Methods for Linear and Differential Cryptanalysis of Elastic Block Ciphers

Debra L. Cook[1], Moti Yung[2], Angelos D. Keromytis[3]

[1] Bell Labs, New Providence, NJ, USA
dcook@cs.columbia.edu**
[2] Google, Inc. and Department of Computer Science, Columbia University, New York, NY, USA
moti@cs.columbia.edu
[3] Department of Computer Science, Columbia University, New York, NY, USA
angelos@cs.columbia.edu

Abstract. The elastic block cipher design employs the round function of a given, b-bit block cipher in a black box fashion, embedding it in a network structure to construct a family of ciphers in a uniform manner. The family is parameterized by block size, for any size between b and $2b$. The design assures that the overall workload for encryption is proportional to the block size. When considering the approach taken in elastic block ciphers, the question arises as to whether cryptanalysis results, including methods of analysis and bounds on security, for the original fixed-sized cipher are lost or, since original components of the cipher are used, whether previous analysis can be applied or reused in some manner.

With this question in mind, we analyze elastic block ciphers and consider the security against two basic types of attacks, linear and differential cryptanalysis. We show how they can be related to the corresponding security of the fixed-length version of the cipher. Concretely, we develop techniques that take advantage of relationships between the structure of the elastic network and the original version of the cipher, independently of the cipher.

This approach demonstrates how one can build upon existing components to allow cryptanalysis within an extended structure (a topic which may be of general interest outside of elastic block ciphers). We show that any linear attack on an elastic block cipher can be converted efficiently into a linear attack on the fixed-length version of the cipher by converting the equations used to attack the elastic version to equations for the fixed-length version. We extend the result to any algebraic attack. We then define a general method for deriving the differential characteristic bound of an elastic block cipher using the differential bound on a single round of the fixed-length version of the cipher. The structure of elastic block ciphers allows us to use a state transition method to compute differentials for the elastic version from differentials of the round function of the original cipher.

Key words: security analysis, linear cryptanalysis, differential cryptanalysis.

1 Introduction

Elastic block ciphers were designed to convert existing fixed-length block ciphers into variable-length block ciphers in an efficient manner. Furthermore, the design allows certain

** This work was performed primarily while the author was at Columbia University.

properties of the fixed-length cipher to remain intact in the elastic version, creating a well-defined relationship between the security of the elastic and fixed-length versions [3, 4]. Exploiting existing ciphers' components in the design of new ciphers is not uncommon. In the elastic block cipher case, since the cipher attempts to cover a large range of block sizes, a specific design for each size was traded against a general design methodology. Naturally, in a general design, as opposed to an optimized design for a specific block size, one may lose the ability to provide tight security bounds, but security analysis is required nevertheless. A natural approach when building upon existing components is to reuse the security properties of the building blocks. Thus, our work is concerned with how the security of an elastic block cipher relates to the security of the fixed-length version.

In more detail, we view elastic block ciphers as a category of block ciphers with (somewhat generic) design rules, and we consider how to evaluate their security against the two most basic types of cryptanalysis: linear [6] and differential cryptanalysis [1]. The elastic design is a generic approach that inserts the round function from an existing block cipher into a network structure (the elastic network). Therefore, new methods are needed to perform our analysis that are derived from the structure of the elastic network. Since the approach taken in forming elastic block ciphers is non-traditional in the sense that it does not focus on optimizing the design for a specific block size, one may dismiss the entire idea and stick to usual designs of ciphers of fixed size; however, we believe that the idea of having a substitution-permutation network that is size-flexible (*i.e.*, the elastic network) and is somewhat generic is an interesting subject that deserves investigation. This work is a step in this direction.

Concretely, we first prove that any linear attack on an elastic block cipher can be converted in polynomial time and memory into a linear attack on the fixed-length version of the cipher. This is done by showing how to convert the equations for such an attack on the elastic version to an attack on the fixed-length version. Therefore, if the fixed-length version is immune to linear cryptanalysis, the elastic version is also immune. We extend the result to any algebraic attack. We then define a general method for deriving the differential characteristic bound of an elastic block cipher from the differential bound on a round of the fixed-length version. We summarize our application of the method to elastic versions of AES [9] and MISTY1 [7].

The remainder of the paper is organized as follows. In Section 2, we briefly review the construction of elastic block ciphers. In Section 3, we prove that a linear attack, or more generally any algebraic attack, on an elastic block cipher implies that such an attack exists on the fixed-length version of the block cipher. In Section 4, we define our method for deriving differential bounds on an elastic block cipher. Section 5 concludes the paper.

2 Elastic Block Cipher Review

2.1 Overview

We briefly review the method presented by Cook, *et. al*, for creating elastic block ciphers [3]. The method converts the encryption and decryption functions of existing block ciphers to accept blocks of size b to $2b$ bits, where b is the block size of the original block cipher. The general structure of an elastic block cipher is shown in Figure 1. An elastic version of a block cipher is created by inserting the cycle of the original fixed-length block cipher

into the network structure to form the round function of the elastic version. In each round the leftmost b bits are processed by the round function and the rightmost y bits are omitted from the round function. Afterwards, the rightmost y bits are XORed with a subset of the leftmost b bits and the results swapped. This swapping of bits may be omitted after the last round. The elastic version also includes initial and end of round whitening, and an initial and final key dependent permutation. The number of expanded-key bits required varies based on the block size and the original block cipher. The key schedule of the original cipher is replaced with a generic key schedule that generates as many expanded-key bits as needed. In theory, the expanded key bits can take on any value and we view the expanded key bits in this manner in our analysis. For actual implementations, a stream cipher was suggested as one option for the key schedule [3].

