Linear Cryptanalysis

Linear cryptanalysis is a powerful method of cryptanalysis introduced by Matsui in 1993 [11]. It is a known plaintext attack in which the attacker studies the *linear approximations* of parity bits of the plaintext, ciphertext and the secret key. Given an approximation with high probability and counting on the parity bits of the known plaintexts and ciphertexts one obtains an estimate of the parity bit of the key. Using auxiliary techniques one can usually extend the attack to find more bits of the secret key.

In slightly more detail: following Matsui we denote by $A[i]$ the i-th bit of A and by $A[i_1, i_2, \ldots, i_k]$ the parity $A[i_1] \oplus A[i_2] \oplus \ldots \oplus A[i_k]$. For such simple linear operations as XOR with the key or a permutation of bits very simple linear expressions can be written which hold with probability one. For non-linear elements of a cipher such as S-boxes one tries to find linear approximations with probability p that maximize the deviation $|p - \frac{1}{2}|$. Approximations for single operations inside a cipher can be further combined into approximations that hold for a single round of a cipher. For the whole cipher one receives an approximation of the type:

$$P[i_1, i_2, \ldots, i_a] \oplus C[j_1, j_2, \ldots, j_b] = K[k_1, k_2, \ldots, k_c] \tag{1}$$

(where $i_1, i_2, \ldots, i_a$, $j_1, j_2, \ldots, j_b$ and $k_1, k_2, \ldots, k_c$ denote fixed bit locations) which can be obtained by appropriate concatenation of one-round approximations. Such approximation is interesting only if it holds with $p \neq 1/2$. Matsui found such an approximation for DES with probability $\frac{1}{2} + 2^{-24}$. Using this approximation, a simple algorithm based on the maximum likelihood method can be used to find one parity bit $K[k_1, k_2, \ldots, k_c]$ of the key:

Given a pool of N random known plaintexts, let T be the number of plaintexts such that the left side of the equation 1 is zero.
if $(T - N/2) \cdot (p - 1/2) > 0$ **then**
  $K[k_1, ..., k_c] = 0$
else
  $K[k_1, ..., k_c] = 1$
end if

More efficient algorithms for linear cryptanalysis, which find more key bits are described in [11].

1 Piling-up Lemma

The first stage in linear cryptanalysis consists in finding useful approximations for a given cipher (or in demonstrating that no useful approximations exist, which is usually much more difficult). Although the most biased linear approximation can easily be found in an exhaustive way for a simple component such as an S-box, a number of practical problems arise when trying to extrapolate this method to full-size ciphers. The first problem concerns the computation of the probability

of a linear approximation. In principle, this would require the cryptanalyst to run through all possible combinations of plaintexts and keys, which is clearly infeasible for any practical cipher. The solution to this problem is to make a number of assumptions and to approximate the probability using the so-called *Piling-up Lemma*:

Lemma 1. *Given* n independent *random variables* $X_1, X_2, \ldots, X_n$ *taking on values from* $\{0, 1\}$*, then the bias* $\epsilon = p - 1/2$ *of the sum* $X = X_1 \oplus X_2 \oplus \ldots \oplus X_n$ *is given by:*

$$\epsilon = 2^{n-1} \prod_{j=1}^{n} \epsilon_j \,, \tag{2}$$

where $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$ *are the biases of the terms* $X_1, X_2, \ldots, X_n$*.*

Notice that the lemma can be further simplified by defining $c = 2 \cdot \epsilon$, known as the *imbalance* or the *linear probability* of an expression. With this notation (2) reduces to $c = \prod_{j=1}^{n} c_j$.

In order to estimate the probability of a linear approximation using the Piling-up Lemma, the approximation is written as a chain of connected linear approximations, each spanning a small part of the cipher. Such a chain is called a *linear characteristic*. Assuming that the biases of these partial approximations are statistically independent and easy to compute, then the total bias can be computed using (2).

Although the Piling-up Lemma produces very good estimations in many practical cases, even when the approximations are not strictly independent, it should be stressed that unexpected effects can occur when the independence assumption is not fulfilled. In general, the actual bias in these cases can be both much smaller and much larger than predicted by the lemma.

2 Matsui's Search for the Best Approximations

The Piling-up Lemma in the previous paragraph provides a useful tool to estimate the strength of a given approximation, but the problem remains how to find the strongest approximations for a given cipher. For DES, this open problem was solved by Matsui [13] in 1994. In his second paper, he proposes a practical search algorithm based on a recursive reasoning. Given the probabilities of the best i-round characteristic with $1 \leq i \leq n-1$, the algorithm efficiently derives the best characteristic for n rounds. This is done by traversing a tree where branches are cut as soon as it is clear that the probability of a partially constructed approximation cannot possibly exceed some initial estimation of the best n-round characteristic.

