Strict Key Avalanche Criterion

E Dawson, H Gustafson and A N Pettitt

School of Mathematics and
Information Security Research Centre
Queensland University of Technology
GPO Box 2434
Brisbane Qld 4001

Abstract.

A block cipher, such as the DES cipher, is used to encrypt binary plaintext in discrete blocks of length n using a key of length m, to form ciphertext blocks of length n. A block cipher is said to satisfy the strict key avalanche criterion if a one bit change in the key causes on the average one half of the ciphertext bit positions to change. In this paper the importance of the strict key avalanche criterion for assessing the strength of a block cipher is highlighted. Statistical based methods for determining whether a block cipher satisfies the strict key avalanche criterion are given.

1. INTRODUCTION

A block cipher is the process of encryption whereby a message is divided into a fixed number of characters (or blocks), and encrypted block by block. In the case of binary plaintext the message is divided into plaintext blocks of length n bits, denoted by P. A ciphertext block of n bits, denoted by C, is formed by applying an encryption function E to P under the influence of a key K of m bits where

$$C = E(P,K). \tag{1}$$

To recover P there is a decryption function D which is applied to C under the influence of a key K' where

$$P = D(C,K'). \tag{2}$$

There are two types of block ciphers namely symmetric and public key (asymmetric). In a symmetric cipher the keys K and K' are the same, or can be easily deduced from one another. The standard symmetric cipher is the DES (Data Encryption Standard) algorithm as described in [1]. In a public key cipher the keys K and K' are distinct, and cannot be easily deduced from one another. The standard public key cipher is the RSA (Rivest, Shamir and Adelman) algorithm as described in [1].

In this paper two properties from [7], namely plaintext avalanche effect and the key avalanche effect, which are commonly used to measure the strength of symmetric block ciphers, will be described . However they can equally be applied to measure the strength of public key block ciphers. Throughout the rest of this paper a block cipher will be assumed to be a symmetric cipher although in most cases the cipher could be public key as well, in relation to a particular property being examined.

An important property for a secure block cipher is the plaintext avalanche effect. A block cipher satisfies the plaintext avalanche effect if for a fixed key a small change in the plaintext causes a large change in the resulting ciphertext block. In [13] and [14] a more specialised property is defined namely the strict plaintext avalanche criterion, which will be denoted by SPAC. A block cipher satisfies the SPAC if for a fixed key each bit of a ciphertext block changes with the probability of one half whenever any bit of a plaintext block changes. There have been several papers in the literature containing a study of the SPAC including [5], [8], [13] and [14].

A second important property for a secure block cipher is the key avalanche effect. A block cipher satisfies the key avalanche effect if for a fixed plaintext block a small change in the key causes a large change in the resulting ciphertext block. In this paper a more specialised property is defined namely the strict key avalanche criterion, which will be denoted by SKAC. A block cipher satisfies the SKAC if for a fixed plaintext block each bit of a ciphertext block changes with the probability of one half whenever any bit of the key changes. Besides the discussion in [7] there has been little mention of the key avalanche effect in the literature.

In Section 2 a comparison will be given between the SKAC and the SPAC. The importance of using the SKAC to measure the security of a block cipher will be highlighted. In Section 3 systematic methods for testing for the SKAC under the black box assumption will be given using the Fisher-Pearson and the Kolmogorov Smirnov Tests. Under this black box assumption the person examining the security of the cipher does not know the encryption algorithm. Only the key and block lengths are known by this person who has available a box containing the cipher. It is assumed that this person is able to input keys and plaintext blocks and examine the output ciphertext blocks. As shown in [3] a similar test procedure can be used for testing for the SPAC under the black box assumption. In Section 4 the results of applying the tests from Section 3 to certain block ciphers will be given.