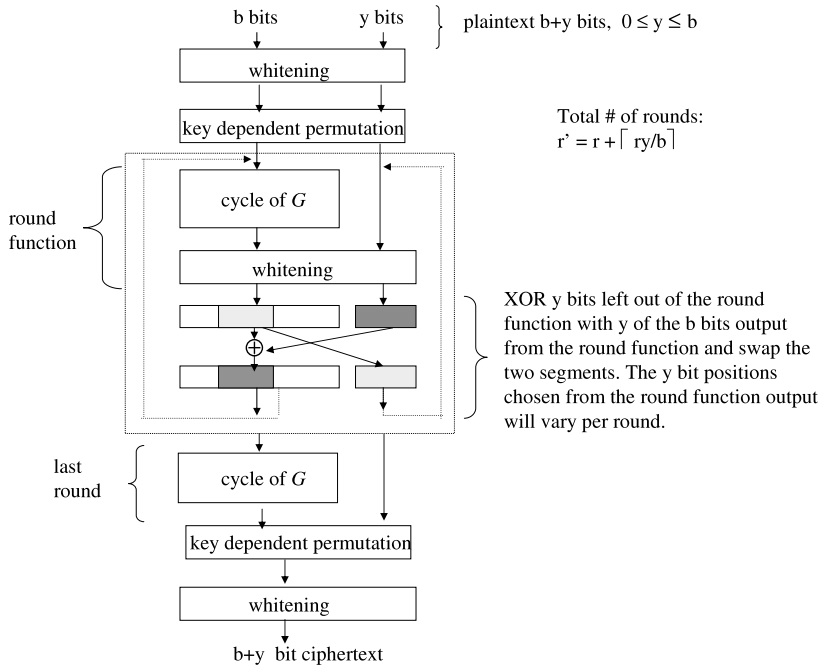

Fig. 1. Elastic Block Cipher Structure [3]

We use the following notation:

- G denotes any existing fixed-length block cipher.
- r denotes the number of cycles in G, where a cycle in G is the point at which all b bits of the block have been processed by the round function of G. For example, if G is a Feistel network, a cycle is the sequence of applying the round function of G to the left and right halves of the b-bit block. In AES, the round function is a cycle.
- b denotes the block length of the input to G in bits.
- y is an integer in the range $[0, b]$.
- G' denotes the elastic version of G with a $(b + y)$-bit input for any valid value of y.

– r' denotes the number of rounds in G'. $r' = r + \lceil \frac{ry}{b} \rceil$.
– The round function of G' will refer to one entire cycle of G.
– The swap step will refer the step in which the rightmost y bits are XORed with a subset of the leftmost b bits and the results swapped.

3 Linear Cryptanalysis

We consider linear attacks and algebraic attacks on elastic block ciphers in general. We prove that any practical linear or algebraic attack on an elastic block cipher, G', can be converted into a polynomial time related attack on the original cipher, G, independently of the specific block cipher used for G. We take advantage of the elastic block cipher structure to define a linear relationship, if one exists, across r rounds of G' in terms of any linear relationship in a cycle of G.

Linear cryptanalysis involves finding equations relating plaintext, ciphertext and key (usually expanded-key) bits via XORs that hold with probability $\frac{1}{2} + \alpha$ for non-negligible α. Without loss of generality, we assume the equations are in the form such that $0 < \alpha \leq \frac{1}{2}$, and that the equations involve the expanded-key bits. We omit the initial and final key-dependent permutations in the elastic block cipher construction when performing our analysis in order to focus on the core structure of elastic block ciphers. The two permutations do not impact any relationship that exists across the rounds of G'.

We show that a linear relationship across r rounds of G' implies such a relationship across r cycles of G. If any such linear relationship holds with a probability such that fewer than $2^{(b-1)}$ (plaintext, ciphertext) pairs are required for an attack, then G is subject to a linear attack that requires fewer plaintexts, on average, than an exhaustive search over all plaintexts. Whether or not using the equations is computationally feasible depends on number of (plaintext, ciphertext) pairs and the number of equations that must be computed. If at least $2^{(b-1)}$ plaintext, ciphertext pairs are required for an attack on r rounds of G', then either the attack is infeasible on r rounds of G' from a practical perspective or G is subject to a brute force attack in practice. Note that we are dealing with an attack on only r rounds of G' and the probability of a linear relationship holding across $r' = r + \lceil \frac{ry}{b} \rceil$ rounds of G' will be less than that for r rounds. More specifically, if the attack on G' involves a maximum correlation between plaintext, ciphertext and key bits which occurs with probability $\leq 2^{-b}$ on r rounds (thus requiring in practice $\geq 2^b$ plaintexts), then an attack on $2r$ rounds involves a maximum correlation that occurs with probability $\leq 2^{-2b}$ and requires $> 2^{2b}$ plaintexts. In this case, G' is practically secure against a linear attack when $\lceil \frac{ry}{b} \rceil = r$. A direct implication of our result is that if G' is subject to an attack using any algebraic equations, as opposed to just linear equations, then so is G.

