mapping of dimension $m \times n$ is a mapping from a subset of m output bits of the target S-box to an n-bit ciphertext sub-block.

III. Key-Dependent Degeneracy

The cryptanalysis exploits the highly probable occurrence of degeneracy in sub-block mappings. In general, if an $n \times n$ mapping is randomly selected, there is a non-zero probability that it is degenerate. It will be shown that sub-block mappings within the network often have a much higher probability of

being degenerate than that of a randomly selected mapping. In such cases, by maintaining a count of the occurrence of such degeneracies for each possible target sub-key, the correct sub-key can be derived with high probability. We refer to the consistent occurrence of degeneracy for the correct target sub-key as *key-dependent degeneracy*. Key-dependent degeneracy is very high in networks with a small number of rounds and decreases as the number of S-box rounds is increased.

In the most difficult cryptanalysis scenario, each S-box in the network has a number of associated key bits that are independent of the other key bits in the network. The cryptanalysis begins by selecting a target S-box in the first round of the network. An appropriate number of chosen plaintexts are selected so that the target sub-key may be determined with reasonable statistical confidence. Subsequently, the remaining first round S-boxes are targeted and the associated key bits determined. Once the first round key is known, the appropriate partial encryption can be used in targeting the second round of S-boxes with chosen inputs. The attack may proceed by stripping off rounds of S-boxes as their key bits are determined. As the unknown portion of the network decreases in size, the number of required chosen plaintexts to discover the target sub-key decreases significantly.

Consider the target S-box to be in the first round. In general, an SPN may be represented by the first round S-boxes, the last round S-boxes, and an inner network, as in Figure 2. The input and output to the inner network are denoted by

$\mathbf{U} = U_1 ... U_N$ and $\mathbf{V} = V_1 ... V_N$, respectively. The attack to determine the target sub-key consists of a number of trials, each trial entailing 2^n chosen plaintexts. The plaintexts in a trial are selected such that the network inputs which are not inputs to the target S-box are arbitrarily fixed and the n inputs to the target S-box are cycled through all 2^n possibilities. In this scenario, the output of the target S-box forms an n-bit input into N boolean functions corresponding to each output of the inner network. The n-input boolean function, f_i, corresponding to output V_i, is arbitrarily determined by the $N - n$ fixed inputs of $\mathbf{U}$ (coming from the outputs of the non-target S-boxes). For inner networks with a small enough number of rounds, f_i has a significant probability of being degenerate in one or more bits.

If function f_i has a high probability of being degenerate in a particular input, U_j, then there is a high probability that all n inputs to a last round S-box are degenerate in U_j as well. When this occurs, the input to the last round S-box is a degenerate mapping from the target S-box output and the corresponding sub-block mapping from the target S-box output to the ciphertext sub-block will be degenerate. However, since the target sub-key is unknown, the outputs of the target S-box and, hence, the n inputs to the sub-block mapping are not known. Therefore, there is a set of K_t possible mappings for each sub-block where K_t represents the number of possible target sub-key values. One of these mappings corresponds to the correct sub-key and is the actual sub-block mapping.

Each trial, consisting of 2^n chosen plaintexts, may be considered conceptually as

illustrated in Figure 3. Assume that the target sub-key consists of one bit used to select between S-boxes S' and S''. The output of the target S-box is mapped to a ciphertext sub-block, $\mathbf{Z}$, through $\widehat{S} : \mathbf{Y} \rightarrow \mathbf{Z}$. There are two possible values for $\widehat{S}$, denoted $\widehat{S}'$ and $\widehat{S}''$, corresponding to S' and S'' respectively. The actual mapping of $\widehat{S}$ corresponding to the correct sub-key is selected arbitrarily for each trial according to the fixed network inputs. The correct sub-key may be deduced by executing several trials and counting the number of times $\widehat{S}'$ and $\widehat{S}''$ are degenerate. We expect (and experimental results confirm) that the correct sub-key will typically exhibit mapping degeneracies most often. The number of trials (and hence chosen plaintexts) required to determine the sub-key should be enough to allow the degeneracy counts to clearly distinguish the correct target sub-key.