Matsui's algorithm can be applied to many other block ciphers, but its efficiency varies. In the first place, the running time strongly depends on the accuracy of the initial estimation. Small estimations increase the size of the search tree. On the other hand, if the estimation is too large, the algorithm will not

return any characteristic at all. For DES, good estimations can easily be obtained by first performing a restricted search over all characteristics which only cross a single S-box in each round. This does not work as nicely for other ciphers however. The specific properties of the S-boxes also affect the efficiency of the algorithm. In particular, if the maximum bias of the S-box is attained by many different approximations (as opposed to the distinct peaks in the DES S-boxes), this will slow down the algorithm.

3 Linear Hulls

Estimating the bias of approximations by constructing linear characteristics is very convenient, but in some cases, the value derived this way diverges significantly from the actual bias. The most important cause for this difference is the so-called *linear hull* effect, first described by K. Nyberg in 1994 [14]. The effect takes place when the correlation between plaintext and ciphertext bits, described by a specific linear approximation, can be explained by multiple linear characteristics, each with a non-negligible bias, and each involving a different set of key bits. Such a set of linear characteristics with identical input and output masks is called a *linear hull*. Depending on the value of the key, the different characteristics will interfere constructively or destructively, or even cancel out completely. If the sets of keys used in the different linear characteristics are independent, than this effect might considerably reduce the average bias of expression (1), and thus the success rate of the simple attack described above. Nyberg's paper shows however that the more efficient attacks described in [11], which only use the linear approximations as a distinguisher, will in general benefit from the linear hull effect.

4 Provable Security Against Linear Cryptanalysis

The existence of a single sufficiently biased linear characteristic suffices for a successful linear attack against a block cipher. A designer's first objective is therefore to ensure that such characteristic cannot possibly exist. This is usually done by choosing highly non-linear S-boxes and then arguing that the diffusion in the cipher forces all characteristics to cross a sufficiently high minimal number of "active" S-boxes.

The approach above provides good heuristic arguments for the strength of a cipher, but in order to rigorously prove the security against linear cryptanalysis in general, the designer also needs to take into account more complex phenomena such as the linear hull effect. For DES-like ciphers, such security proofs were studied by L. Knudsen and K. Nyberg, first with respect to differential cryptanalysis [15], and then also applied to linear cryptanalysis [14]. The results inspired the design of a number of practical block ciphers such as MISTY (or its variant KASUMI), AES, Camellia and others. Later, similar proofs were formulated for ciphers based on SP-networks [5, 7].

A somewhat more general theory for provable security against a class of attacks, including basic linear cryptanalysis, is based on the notion of decorrelation, introduced by S. Vaudenay [21]. The theory suggests constructions were a so-called *Decorrelation Module* effectively blocks the propagation of all traditional linear and differential characteristics.

An important remark with respect to the previous notions of provable security, however, is that ciphers which are provably optimal against some restricted class of attacks often tend to be weak when subject to other types of attacks [19, 22].

5 Comparison with Differential Cryptanalysis

Linear cryptanalysis has many methodological similarities with differential cryptanalysis as is noted in [1]. *Differential characteristics* correspond to *linear approximations*. *Difference distribution tables* are replaced by *linear approximation tables*. Concatenation rule for differential characteristics: "match the differences, multiply the probabilities" corresponds to concatenation rule for linear approximations (the piling-up lemma): "match the masks, multiply the imbalances". The algorithms that search for the best characteristic or the best linear approximation are essentially the same. The notion of *differentials* has a corresponding notion of *linear hulls*. Together with striking methodological similarity between the two techniques there is also *duality* [13] of operations: "XOR branch" and "three-forked branch" are mutually dual regarding their action on differences and masks respectively. Important distinction between the two methods is that differential cryptanalysis works with blocks of bits while linear cryptanalysis typically works with a single bit. The bias of the linear approximation has a sign. Thus given two approximations with the same input and output masks and equal probability but opposite signs, the resulting approximation will have zero bias, due to the cancellation of the two approximations by each other.

6 Extensions

Linear cryptanalysis technique has received much attention since its invention and has enjoyed several extensions. One technique is a combined differential-linear approach proposed by Langford and Hellman. Other extensions include *key-ranking* which allows for a tradeoff between data and time of analysis [6, 12, 17]; *partitioning cryptanalysis* [4] which studies correlation between partitions of the plaintext and ciphertext spaces (no practical cipher has been broken via this technique so far); *chi-square cryptanalysis* [9, 20] has been applied successfully against several ciphers, including round-reduced versions of RC6; the use of non-linear approximations was suggested [10, 18], but so far it provided only small improvements over the linear cryptanalysis. Full non-linear generalization still remains evasive. Idea to use multiple approximations has been expressed in [2] but no significant improvement over the basic technique has been demonstrated.

A conversion of a known plaintext linear attack to a chosen plaintext linear attack has been proposed in [8]. Finally note that similar techniques have been applied to stream ciphers (see Linear Cryptanalysis for Stream Ciphers).

–Alex Biryukov, Christophe De Cannière.

References

[1] E. Biham, "On matsui's linear cryptanalysis," in Santis [16], pp. 341–355.