Theorem 1. *Given a block cipher G with a block size of b bits and r cycles, and its elastic version G' with a block size of $b + y$ bits for $0 \leq y \leq b$, if G' is subject to a linear attack on r rounds then either G is subject to a linear attack or the resources exist to perform an exhaustive search on G over all plaintexts, assuming the key schedules of G and G' do not produce message-dependent expanded keys, meaning any expanded-key bits depend only on the key and do not vary based on the plaintext or ciphertext input to the cipher.*

Proof. We first note that if the linear attack on r rounds of G' requires at least 2^b (plaintext, ciphertext) pairs then either the attack is computationally infeasible or G is insecure

independent of the attack (since the attacker has the resources to encrypt 2^b plaintexts). Therefore, it can be assumed that the attack on G' requires $< 2^b$ (plaintext, ciphertext) pairs. The assumption that the expanded key bits do not depend on the input to the cipher (the plaintext or ciphertext) is true of block ciphers used in practice and of elastic block ciphers. The theorem is proved by showing how a linear attack on G' can be converted into an attack in G. With no further assumptions about the key schedules, the result is an attack that finds an expanded key for G that produces the $(plaintext, ciphertext)$ pairs consistent with G, but which may or may not adhere to the key schedule of G. If the expanded key is inconsistent with the key schedule of G, this itself indicates another weakness in G because it means there is some expanded key that is not produced by the key schedule of G but which produces the same $(plaintext, ciphertext)$ pairs that G would produce when using some key generated by G's key schedule (*i.e.* the attack finds an equivalent key). If the following three assumptions are placed on the expanded key bits of G', then the attack on G will find a key consistent with the key schedule of G:

- The rightmost y bits of each whitening step in G' can take on any value and are independent of any other expanded-key bits.
- Any expanded-key bits used in the round function of the first r consecutive rounds of G' can take on the same values as the expanded-key bits used in the cycles of G.
- If G contains initial and end of cycle whitening, any expanded-key bits used for the leftmost b bits of each whitening step in the first r consecutive rounds of G' can take on the same values as the corresponding whitening bits in G.

To understand how a linear relationship (if one exists) between the plaintext, ciphertext and expanded-key bits is determined for G', we first consider how a linear relationship is derived for a block cipher structured as a series of rounds with block length b and then add the impact of the whitening and swap step to these relationships. We number the rounds from 1 to r. We will refer to any initial whitening step that occurs prior to the first round as round 0 and the round function of round 0 is just the initial whitening. The relationship between the output of the j^{th} round/cycle and the input to the $(j+1)^{st}$ round/cycle is depicted in Figure 2 for both G and G'.

We use the following notation for describing the relationships across the rounds of G':

- Two bits, $x1$ and $x2$, cancel each other in an equation means $x1 \oplus x2 = 0$ with probability 1.
- Let u_{ji} denote the i^{th} bit of the input to the round function in round j, $1 \leq i \leq b$, $0 \leq j \leq r$.
- Let v_{ji} denote the i^{th} bit of the output from the round function in round j, $1 \leq i \leq b$, $0 \leq j \leq r$.
- Let n_j denote the number of expanded-key bits used in the round function in round j, $0 \leq j \leq r$. This does not include any end of round whitening added to form G', but does include the end of round whitening if it is part of the cycle of G (as is the case with AES). If G does not contain initial whitening, the round function in round 0 is the identity function and $n_0 = 0$.
- Let ek_{ji} denote the i^{th} expanded-key bit in the round function in round j, $1 \leq i \leq n_j$.
- Let $L_j([u_{j1}, ...u_{jb}] \oplus [v_{j1}, ...v_{jb}] \oplus [ek_{j1}, ...ek_{jn_j}])$ denote the set of linear equations (if any) relating the input, output and round key bits with non-negligible probability for the round function in round j, $0 \leq j \leq r$. We will abbreviate this as L_j. An equation

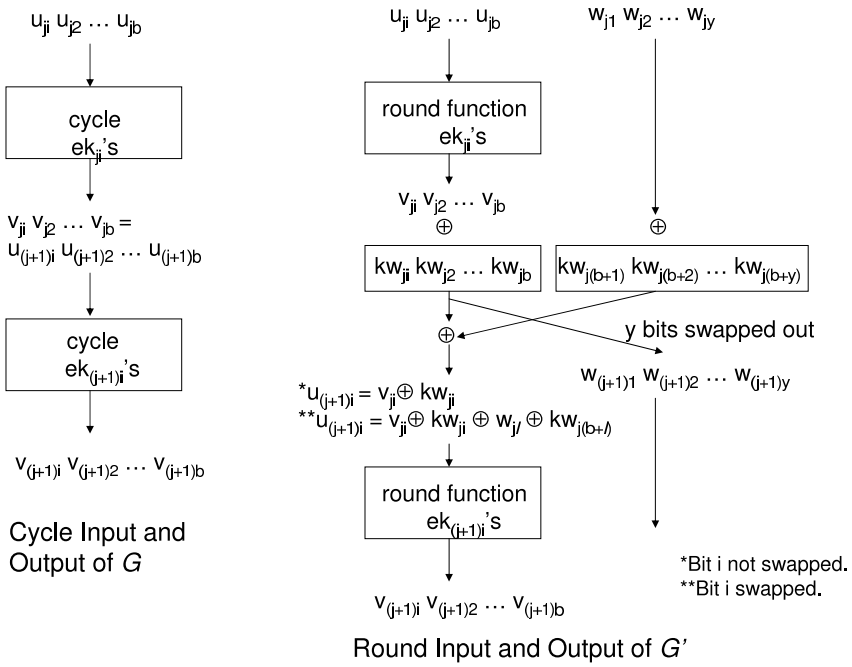

Fig. 2. Linear Relationship Between Round j's Output and Round $(j+1)$'s Input

in L_j holds with probability $\frac{1}{2}+\alpha$ for some non-negligible α such that $0 < \alpha \leq \frac{1}{2}$. For example, if $u_{12} \oplus v_{13} \oplus ek_{15} = 0$ with probability 0.75, this equation will be in L_1. Any equation which reflects a negative relationship, meaning the equation holds with probability $\frac{1}{2}-\alpha$, is rewritten as an equation holding with probability $\frac{1}{2}+\alpha$.