Example: The target S-box is selected by a key bit to be either S' or S''. The results of one trial are listed in Table 1: the outputs of one sub-block corresponding to the target S-box inputs (the remaining network inputs having been arbitrarily fixed) are given along with the possible target S-box outputs corresponding to S' and S''. From this information, Table 2 is compiled to conveniently display sub-block mapping possibilities $\widehat{S}'$ and $\widehat{S}''$. It is obvious that $\widehat{S}'$ is a degenerate mapping in input Y_3 and that $\widehat{S}''$ is not degenerate.

IV. Enhancement of the Attack Using Partial Mappings

The success of the cryptanalysis can often be enhanced by considering the

target S-box input $X_1X_2X_3X_4$	S' *output* $Y_1Y_2Y_3Y_4$	S'' *output* $Y_1Y_2Y_3Y_4$	*sub-block output* $Z_1Z_2Z_3Z_4$
0000	0100	1101	1010
0001	0001	1000	0001
0010	1110	1010	0110
0011	1000	0001	0011
0100	1101	0011	0000
0101	0110	1111	1010
0110	0010	0100	0111
0111	1011	0010	1110
1000	1111	1011	0000
1001	1100	0110	0110
1010	1001	0111	1110
1011	0111	1100	1001
1100	0011	0000	0001
1101	1010	0101	0011
1110	0101	1110	1001
1111	0000	1001	0111

Table 1. Key-Dependent Degeneracy Example

degeneracy of partial outputs of the target S-box. For example, a network with 4×4 S-boxes which displays significant key-dependent degeneracy in the 4×4 sub-block mapping of the target S-box output to ciphertext sub-block will also display these degeneracy traits when considering a 2×4 or 3×4 sub-block *partial mapping*. A partial mapping is a mapping from a group of 2 or 3 target S-box outputs to the ciphertext sub-block. The same set of chosen plaintexts used to

target S-box output $Y_1Y_2Y_3Y_4$	*sub-block output for* S' $Z_1Z_2Z_3Z_4$	*sub-block output for* S'' $Z_1Z_2Z_3Z_4$
0000	0111	0001
0001	0001	0011
0010	0111	1110
0011	0001	0000
0100	1010	0111
0101	1001	0011
0110	1010	0110
0111	1001	1110
1000	0011	0001
1001	1110	0111
1010	0011	0110
1011	1110	0000
1100	0110	1001
1101	0000	1010
1110	0110	1001
1111	0000	1010

Table 2. Sub-block Mappings Corresponding to Sub-keys

examine the full 4×4 sub-block mapping is also easily analyzed for degeneracies in the partial mappings.

When considering partial mappings from a trial of 16 chosen plaintexts, the bits that are not included as part of the mapping under examination must be fixed. Hence, for any 3 bits of the target S-box output, there are two 3×4 mappings to be examined: one corresponding to the fourth bit equal to 0 and one corresponding to

target S-box input $X_1X_2X_3X_4$	*S' output* $Y_1Y_2Y_3Y_4$	*sub-block output* $Z_1Z_2Z_3Z_4$
0011	1000	0011
0111	1011	1110
1010	1001	1110
1101	1010	0011

Table 3. Partial Mapping Example

the fourth bit equal to 1. Since there are four 3–bit groups, over all 16 target S-box inputs we have a total of eight 3×4 sub-block mappings to consider. Similarly, each 2 bit combination of outputs generates four 2×4 mappings, corresponding to the four possible values of the 3rd and 4th bits. With six ways of selecting the two outputs to consider from the target S-box, there are a total of 24 2×4 mappings. In general, for an m-bit partial output, there are $\binom{n}{m}2^{n-m}$ possible $m \times n$ mappings to be examined for degeneracy from a trial of 2^n plaintexts.

Example: Consider the example of Table 1. A portion of the table is reproduced in Table 3 in order to illustrate a case where, if the first 2 inputs to the sub-block mapping for S' are fixed at $Y_1Y_2 = 10$, the 2×4 sub-block mapping $Y_3Y_4 \rightarrow Z_1...Z_4$ is degenerate in Y_3.

Often, the correct sub-key can be easily distinguished with fewer plaintext-ciphertext pairs by analyzing partial mappings rather than the full sub-block mapping. Although randomly selected mappings with fewer inputs have a higher

probability of being degenerate, in many cases the key-dependent degeneracy is significant enough to allow identification of the correct key.