[2] M. J. R. Burton S. Kaliski, "Linear cryptanalysis using multiple approximations," in Desmedt [3], pp. 26–39.

[3] Y. Desmedt, ed., *Advances in Cryptology – CRYPTO'94*, vol. 839 of *Lecture Notes in Computer Science*, Springer-Verlag, 1994.

[4] C. Harpes and J. L. Massey, "Partitioning cryptanalysis," in *Fast Software Encryption, FSE'97* (E. Biham, ed.), vol. 1267 of *Lecture Notes in Computer Science*, pp. 13–27, Springer-Verlag, 1997.

[5] S. Hong, S. Lee, J. Lim, J. Sung, D. Cheon, and I. Cho, "Provable security against differential and linear cryptanalysis for the SPN structure," in *Proceedings of Fast Software Encryption – FSE'00* (B. Schneier, ed.), no. 1978 in Lecture Notes in Computer Science, pp. 273–283, Springer-Verlag, 2000.

[6] P. Junod and S. Vaudenay, "Optimal key ranking procedures in a statistical cryptanalysis," in *Fast Software Encryption, FSE 2003* (T. Johansson, ed.), vol. 2887 of *Lecture Notes in Computer Science*, pp. 1–15, Springer-Verlag, 2003.

[7] L. Keliher, H. Meijer, and S. E. Tavares, "New method for upper bounding the maximum average linear hull probability for SPNs," in *Proceedings of Eurocrypt'01* (B. Pfitzmann, ed.), no. 2045 in Lecture Notes in Computer Science, pp. 420–436, Springer-Verlag, 2001.

[8] L. R. Knudsen and J. E. Mathiassen, "A chosen-plaintext linear attack on DES," in *Fast Software Encryption, FSE 2000* (B. Schneier, ed.), vol. 1978 of *Lecture Notes in Computer Science*, pp. 262–272, Springer-Verlag, 2001.

[9] L. R. Knudsen and W. Meier, "Correlations in RC6 with a reduced number of rounds," in *Proceedings of Fast Software Encryption – FSE'00* (B. Schneier, ed.), no. 1978 in Lecture Notes in Computer Science, pp. 94–108, Springer-Verlag, 2000.

[10] L. R. Knudsen and M. J. B. Robshaw, "Non-linear approximations in linear cryptanalysis," in *Proceedings of Eurocrypt'96* (U. Maurer, ed.), no. 1070 in Lecture Notes in Computer Science, pp. 224–236, Springer-Verlag, 1996.

[11] M. Matsui, "Linear cryptanalysis method for DES cipher," in *Advances in Cryptology – EUROCRYPT'93* (T. Helleseth, ed.), vol. 765 of *Lecture Notes in Computer Science*, pp. 386–397, Springer-Verlag, 1993.

[12] M. Matsui, "The first experimental cryptanalysis of the data encryption standard," in Desmedt [3], pp. 1–11.

[13] M. Matsui, "On correlation between the order of S-boxes and the strength of DES," in Santis [16], pp. 366–375.

[14] K. Nyberg, "Linear approximations of block ciphers," in Santis [16], pp. 439–444.

[15] K. Nyberg and L. R. Knudsen, "Provable security against a differential attack," *Journal of Cryptology*, vol. 8, no. 1, pp. 27–38, 1995.

[16] A. D. Santis, ed., *Advances in Cryptology – EUROCRYPT'94*, vol. 950 of *Lecture Notes in Computer Science*, Springer-Verlag, 1995.

[17] A. A. Selcuk, "On probability of success in differential and linear cryptanalysis," Technical report, Network Systems Lab, Department of Computer Science, Purdue University, 2002. previously published at SCN 2002.

[18] T. Shimoyama and T. Kaneko, "Quadratic relation of s-box and its application to the linear attack of full round des," in *Advances in Cryptology – CRYPTO'98* (H. Krawczyk, ed.), vol. 1462 of *Lecture Notes in Computer Science*, pp. 200–211, Springer-Verlag, 1998.

[19] T. Shimoyama, S. Moriai, T. Kaneko, and S. Tsujii, "Improved higher order differential attack and its application to Nyberg-Knudsen's designed block cipher," *IEICE Trans. Fundamentals*, vol. E82-A, no. 9, pp. 1971–1980, 1999. `http://search.ieice.or.jp/1999/files/e000a09.htm#e82-a,9,1971`.

[20] S. Vaudenay, "An experiment on DES statistical cryptanalysis," in *3rd ACM Conference on Computer and Communications Security, CCS*, pp. 139–147, ACM Press, 1996.

[21] S. Vaudenay, "Decorrelation: A theory for block cipher security," *Journal of Cryptology*, vol. 16, no. 4, pp. 249–286, 2003.

[22] D. Wagner, "The boomerang attack," in *Fast Software Encryption, FSE'99* (L. R. Knudsen, ed.), vol. 1636 of *Lecture Notes in Computer Science*, pp. 156–170, Springer-Verlag, 1999.