- Without loss of generality, the equations in L_j are in reduced form; for example, $u_{j2} \oplus u_{j2} \oplus u_{j2} = 1$ will be reduced to $u_{j2} = 1$.
- Internal variables will refer to the set of u_{ji} for $1 \leq j \leq r$ and v_{ji} for $0 \leq j \leq r-1$, with $1 \leq i \leq b$. *i.e.*, any variable corresponding to an input bit for rounds 1 to r or to an output bit of rounds 0 (initial whitening step) to $r-1$.

A linear relationship across consecutive rounds is obtained by combining the linear equations for each of the rounds, with v_{ji} becoming $u_{(j+1)i}$. A linear relationship exists that involves only plaintext, ciphertext and expanded-key bits if the intermediate round inputs and outputs (the internal variables) cancel when combining the per round equations, leaving equation(s) involving only u_{0i}'s, v_{ri}'s and expanded-key bits. For example, if in G with two cycles: $u_{11} \oplus v_{12} = ek_{11}$ and $u_{22} \oplus v_{26} = ek_{23}$. Then, since $v_{12} = u_{22}$, $u_{11} \oplus ek_{11} \oplus v_{26} = ek_{23}$.

We now consider how the steps between the rounds in G' impact the linear relationships across the rounds.

- Let Y denote the rightmost y bits of the data block for a $(b+y)$-bit data block.
- Let Γ' refer to the set of the equations used in a linear attack on r rounds of G' formed from combining the L_j's for the individual rounds along with the end of round whitening and swap steps.
- Let Γ refer to a set of linear equations for G formed from equations in Γ'.

– Let kw_{ji} denote the i^{th} key bit used for the whitening step added in round j when constructing G', $1 \leq i \leq b+y$ and $1 \leq j \leq r$. $kw_{ji} = 0$ for $1 \leq i \leq b$ if the cycle of G includes end of cycle whitening and $kw_{0i} = 0$ for $1 \leq i \leq b$ if G contains initial whitening because G' does not add whitening to the b bits when it is already present.
– Let w_{jl} denote the l^{th} bit of the Y portion of the data, for $1 \leq l \leq y$ and $2 \leq j \leq r$. $w_{jl} = v_{(j-1)h} \oplus kw_{(j-1)h}$ where $1 \leq h \leq b$ and h is the bit position swapped with bit position l in the previous swap. When $j = 1$, $w_{1l} = w_{0l} \oplus kw_{0(b+l)}$, the initial input bit XORed with the initial whitening applied.

With the addition of the whitening and swap steps, the input to the round function is now defined as:

– $u_{(j+1)i} = v_{ji} \oplus kw_{ji}$ when v_{ji} is not involved in the swap step.
– $u_{(j+1)i} = v_{ji} \oplus kw_{ji} \oplus w_{jl} \oplus kw_{j(b+l)}$ when v_{ji} is involved in the swap step. When $j \geq 2$, this can be written as $u_{(j+1)i} = v_{ji} \oplus kw_{ji} \oplus v_{(j-1)h} \oplus kw_{(j-1)h} \oplus kw_{j(b+l)}$.

Notice that the steps between applications of the round function in G' maintain a linear relationship between the output of one round and the input of the next round.

If the key schedule of G' produces whitening bits which are created independently of the key bits used within the round function (to the extent that the key bits are pseudo-random), and of the round function's input and output, these whitening bits will cancel with any v_{ji}, u_{j+1} and/or ek_{ji} with probability $\frac{1}{2} + e$ for negligible e (*i.e.,* there is no discernable relationship between these whitening bits and any of the plaintext, ciphertext and expanded-key bits used internal to the round function by definition of the key schedule). Thus, the kw_{ji}'s added when forming G' will not increase the probability of a linear relationship between plaintext bits, ciphertext bits and expanded-key bits used in the round function. If a key schedule is used for G' that does not guarantee independence amongst the kw_{ji}'s and that results in cancellation among some kw_{ji}'s, this is merely cancelling variables that are not present in the linear equations for the round function and thus will not simplify the equations or increase the probability that an equation holds across r applications of the round function.

Now we assume a set of equations, Γ', exist for G' that contains no internal variables and show how to convert them to a set of equations for G. Given the sets, L_j's, of linear equations for the round function in G', these same sets of equations hold for G because the elastic version does not alter the cycle of G. These equations are combined across cycles as was done for the rounds of G', except to form the input to one cycle from the output of the previous cycle, the impact of the swap step and any whitening added when forming G' is removed as follows:

– Set kw_{ji} to 0 for $0 \leq j \leq r$ and $1 \leq i \leq b$ so these whitening bits are omitted from the resulting equations. This removes any initial and end of round whitening that was added to the leftmost b bits when forming G'. Recall that if G had initial and end of cycle whitening, it was treated as part of the round function of G and additional whitening on the leftmost b bits in each round was not added when forming G' (*i.e.* kw_{ji} was already 0 in the equations for G' for $0 \leq j \leq r$ and $1 \leq i \leq b$).
– Set $kw_{0(b+l)} = 0$ and $kw_{1(b+l)} = 0$ for $1 \leq l \leq y$. This sets the rightmost y bits of the initial whitening and of the end of round whitening in the first round to 0. By using plaintexts that have the rightmost y bits set to 0, this results in the rightmost y bits in the first round having no impact on the equations.