V. Effectiveness of the Algorithm

In general, it is hard to derive explicitly the complexity or the probability of success of the attack. The effectiveness of the cryptanalysis depends largely on the properties of the S-boxes and the permutations used. In analyzing the attack, it is of interest to determine (1) the likelihood that different target sub-keys cannot be distinguished and (2) the likelihood of the inner network being degenerate with a probability significantly greater than is expected for a randomly selected mapping. If we cannot distinguish between the correct sub-key and all incorrect sub-keys or if degeneracy occurs with the same frequency as expected in a random mapping, then the cryptanalysis will be unsuccessful.

Distinguishing Between Keys

It is quite possible that a particular trial will display degeneracies for the sub-block mappings of different sub-keys, one of which may or may not be the correct sub-key. The success of the attack relies on the correct sub-key displaying degeneracy more often than incorrect sub-keys. Assuming that the probability of degeneracy is large and a suitable number of chosen plaintexts is available, only under exceptional circumstances will it be impossible to distinguish between the correct key and an incorrect key. The relationship between S-box mappings which

will allow this to occur and the subsequent likelihood of randomly selected S-boxes being indistinguishable is given in the following theorem and corollary.

Theorem 1: Two $n \times n$ bijective S-boxes, S' and S'', will be indistinguishable if, and only if, each boolean function of S' is identical to a boolean function or the complement of a boolean function of S''.

Proof (Sketch):

Let two functions of S' and S'' be defined to be *similar* if they are identical or one is the complement of the other. Changes in the output of similar functions occur for the same input changes. Assume that S' and S'' are related as stated in the theorem. Then, for any subset of the function of S'', all functions in the subset are similar to output function of S'' and since degeneracies are detected based on changes in the cipertext sub-block, any degeneracies which are observed can be associated with both S-boxes and the S-boxes cannot be distinguished.
Consider the case now where one of the boolean functions of S', f_i, is not similar to a function from S''. Then in the scenario where S' is used for the cipher and the sub-block mapping is degenerate in all inputs other than the input corresponding to f_i, there is no degeneracy of S'' that is equivalent to this degeneracy of S' and the S-boxes can be distinguished. Hence, in order for S-boxes to be indistinguishable, they must be of the format suggested by the theorem. □

Corollary 1: The probability of two randomly selected $n \times n$ bijective mappings being indistinguishable is given by

$$\frac{2^n n!}{2^n!}. \tag{1}$$

Proof:

The number of possible mappings for S'' that are indistinguishable from S' is simply given by the number of ways of selecting, for all n functions of S', either the function or its complement and permuting the n functions within the mapping. This is divided by the number of possible bijective mappings to give (1) above.□

From Corollary 1, it is apparent that, if an S-box is keyed by selecting between two randomly selected mappings and assuming a sufficient number of chosen plaintexts to allow distinguishing, it is very unlikely that the two S-boxes will be indistinguishable and it will only occur for the constrictive relationship of Theorem 1. For example, if $n = 4$, the probability of two randomly selected S-boxes being indistinguishable is 1.84×10^{-11}.

Degeneracy in Random Mappings

Success of the cryptanalysis requires that the probability of degeneracy for the full and partial sub-block mappings is significantly different than the degeneracy of a random mapping so that the correct sub-key is obvious for the number of chosen plaintexts available. It is of interest therefore to determine the probability of a randomly selected $m \times n$ mapping being degenerate. As the number of rounds in

the network increase, the probability of degeneracy approaches this value and it becomes infeasible to distinguish the correct sub-key from wrong sub-keys.

Theorem 2:

The probability of a randomly selected $m \times n$ mapping being degenerate in one or more inputs is given by:

$$P_{deg} = \sum_{k=1}^{m} (-1)^{k+1} \binom{m}{k} \cdot P_{nk} \tag{2}$$

where P_{nk} represents the probability of the mapping being degenerate in k particular inputs and is given by:

$$P_{nk} = \prod_{i=1}^{k} (1/2)^{n \cdot 2^{m-i}}. \tag{3}$$

Proof:

To see how (2) is derived, consider first the probability of a random m-input boolean function, $f(X_1, ..., X_m)$, being degenerate. Since each output of the function is independently selected to be either 0 or 1 with a probability of 1/2, the probability that a change in only input X_i not causing a change in the output over all $\mathbf{X} \in \{0,1\}^m$ is given by $P_{11} = (1/2)^{2^{m-1}}$. This is derived by considering 2^{m-1} pairs of outputs in the truth table, each pair corresponding to values of $\mathbf{X}$ differing in only X_i.