2. COMPARISON OF SKAC AND SPAC

If a block cipher has a poor plaintext avalanche effect in several bit positions a cryptanalyst may be able to use such information to conduct a chosen plaintext attack as shown in [17]. In this attack the cryptanalyst has one or more ciphertext blocks which he/she wants to decrypt. Under the chosen-plaintext attack assumption the cryptanalyst is able to select various plaintext blocks of his/her choice and derive the corresponding ciphertext blocks using the same key as was used to form the ciphertext blocks being attacked. This scenario could occur in the situation where a company uses a block cipher and the cryptanalyst has available an official of the company (i.e. a person thought to be trustworthy by the company but is actually working for the cryptanalyst) who can insert plaintext blocks to the block cipher using the same key as that for the ciphertext block being attacked. Suppose that the cryptanalyst wants to find the matching plaintext block for intercepted ciphertext block C. The procedure would be to select as chosen plaintext various plaintext blocks until the Hamming distance between a resulting ciphertext block and C is small where the Hamming distance between two binary vectors denotes the number of places where the vectors differ. Let P be the plaintext block which defines the ciphertext block which is close to C in Hamming distance. Using the weakness in the cipher indicated by the plaintext avalanche effect, the cryptanalyst will try to discover the correct plaintext block to match C. The method would be to find, as chosen plaintext, vectors which differ from P in only a few places.

It should be noted that the chosen plaintext attack mentioned above does not reveal the key and is conducted on a single ciphertext block at a time. In addition this attack is very difficult to implement. Nevertheless the possibility of such an attack does indicate a possible weakness in the cipher. This being the case a block cipher should only be used in the case where the cipher approximately satisfies the strict plaintext avalanche criterion.

If a block cipher has a poor strict key avalanche effect in several positions a cryptanalyst may be able to use such information to conduct a known plaintext attack in order to find the key. Under the assumption of this attack the cryptanalyst knows a few plaintext-ciphertext pairs of blocks (P,C). In addition the assumption is made that the cryptanalyst has a similar cipher. The aim of this attack is to derive the key. Success in this would enable the cryptanalyst to decrypt an entire intercepted cryptogram. The strategy of the attack would be to use a known plaintext-ciphertext pair (P,C) and to search for keys that will match P and C. The procedure would be to encrypt P with various keys until the Hamming distance between the resulting ciphertext block is small. Let K be the key which encrypts P to a ciphertext block which is close to C. Using the weakness in the cipher indicated by the dependence matrices used for examining the strict key avalanche criterion, the cryptanalyst will try to discover the correct key to match P and C. The method would be to encrypt P using keys which differ from K in only a few places.

Depending on the degree of weakness of the cipher in relation to the SKAC the above known plaintext attack can be much faster than an exhaustive key search using a known plaintext attack that does not use such information. In this fashion a weakness in the cipher in relation to the SKAC indicates a more serious structural defect in the cipher than a similar weakness in the cipher in relation to the SPAC. The former weakness may lead to the known plaintext attack described above which may greatly reduce the number of keys required to search in an exhaustive key attack. On the other hand the latter weakness only leads to the far more difficult to implement chosen plaintext attack described in [13] and [14]. A brief description of the basis of the known plaintext attack described above in relation to the SKAC is given below:

Suppose that the probability is close to zero or one that the entry in ciphertext position i changes when the only change to the key is that the entry in position j is changed. This information can help in reducing the number of keys required to test in an exhaustive key search. Assume in this fashion that the cryptanalyst has a known plaintext-ciphertext pair (P,C). The cryptanalyst encrypts P using key K resulting in ciphertext block C'. It shall be assumed that C and C' are not equal so that K is not the correct key. Let K' denote the key which differs from K in only the j position.

Firstly assume that the probability is one (or close to one) that the entry in ciphertext position i changes when the entry in key position j changes. If the entries in in position i of C and C' are the same then the cryptanalyst knows for certain (or close to certain depending on how close the above probability is to one) that key K' is impossible, without having to test this key. If the entries in position i of C and C' are not the same then the cryptanalyst will need to test the key K'.

Secondly assume that the probability that the entry in ciphertext position i changes when the entry in key position j changes is zero (or close to zero). If the entries in position i of C and C' are different then the cryptanalyst knows for certain (or close to certain depending on how close the above probability is to zero) that the key K' is impossible without having to test this key. If the entries in position i of C and C' are the same then the cryptanalyst will need to test the key K'.

The above attack strategy clearly will lead to finding the key in faster time in general than required for an exhaustive key search alone. If there are a significant number of positions i,j where the probabilty is either close to zero or one that the entry in ciphertext position i changes whenever the entry in key position j changes then this information can be useful. This could lead to a very fast known plaintext attack on the cipher.