– Set $kw_{j(b+l)}$ to $v_{(j-1)h}$ for $2 \leq j \leq r-1$ and $1 \leq l \leq y$, where h is the index in the leftmost b bits corresponding to the bit position swapped with the l^{th} bit of the rightmost y bits. This removes the impact of the swap steps by having the rightmost y bits of whitening in each round cancel with the y bits omitted from each round. These settings are needed only on rounds 2 through $r-1$. The output of the r^{th} round function is the ciphertext so the swap step is not applicable after the r^{th} round. Per the previous item, the rightmost y bits in the first round can be set to have no impact on the equations. Each such setting can add an internal variable, $v_{(j-1)h}$, which now equals u_{jh}, to the equations.

These settings result in each input bit to the $(j+1)^{st}$ round function being of the form $u_{(j+1)i} = v_{ji}$ and the impact of any added end of round whitening and the swap step being removed. The equations will combine to form a set of equations, Γ from the equations in Γ' with any kw_{ji}'s which appear in Γ' removed and with at most $(r-2)y$ internal variables added to the equations. Before explaining how these variables can be accommodated, we first state a few additional notes on the resulting equations. The equations in Γ may contain up to y extra plaintext bits and up to y extra ciphertext bits beyond the b-bit block size of G since G' processes $b+y$ bit blocks. The attacker can set these extraneous y plaintext bits to any value (the whitening bits were set in the conversion based on these plaintext bits being set to 0) and the extra y ciphertext bits are identical to y of the bits output from the next to last round function. For any equation $Eq' \in \Gamma'$ that holds with probability $\frac{1}{2}+\alpha$, the corresponding equation, $Eq \in \Gamma$, formed by removing the $kw'_{ji}s$ from Eq' will also hold with probability $\frac{1}{2}+\alpha$. Furthermore, only variables representing whitening bits not present in G are deleted when converting Γ' to Γ and no equations are added or removed. An equation will not disappear when removing kw_{ji} variables because that would imply the equation did not involve plaintext and/or ciphertext bits.

We now address the presence of the internal variables in Γ. Since it was assumed Γ' consists entirely of equations involving only plaintext, ciphertext and expanded-key bits, the removal of the swap step can introduce up to y internal variables, $(v_{ji's})$, per round (cycle) into the equations. The removal of the swap step impacts $r-2$ rounds (cycles), resulting in a maximum of $(r-2)y$ internal variables in the equations in Γ. If equations in Γ' corresponding to some $y > 0$ are converted directly into equations for the original cipher ($y = 0$), this results in at most $2^{(r-2)y}$ possible values to try for the internal variables. However, it is possible to make the number of such values to test linear in y instead of exponential in y. Instead of converting the attack on G' directly to an attack on G, repeatedly decrease y one bit at a time (decrease the block size of G') converting the attack on G' with a $b+n$ bit block size to an attack on G' with a $b+n-1$ bit block size, for $n = y, y-1, ...1$. When Γ' is converted into a set of equations for the cipher corresponding to a $b+y-1$ blocksize, there are at most $r-2$ internal values, one for each of rounds 2 to $r-1$, and therefore at most 2^{r-2} possible combinations of values for the internal values. Let Γ'_{b+y-1} denote this set of equations. Using (plaintext,ciphertext) pairs with a $b+y-1$ bit block size, solve the equations, setting the $r-2$ internal variables in the equations to the specific values that result in a solution consistent with the (plaintext, ciphertext) pairs. In the worst case, all possible combinations of values for the interal variables must be tested in the equations, resulting in at most $2^{(r-2)}$ combinations to test. Then repeat the process, decreasing the block size one bit at a time. In each iteration, there are at most $r-2$ internal variables whose values need to be determined.

More formally, given G' with a block size of $b + y$ bits, where $0 \leq y \leq b$ and the set of linear equations Γ' used to attack r consecutive rounds of G':

- Let G'_{b+n} refer to an elastic version of G with a $(b + n)$- bit block size, where $0 \leq n \leq y$.
- Let Γ^*_{b+n} refer to the set of linear equations for r consecutive rounds of G'_{b+n} with at most $r - 2$ internal variables present in the equations.
- Let Γ'_{b+n} refer Γ^*_{b+n} with the values of the internal variables determined. This is a set of linear equations involving only plaintext, ciphertext and expanded key bits for r rounds of G'_{b+n}.
- Let A_{b+n} refer to the attack on G'_{b+n} using Γ'_{b+n}.

Convert the attack on G' to an attack on G as follows:

$n = y$
$\Gamma'_{b+n} = \Gamma'$
while $(n > 0)$ {
    convert Γ'_{b+n} to Γ^*_{b+n-1}
    Using (plaintext,ciphertext) pairs for G'_{b+n-1}, solve for any
    internal variables in Γ^*_{b+n-1} to obtain Γ'_{b+n-1}.
    $n \leftarrow n - 1$
}

The set of equations, Γ, used to attack G will be Γ'_b. This results in at most $\sum_1^y 2^{(r-2)} = y2^{(r-2)}$ possible combinations of the internal variables to try as opposed to $\leq 2^{(r-2)y}$ combinations. Since r is constant (and small in practice) and y is bounded by b, which is constant, the amount of work in converting the attack on G' to an attack on G is polynomial in the time to attack G', specifically, the work is bounded by a constant times the time to attack G'. For example, in AES with a 128-bit key, $b = 128$ and $r = 10$, thus $y \leq 128$ and $y(2^{(r-2)}) \leq 128 * 256 = 32768$. The amount of memory required is linear in the amount of memory required to attack G'. In the worst case, a separate amount of memory is required when forming each Γ'_{b+n}. Thus, a linear attack on a r-round version of G' that requires less than 2^b (plaintext, ciphertext) pairs implies a linear attack exists on G.