The probability of the function being degenerate in two inputs, X_i and X_j, is given by the probability of being degenerate in X_i multiplied by the probability

mapping size	*probability of degeneracy*
2x4	7.5684 E-03
3x4	4.5601 E-05
4x4	9.3130 E-10

Table 4. Probability of Degeneracy for Random Mappings

of being degenerate in X_j given degeneracy in X_i. Given that the function is degenerate in X_i, since there is no change in the output when only X_i changes, we need only consider half the function output values (for example, when $X_i = 0$). Hence, if the function is to be degenerate in X_j given that it is already degenerate in X_i, there are only $2^{m-1}/2$ cases for which a change in only input X_j must not change the output. Therefore, the probability of the function being degenerate in two particular input bits is given by $P_{12} = (1/2)^{2^{m-1}} \cdot (1/2)^{2^{m-2}}$. The remaining cases for degeneracy in more than two inputs can be derived similarly and in general

$$P_{1k} = \prod_{i=1}^{k} (1/2)^{2^{m-i}}.$$

Considering now the random $m \times n$ mapping, since all output functions of the mapping are independent, the probability of all n outputs being degenerate in k particular inputs is given by (3). Using the principle of inclusion-exclusion from set theory [10] and noting the symmetric nature of the degeneracy, the probability of the $m \times n$ mapping being degenerate in one or more inputs is simply given by (2). □

stage r	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
stage r+1	1	17	33	49	2	18	34	50	3	19	35	51	4	20	36	52
stage r	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
stage r+1	5	21	37	53	6	22	38	54	7	23	39	55	8	24	40	56
stage r	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
stage r+1	9	25	41	57	10	26	42	58	11	27	43	59	12	28	44	60
stage r	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
stage r+1	13	29	45	61	14	30	46	62	15	31	47	63	16	32	48	64

Table 5. Permutation Used in 64x64 SPN Experiments

The probability of degeneracy for different size mappings is given in Table 4.

VI. Experimental Results

This section highlights some of the results of the cryptanalysis applied to different SPNs. We analyzed a 64×64 network comprised of 4×4 S-boxes and a permutation, listed in Table 5, selected from the class of permutations suggested by Ayoub [11]. Two types of networks were analyzed: one using S-boxes arbitrarily selected from the rows of the Data Encryption Standard (DES) S-boxes [4] and one using randomly selected bijective, non-degenerate S-boxes. A set of 16 S-boxes was selected for each network and used in all rounds. XOR mask keying was used for each round with the key bits randomly selected.

Experimental results were compiled for 4×4 and 2×4 sub-block mappings. A large number of trials was executed for each network and the probability of degeneracy, P_{deg}, was determined as a function of the number of rounds for SPNs

Round	4 x 4 Sub-block Mapping		2 x 4 Sub-block Mapping	
	DES Boxes	Random S-boxes	DES S-boxes	Random S-boxes
3	1	1	1	1
4	.9591	.9961	.7949	.9319
5	.3444	.8491	.3095	.7121
6	1.288×10^{-3}	.3189	4.054×10^{-2}	.3745
7	$< 6.25 \times 10^{-7}$	4.606×10^{-2}	9.120×10^{-3}	.1378
8	$< 6.25 \times 10^{-7}$	2.644×10^{-3}	7.643×10^{-3}	4.274×10^{-2}
9	$< 6.25 \times 10^{-7}$	8.125×10^{-5}	7.583×10^{-3}	1.664×10^{-2}
10	-	$< 6.25 \times 10^{-7}$	-	9.750×10^{-3}
11	-	$< 6.25 \times 10^{-7}$	-	8.191×10^{-3}
12	-	$< 6.25 \times 10^{-7}$	-	7.741×10^{-3}
13	-	$< 6.25 \times 10^{-7}$	-	7.622×10^{-3}

Table 6. Experimental Degeneracy Probabilities

with DES S-boxes and random S-boxes. In Figure 4 the measured degeneracy probabilities for the full 4×4 and the partial 2×4 mappings of various size networks are compared to the values expected for a random mapping from Table 4. As well, the experimental values are tabulated in Table 6. The general trend of convergence towards the random mapping values is evident in both network types. However, it is also clear that the network utilizing DES S-boxes approaches the desired asymptote much more quickly than the randomly selected S-boxes. Intuitively, this is likely due to the strong diffusion properties of the DES S-boxes. In particular, the property that, for a single input bit change, at least two output

bits change is very useful in diffusing changes through the network and thereby minimizing degeneracies. This property is unlikely to occur in randomly selected S-boxes and, therefore, it is not surprising that the frequency of key-dependent degeneracy is higher in networks where the S-boxes are simply randomly selected.