3. STATISTICAL TESTS FOR SKAC

To test the strength of a cipher for the SKAC it is important to obtain a measure for each key-ciphertext pair. For a fixed plaintext block an n by m avalanche matrix B, as suggested in [13] and [14] (to examine SPAC), will be defined using the following procedure:

To begin the procedure select a random plaintext block P and generate a large number of random keys. Suppose that r random keys, K_k for k=1,..,r, are generated. Let K_{kj} for j=1,...,m be key vectors that differ from K_k in the j th coordinate. Let C_k and C_{kj} denote ciphertext vectors that result from encrypting P using keys K_k and K_{kj} respectively. Define avalanche vectors $U_{kj} = C_k \oplus C_{kj}$ for k = 1,...,r and j =1,...,m where $\oplus$ denotes bitwise modulo two addition of binary vectors. For k=1,...,r add the n entries of U_{kj} to each corresponding entry in column j of an n by m matrix whose initial values are all zero. After finishing the above process divide each entry of the matrix by r. This defines the n by m matrix B.

A statistical test, for analysing the avalanche matrices as defined above, will be suggested in order to determine whether a particular block cipher satisfies the SKAC. The entries b_{ij} of B are between zero and one and b_{ij} refers to the proportion of times that the entry in ciphertext position i changes when the entry in key position j changes. If the block cipher satisfies the SKAC each entry of B should be close to one half, showing only statistical variation from this value. The test to measure the SKAC examines the hypothesis that the entry in a particular ciphertext position i changes with probability of one half whenever the entry in key position j changes. This is the actual information that is needed for the mn choices of i and j to understand to what degree the known plaintext attack described in Section 2 can be avoided. The hypothesis for the test in the ij position is that the expected value of the entry is one half.

$$H_0: \; E(b_{ij}) = 0.5 \qquad i = 1, \dots n; \; j = 1, \dots m$$

The test statistic:

$$z_{ij} = 2\sqrt{r}\,(b_{ij} - 0.5) \tag{3}$$

is calculated, where i = 1,..,n and j = 1,..,m. Under H_0 this statistic defines a standard normal variable. This is a two-tailed test.

In order to gain a better understanding of the SKAC the above procedure should be repeated for a large number of random plaintext blocks, say t in total. For each of the mn entries in B the t statistics obtained can be combined to test H_0 for that entry. The Fisher-Pearson method [4] is applied to each entry using the tail area probabilities p_k , k = 1,..,t of the t statistics for entry z_{ij}, i = 1, ..., n, j = 1, ..., m.

Under H_0 the test statistic from the Fisher-Pearson method is

$$L_{ij} = -2 \sum_{k=1}^{t} (\log p_k) \,. \tag{4}$$

This is a chi-squared variable with 2t degrees of freedom. The significance level of this statistic can be used to test the hypothesis H_0 for entry i,j which defines a key-ciphertext pair.

A single measure for all mn entries in the avalanche matrices can be obtained. For each of the mn chi-squared statistics in (4) find the tail area probability of the equivalent standardised value, given t is large. The Kolmogorov-Smirnov test statistic as described in [12] is used to determine whether these mn independent probability values fit a uniform distribution in [0,1]. The significance level of the Kolmogorov-Smirnov statistic obtained is used to test the hypothesis that the block cipher satisfies the SKAC.

It should be noted that the above statistical tests for SKAC can be conducted under the black box assumption mentioned in Section 1. These tests allow one to make a decison whether a block cipher satisfies the SKAC without knowledge of the actual algorithm of the cipher.

4. EXAMPLES OF SKAC

The tests described in Section 3 were carried out for values of r = 1000 and t = 100 on the FEAL-32 [11] and Madryga [9] ciphers, and the DES (Data Encryption Standard) cipher for r = 100 and t = 100. In each of these ciphers the block size, n, is 64 bits. The key size for the FEAL-32 and Madryga ciphers is m = 64 bits, defining avalanche matrices with 4096 entries. The key size for the DES cipher is m = 56 bits, defining an avalanche matrix with 3584 entries in this case.

Test 1

Fisher-Pearson Test on Each Entry in the Avalanche Matrix

For each cell the L_{ij} statistic was calculated over the t trials and its tail area probability determined. Table 1 presents a summary of the percentage of entries in the avalanche matrix whose tail area probability is less than that stated.