Theorem 1 can be applied to algebraic equations in general. An algebraic attack on a block cipher G is defined in the same manner as the linear attack with the modification that the equations can involve any algebraic operations, not just XORs.

Lemma 1. *Given a block cipher G with a block size of b bits and r cycle, and its elastic version G' with a block size of $b + y$ bits for $0 \leq y \leq b$, if G' is subject to an algebraic attack on r rounds then either G is subject to an algebraic attack or the resources exist to perform an exhaustive search on G over all plaintexts.*

Proof. The proof follows directly from the proof to Theorem 1 by removing the qualification in Theorem 1's proof that the equations in the L_j sets are linear. Now Γ' and Γ contain algebraic equations instead of only linear equations. Γ is formed from Γ' exactly as before (the conversion adds only XORs of variables to the equations). Therefore, if an algebraic attack exists on r rounds of G' then an attack exists on G.

4 Differential Cryptanalysis

4.1 Overview

We consider how the conversion of a block cipher to its elastic form impacts differential cryptanalysis. We define a general method for bounding the probability a differential characteristic occurs in the elastic version of a cipher when given the bound for a single round of the original cipher. We have illustrated the method on elastic versions of AES and MISTY1 in [2]. We use the symbol Δ to refer to the XOR of two bit strings. The sequence of Δ inputs and outputs of the rounds of a block cipher is a differential characteristic. Specifically, let $(P1, C1)$ and $(P2, C2)$ be two (plaintext, ciphertext) pairs for a block cipher with r rounds. $\Delta P = P1 \oplus P2$ and $\Delta C = C1 \oplus C2$. Let λ_{ij} refer to the delta input to round j and let λ_{oj} refer to the delta output of round j. $\lambda_{i1} = \Delta P$. $\lambda_{or} = \Delta C$. Let pr_j be the probability λ_{oj} occurs given λ_{ij}. Let $\Omega = (\lambda_{i1}, \lambda_{o1}, \lambda_{i2}, \lambda_{o2} ... \lambda_{ir}, \lambda_{or})$. The probability Ω ocurrs is $\prod_{j=1}^{j=r} pr_j$. If the block size is b bits, it is sufficient to show that no differential characteristic occurs with probability $\leq 2^{-b}$ in order to prove a cipher is immune to differential cryptanalysis (because this implies $\geq 2^b$ (plaintext, ciphertext pairs) are required for the attack).

The variable block size and the swap step in elastic block ciphers significantly increase the number of cases to explore when determining the probability of a differential characteristic compared to that of the fixed-length version of a block cipher. This is the reason why we had to find a new approach to modelling the differentials instead of using an existing approach, such as the differential trails approach used on AES [5]. Furthermore, the structure of elastic block ciphers allows analysis performed on the fixed-length version to be partially reused when evaluating the elastic version.

The method we use to bound the probabilities of differential characteristics for an elastic block cipher involves defining states representing which bytes in the differential input to a round have a non-zero delta and tracking what sequences of states the cipher can potentially pass through over a number of rounds. Using this method and differential bounds for the round function of the original cipher, we can derive an upper bound on differential characteristics for the elastic version of a cipher. We exclude the initial and final key-dependent mixing steps from our analysis in order to focus on the core structure and these permutations will only reduce the probability of any specific differential characteristic occurring.

4.2 General Observation

The first observation we make regarding differential cryptanalysis of elastic block ciphers is that, unlike linear cryptanalysis where the equations for the elastic version, G', of a block cipher can be converted directly into equations for the original cipher G, a differential characteristic for G' cannot be converted directly into a differential characteristic for G except for one special case.

We use the following notation when describing a differential characteristic of an elastic block cipher.

- ΔY_i is the XOR of two y-bit segments for round i.
- ΔBin_i is the XOR of two b-bit segments input to the round function in round i.

– $\Delta Bout_i$ is the XOR of two b-bit segments output from the round function in round i.
– A b-bit value formed from the XOR of a b-bit value and a y-bit value, where $y \leq b$, refers to the b-bit result when the y bits are XORed with a subset of y bits of the b bits and the remaining $b - y$ bits are unchanged.
– Forming ΔY_{i+1} from $\Delta Bout_i$ refers to setting ΔY_i to the y bits from $\Delta Bout_i$ that are in the bit positions involved in the swap step after round i.
– ΔY, ΔBin and $\Delta Bout$ without a subscript of i refers to a specific delta independently of the round.