Consider the number of chosen plaintexts required to determine a first round target sub-key using the full $n \times n$ sub-block mapping. We can expect the number of sub-block mappings which must be analyzed before observing a degeneracy to be given by $1/P_{deg}$. In cases where degeneracy occurs with a much higher probability than expected for random mappings, it will typically take only a few occurrences of degeneracy to establish the correct key. Considering that analyzing a sub-block mapping requires 2^n chosen plaintexts and that there are N/n sub-blocks to be examined for every ciphertext block, we can determine the number of plaintexts required to reveal a first round target sub-key to be on the order of the nearest multiple of N/n above $\left\{2^n / \left[P_{deg} \cdot (N/n)\right]\right\}$. (Rounding up to the multiple of N/n is necessary since we cannot consider only part of a ciphertext block.) For example, based on an analysis of the full 4×4 sub-block mappings and the experimentally determined P_{deg}, a 6 round network with DES S-boxes requires on the order of 784 plaintexts and an 8 round network with random S-boxes requires on the order of 384 plaintexts. Networks with DES S-boxes composed of 5 or less rounds and networks with random S-boxes composed of 6 or less rounds only require on the order of 16 chosen plaintexts to determine

the target sub-key.

In practice, using partial mappings often requires fewer chosen plaintexts, particularly when P_{deg} becomes small. Figure 5 illustrates typical key counts for an 6 round DES network and an 8 round random S-box network based on 2×4 partial mapping key-dependent degeneracy. In both cases, it is apparent that, with 160 chosen plaintexts, the correct target sub-key is clearly distinguishable.

If the probability of degeneracy approaches the probabilities listed in Table 4, it might take many times more than one degeneracy occurrence to clearly distinguish the correct key. For example, analyzing 2×4 partial mappings, it was found that the correct target sub-key was determined in 7 of 8 experiments, each experiment using 1.6 million chosen plaintexts, for the 8 round network using DES S-boxes. Similarly, it takes on the order of 1.6 million chosen plaintexts to distinguish the correct target sub-key for the 12 round network with random S-boxes.

VII. Applicability of the Attack

Thwarting the Attack

There are a number of ways to minimize the impact of the attack. The rapid diffusion or avalanche of bit changes is effective in decreasing the probability of degeneracies occurring. There are several techniques that can be used to improve the diffusion in an SPN. These include:

(1) using larger S-boxes

(2) using S-boxes with good diffusion, and

(3) using a diffusive linear transformation (LT) as the interconnection between rounds of S-boxes.

The effect on the degeneracy probability as a function of the number of rounds in the network is illustrated in Figure 6. Experimental results are presented for randomly selected 4×4 and 8×8 S-boxes, as well as for S-boxes with good diffusion and a diffusive linear transformation (LT). Note, in particular, the dramatic effect of the diffusive linear transformation.

S-boxes which have good diffusion are capable of taking small input changes and converting these to a larger number of output changes. The S-boxes used in Figure 6 guarantee that a one bit input change will result in an output change of at least two bits.

A diffusive linear transformation can be generated by having each input to a round of S-boxes be determined by the sum of a number of output bits from the previous round output. For example, the linear transformation used in Figure 6 is derived by adding a parity bit to each bit after applying the permutation of Table 5 where the parity bit consists of the XOR sum of all bits.

It should also be noted that comparing the degeneracy probabilities of the networks based on 4×4 and 8×8 S-boxes is somewhat misleading since the expected degeneracy probabilities for random 2×4 and 2×8 mappings are different $\left(7.57 \times 10^{-3} \text{ and } 1.52 \times 10^{-5}, \text{ respectively}\right)$. However, for networks with a

small number of rounds, both random degeneracy probabilities are much smaller than the experimental probabilities and, hence, may be considered to be negligible. In this case, it is clear that the degeneracy probabilities for the SPN using 8×8 S-boxes are much smaller and therefore more chosen plaintexts are required to successfully cryptanalyze by exploiting key-dependent degeneracy.