Cipher	r	t	5%	1%	0.1%
FEAL-32	1000	100	4.86	0.90	0.02
Madryga	1000	100	5.64	1.22	0.17
DES	100	100	4.85	0.98	0.06

Test 2

Kolmogorov-Smirnov Test on All Entries in the Avalanche Matrix

The Kolmogorov-Smirnov statistic, D, was calculated for each of the three ciphers stated on the entries in the avalanche matrix obtained in Test 1. These statistics are compared with the 5% significance level values for the number of cells, n.

Cipher	n	D	5% critical value
FEAL-32	4096	0.0107	0.0212
Madryga	4096	0.0190	0.0212
DES	3584	0.0146	0.0227

Results

The three ciphers perform satisfactorily for both tests indicating that they all satisfy the SKAC.

6. CONCLUSION

In this paper the use of the SKAC to measure the security of a block cipher was highlighted. A comparison of the SKAC with the SPAC in Section 2 demonstrated the importance of a secure block cipher satisfying the SKAC since otherwise it may be possible to use a known plaintext attack on the cipher to derive the key. In Section 3 two statistical tests for examining the SKAC were given. These tests allow one to systematically examine the SKAC for a block cipher under the black box assumption.

The SKAC and SPAC are only two of several properties of a block cipher that can be assessed using a statistical approach. A description of other properties of a block cipher which can be assessed by using a statistical approach can be found in [3], [6], [7] and [11].

It should be noted that a block cipher may satisfy all of these properties and still be attacked in the case where the cryptanalyst knows the actual algorithm and has found some other structural weakness in the cipher. For example in [10] it is shown that in the DES algorithm after five rounds every ciphertext bit is dependent on every plaintext bit and every key bit. However, as shown in [2] it is possible to conduct a chosen plaintext attack on a six round version of the DES algorithm on a personal computer in 0.3 of a second using the characteristics of the algorithm.

REFERENCES

1. H. Beker and F. Piper, **Cipher Systems: The Protection of Communications**, JohnWiley and Sons, 1982.

2. E. Biham and A. Shamir, "Differential cryptanalysis of DES-like cryptosystems", in **Abstracts of CRYPTO 90**, Santa Barbara Calif., August 1990, pp. 1-19.

3. E. Dawson, **Design and Cryptanalysis of Symmetric Ciphers**, PhD Thesis, Queensland University of Technology, 1991.

4. L.J. Folks, "Combination of Independent Tests", in **Handbook of Statistics**, Vol. 4, Elsevier Science Publishers, 1984, pp. 113-121.

5. R. Forre, "The strict avalanche criterion: Spectral properties of Boolean functions and an extended definition" in **Advances in Cryptology: Proceedings of Crypto 88**, Springer-Verlag, 1990, pp. 450-468.

6. H. Gustafson, E. Dawson and W. Caelli, "Comparison of block ciphers" in **Advances in Cryptology: Proceedings of AUSCRYPT90**, edited J. Seberry and J. Pieprzyk, Springer-Verlag, 1990, pp. 208-220.

7. A.G. Konheim, **Cryptography: A Primer**, John Wiley and Sons, 1981.

8. S. Lloyd, "Counting functions satisfying a higher order strict avalanche criterion" in **Advances in Cryptology: Proceedings of Eurocrypt 89**, Springer-Verlag, 1989.

9. W.E. Madryga, "A High Performance Encryption Algorithm", **Computer Security**: A Global Challenge. Elsevier Science Publishers B.V., 1984, pp. 557-570.

10. C.H. Meyer and S.M. Matyas, **Cryptography: A New Dimension in Computer Data Security**, John Wiley and Sons, 1982.

11. A. Shimizu and S. Miyaguchi, "FEAL - Fast Data Encipherment Algorithm", **Systems and Computers in Japan**, Vol. 19, No. 7, 1988, pp. 20-34.

12. M.A. Stephen and R.B. D'Agostino, "Tests Based on EDF Statistics", **Goodness of Fit Techniques**, in Statistics, Textbooks and Monographs; Vol. 68, Marcel Dekker Inc., 1986, pp. 97-193.

13. A.F. Webster, **Plaintext/Ciphertext Bit Dependencies in Cryptographic Algorithms**, M.Sc. Thesis, Queen's University at Kingston, 1985.

14. A.F. Webster and S.E. Tavares, "On the design of S-Boxes" in **Advances in Cryptology**: Proceedings of CRYPTO'85, Springer-Verlag, 1986, pp. 523-530.

ACKNOWLEDGEMENT

This work was supported by an Australian Research Council Grant.

(Received 25/4/91)