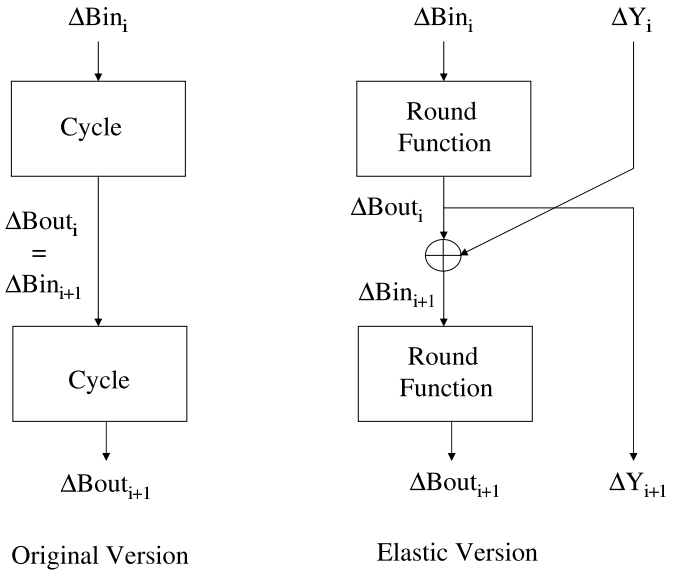

Fig. 3. Differential in Original and Elastic Versions of a Cipher

In the elastic version of a cipher, ΔBin_{i+1} is determined by $\Delta Bout_i$ and ΔY_i. If $\Delta Y_i \neq 0$ then $\Delta Bin_{i+1} \neq \Delta Bout_i$; whereas, $\Delta Bin_{i+1} = \Delta Bout_i$ in the original block cipher. This is shown in Figure 3. Therefore, a sequence of deltas ocurring across multiple rounds in the elastic version will not hold across the original version unless $\Delta Y_i = 0$ for r sequential rounds.

Now we consider the special case where r consecutive ΔY_i's are 0.

Lemma 2. *If a differential characteristic occurs in the elastic version, G', of a block cipher that contains r consecutive rounds with $\Delta Y_i = 0$ and this characteristic can be used to attack G', then it can be used to attack G.*

Proof. Let Ω' be the characteristic corresponding to the ΔBin_i values and $\Delta Bout_i$ values for the r consecutive rounds each with $\Delta Y_i = 0$. Ω' is also a characteristic for the r rounds of G. Ω' must hold with probability $> 2^{-b-y}$ to be used in an attack on G'. If Ω' holds with probability $2^{-\alpha} > 2^{-b}$, then it can be used to attack G directly, provided the probability is large enough that it is computationally feasible to encrypt $O(2^{\alpha})$ plaintexts.

If it holds with probability $2^{-\alpha}$ such that $2^{-b} > 2^{-\alpha} > 2^{-b-y}$, it can be used to attack G as follows: Using an r round version of G' and (plaintext, ciphertext) pairs consistent with the delta input and delta output of Ω' by setting the leftmost b bits to be consistent with Ω' and the rightmost y bits to have a Δ of 0. Then apply the attack on G' to find the round keys for the r rounds and use these as the keys for the r cycles of G.

However, if this later case where $2^{-b} > 2^{-\alpha} > 2^{-b-y}$ is computationally feasible, it implies it is computationally feasible to encrypt 2^b plaintexts with G. Thus G is insecure because given a ciphertext, C, an attacker can ask for all 2^b plaintexts be encrypted with the same key (which is unknown) used to generate C and see which plaintext produces C. As an estimate of the probability of r consecutive rounds having $\Delta Y = 0$, consider what happens if the y bits left out of each round in G' take on any of the possible 2^y values with equal probability. Then, ignoring the differential for the b-bit portions of each round's input and output, a case where $\Delta Y_i = 0$ for r consecutive rounds may be found for small values of y and r. If each ΔY_i occurs with probability 2^{-y}, then the probability that $\Delta Y_i = 0$ in r consecutive rounds is 2^{-yr}. For example, in MISTY1, $r = 4$ (MISTY1 contains four cycles and a cycle is used as the round function in the elastic version). When $y = 1$, the probability of r consecutive ΔY's being zero is $\frac{1}{16}$.

4.3 State Transition Method

We now consider how to evaluate any elastic block cipher's immunity or susceptibility to differential cryptanalysis by using the bound from a single cycle of the fixed-length version of the cipher.

Theorem 2. *The differential probabilities from the cycle of a fixed-length block cipher G can be used to bound the probability that a differential characteristic occurs in its elastic version G'.*

The general method we use is the tracking of states through the rounds of an elastic block cipher. We devise a method for categorizing the impact of the swapping of bits between rounds on the differentials entering a round. We combine the impact of the swap step with the upper bound on the probability a differential characteristic occurs in a single application of the round function (from available analysis on G) to determine an upper bound the probability of a differential characteristic across multiple rounds in G'. By obtaining a bound, x, on the probability across n rounds in G', the probability across r' rounds can be bounded by $x^{\lfloor \frac{r'}{n} \rfloor}$.

In the case where the round function of G is a cycle, such as in AES, we view the $(b+y)$-bit data block entering a round of G' as a b-bit segment and a y-bit segment. Three main states are defined:

($\Delta Bin = 0$ and $\Delta Y \neq 0$), ($\Delta Bin \neq 0$ and $\Delta Y = 0$), ($\Delta Bin \neq 0$ and $\Delta Y \neq 0$)
The state in which $\Delta Bin = 0$ and $\Delta Yin = 0$ is not of interest because, given a non-zero delta input to the cipher, a delta of zero across all $b + y$ bits cannot occur. Within a main state, the number of bytes for which the delta is non-zero are counted. For example, if the input to the third round has a ΔBin that is 1 in the 2^{nd} and 18^{th} bit positions and is zero in all other bits, then there are two bytes with non-zero deltas in ΔBin. Tracking of states between rounds involves determining what $\Delta Bin||\Delta Y$ can result for the $(i+1)^{st}$ round

based on the delta in the i^{th} round. For example, if $\Delta Bin = 0$ and $\Delta Y \neq 0$ in the input to round i, then $\Delta Bin \neq 0$ and $\Delta Y = 0$ in round $i + 1$. This is because the delta output of the i^{th} round function will be zero, then the non-zero ΔY will be swapped into the b-bit portion input to the $(i + 1)^{st}$ round and a delta of zero will be swapped out to form the ΔY for the $(i + 1)^{st}$ round.