Application of Attack to DES

In an attempt to determine the effectiveness of the attack on the Data Encryption Standard, we ran extensive simulations of the attack on DES. It was found that no degeneracy could be detected beyond 4 or 5 rounds, suggesting that DES is very resistant to this form of cryptanalysis. Other cryptanalysis techniques, such as differential or linear cryptanalysis, have been much more successful against DES.

This result is not surprising in light of the good diffusion characteristics of DES. For example, DES S-boxes have good diffusion of bit changes (since a one bit input change must result in at least a two bit output change) and the effects of bit changes are spread to a large number of S-boxes in the next round due to the expansion operation and the asymmetric S-boxes.

Comparison to Differential Cryptanalysis

Cryptanalysis of SPNs using key-dependent degeneracies takes advantage of the

weaknesses in the dynamic properties of ciphers. These weaknesses can also be typically exploited by differential cryptanalysis. It is difficult to directly compare the effectiveness of the two attacks since the exact complexity of both attacks is difficult to determine. Differential cryptanalysis, for example, requires knowledge of a highly probable characteristic in order to estimate the complexity of the attack.

We have found, however, that cryptanalysis using key-dependent degeneracy can be successful on networks of many rounds and the approach is very systematic and involves no detailed analysis of the SPN before execution. Conversely, while it is likely that differential cryptanalysis could be successful on SPNs of many rounds, performing the attack requires careful analysis of the difference distributions of the S-boxes in order to determine the most probable characteristics that may be used in executing the attack. This is not necessarily a simple process and it is conceivable that it could make differential cryptanalysis much more difficult to implement.

VIII. Conclusion

We have presented a novel cryptanalysis of an important class of private key block cryptosystems referred to as Substitution-Permutation Networks. The attack exploits the likely occurrence of key-dependent degeneracy within the network to determine target sub-keys associated with individual S-boxes. Experimental

results indicate that the cryptanalysis is very effective on networks up to a large number of rounds. As well, it is noted that strong S-box diffusion, as in the DES S-boxes, has significant effect on minimizing the success of the attack as the number of rounds in the network is increased.

References

1. C.E. Shannon, “Communication theory of secrecy systems,” *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.
2. H. Feistel, “Cryptography and computer privacy,” *Scientific American*, vol. 228, no.5, pp. 15–23, 1973.
3. A. Sorkin, “Lucifer: A cryptographic algorithm,” *Cryptologia*, vol. 8, no. 1, pp. 22–35, 1984.
4. National Bureau of Standards, “Data Encryption Standard (DES),” *Federal Information Processing Standard Publication 46*, 1977.
5. M. Sivabalan, S.E. Tavares, and L. E. Peppard, “On the design of SP networks from an information theoretic point of view,” *Advances in Cryptology: Proceedings of CRYPTO ’92*, Springer-Verlag, Berlin, pp. 260–279, 1993.
6. E. Biham and A. Shamir,”Differential cryptanalysis of DES-like cryptosystems,” *Journal of Cryptology*, vol. 4, no. 1, pp. 3–72, 1991.
7. E. Biham and A. Shamir, “Differential cryptanalysis of Snefru, Khafre, REDOC-II, LOKI, and Lucifer,” *Advances in Cryptology: Proceedings of CRYPTO ’91*, Springer-Verlag, Berlin, pp. 156–171, 1992.
8. M. Matsui, “Linear cryptanalysis method for DES cipher,” *Advances in Cryptology: Proceedings of EUROCRYPT ’93*, Springer-Verlag, Berlin, pp. 387–397, 1994.

9. H. M. Heys and S. E. Tavares, “Substitution-Permutation Networks Resistant to Differential and Linear Cryptanalysis,” to appear in *Journal of Cryptology*, 1995.
10. L. J. O’Connor, “The Inclusion-Exclusion Principle and Its Applications to Cryptography” *Cryptologia*, vol. 17, no. 1, pp. 63–79, 1993.
11. F. Ayoub, “The design of complete encryption networks using cryptographically equivalent permutations,” *Computers and Security*, vol. 2, pp. 261–267, 1982.

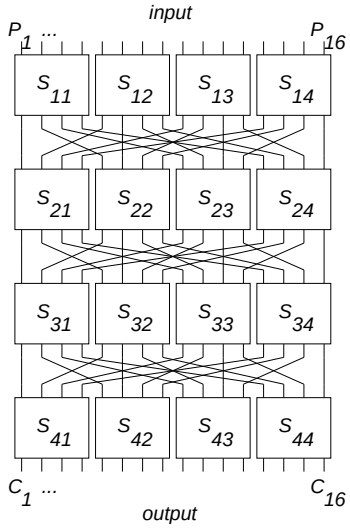

Figure 1. SPN with $N = 16$, $R = 4$, and $n = 4$

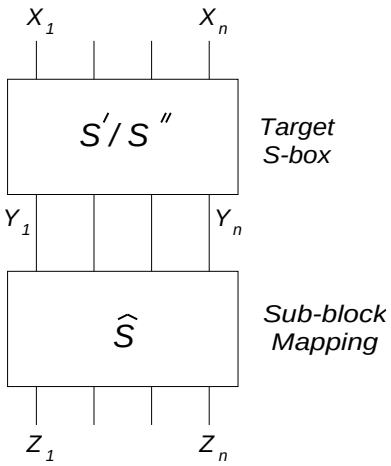

Figure 2. SPN Model Used in Attack

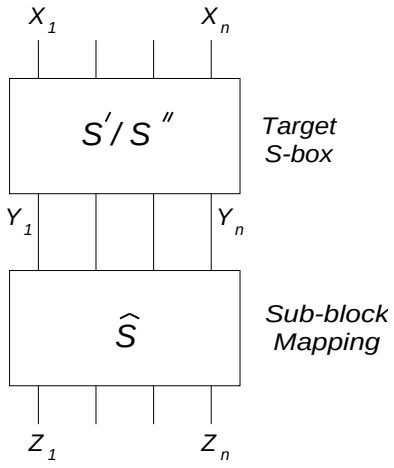

Figure 3. Sub-block Mapping Model

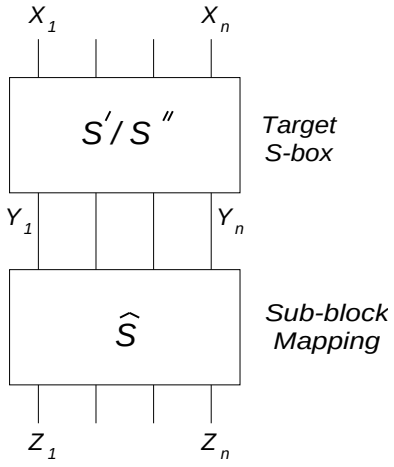

Figure 4. Experimental Degeneracy Probabilities

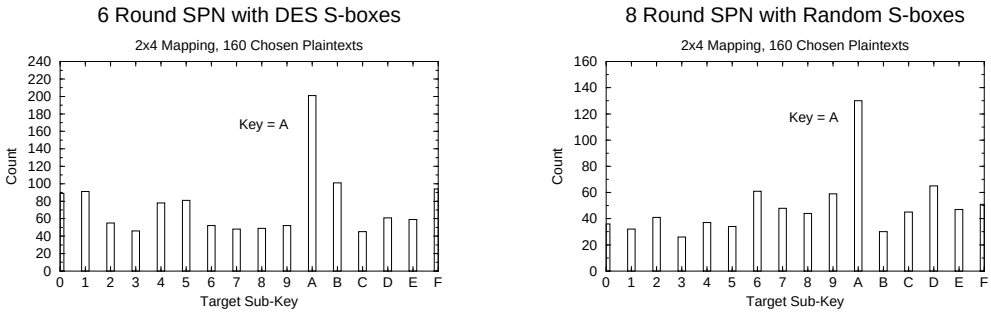

Figure 5. Typical Key Counts

References

[1] C. E. Shannon, “Communication theory of secrecy systems,” *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.

[2] H. Feistel, “Cryptography and computer privacy,” *Scientific American*, vol. 228, no. 5, pp. 15–23, 1973.

[3] A. Sorkin, “Lucifer: A cryptographic algorithm,” *Cryptologia*, vol. 8, no. 1, pp. 22–35, 1984.

[4] National_Bureau_of_Standards, “Data Encryption Standard (DES),” *Federal Information Processing Standard Publication 46*, 1977.

[5] M. Sivabalan, S. E. Tavares, and L. E. Peppard, “On the design of SP networks from an information theoretic point of view,” *Advances in Cryptology: Proceedings of CRYPTO ’92*, Springer-Verlag, Berlin, pp. 260–279, 1993.