When the original cipher is a Feistel network (or is a Feistel network with additional steps as in the case of MISTY1), the ΔBin portion is viewed as a left half (ΔLin) and right half (ΔRin). The main states are the seven combinations of $\Delta L, \Delta R$ and ΔY being $= 0$ or $\neq 0$ with at least one being $\neq 0$.

Using the states, an upper bound (which is not necessarily a tight upper bound) can be determined for the probability of a differential characteristic for r' rounds of G'. The probability of a differential characteristic occurring for a single application of the round function of G and the possible ΔB or $\Delta L||\Delta R$ values entering the round function in each round are used to bound the probability for a round of G'. The possible ΔB or $\Delta L||\Delta R$ and ΔY values in a round determine the possible input states to the next round of G'.

4.4 Examples

We applied the state transition method to the elastic versions of AES and MISTY1 described in [3]. The process and results are described in [2]. We briefly state the results of the work here. Elastic AES is an example in which the input to each round is viewed in the form of $\Delta Bin||\Delta Y$. AES is a 128-bit block cipher with 10 rounds. The number of rounds, r', in the elastic version is $10 + \lceil \frac{10y}{128} \rceil$. Elastic MISTY1 is an example in which the input to each round is viewed in the form of $\Delta L||\Delta R||\Delta Y$. MISTY1 is a 64-bit block cipher involving four cycles of a Feistel network. $r' = 4 + \lceil \frac{4y}{64} \rceil$ in the elastic version of MISTY1.

We analyzed the elastic versions without the initial and final key dependent permutations to simplify the model since these permutations will only decrease the probability that a specific differential characteristic occurs. Our analysis is independent of the key schedule.[4] The swap step is performed by selecting y consecutive bits from the round function's output to XOR and swap with the y bits left out of the round function. In the implementation of elastic AES, the starting position of the y bits selected rotates to the right one byte each round. In elastic MISTY1, the starting position alternates between the left and right halves of the b bit segment in addition to rotating to the right within the half block each round.

When analyzing the state transitions for both elastic AES and elastic MISTY1, we are concerned with how many byte positions have non-zero deltas. Therefore, we only need to consider each block size where Y contains an integer number of bytes. The case for $y = 8x$ where x is an integer such that $1 \leq x \leq \frac{b}{8}$ covers the cases of y such that $8(x - 1) < y \leq 8x$. For example, the lower bound on a differential characteristic occurring for the case of $y = 8$ is also the lower bound for values of y in the range of 1 to 7 because this range of y influences exactly one byte in b-bit portion during each of the swap steps.

In order to analyze the state transitions in elastic AES, we created a program that tracks how many bytes contain a non-zero differential characteristic in each round and

[4] In the constructions from [3], the stream cipher RC4 was used for the key schedule.

determines the possible next states. The number of bytes with a non-zero delta in the b-bit portion in a single round bounds the probability that a differential characteristic holds through that round. A lower bound on the differential probability for a single round of AES is $\leq 2^{-exp}$ where $exp = 6 * |\Delta Bin|$. The multiplication by 6 is due to the fact that the probability a specific difference in two one-byte inputs to AES's S-Box produces a specific difference in the two outputs of the S-Box is 2^{-6} or 2^{-7}, depending on the exact byte values ([5] pages 205-206). For block sizes of 17, 18 ... to 32 bytes, the model was run through three rounds for all possible input states. A lose lower bound for all r' rounds was then calculated by viewing the r' rounds as 3 round segments plus 0 to 2 additional rounds, depending on the exact value of r'. Sequences producing a three round bound which did not exclude the possibility of a differential attack were traced through subsequent rounds, with the number of rounds depending on the exact size of y and the probability produced after each round. The results from our analysis show that the probability of a differential characteristic occurring is $\leq 2^{-128-y}$. Therefore, a differential attack is impossible.

Our analysis of elastic AES is general in terms of block size but only considers a single method for selecting the bits to swap (described previously) after each round as opposed to all possible ways of selecting y bits from 128 bits. In [4] it was proven that an elastic version of a cipher is immune to any practical key-recovery attack if the original cipher is immune to the attack regardless of the specific bit positions chosen for the swap steps. Differential cryptanalysis is covered by this result. The state transition method can be applied to any choice of bits to swap, but it is computationally infeasible to include in one model all $2^{y(r'-1)}$ possible ways of selecting the bits to swap in the first $r'-1$ rounds (recall that the swap step adds no value after the last round and thus can be omitted from round r').

MISTY1 uses two functions, referred to as $F0$ and FL, as building blocks along with a Feistel network. $F0$ is the round function in the Feistel network. In each cycle of the Feistel network, FL is applied to one half of the data and FL^{-1} is applied to the other half. An upper bound of 2^{-56} on the probability a differential characteristic occurs was derived for 4 cycles of the 64-bit version [8] by using a bound of 2^{-14} per cycle due entirely to the bound from the $F0$ function. Using a manual analysis of state transitions and only the bound for the $F0$ function, we derive an upper bound on the elastic version of MISTY1 of $2^{-14(r'-1)}$, where r' is the number of rounds (cycles of MISTY1) in the elastic version. This bound is not tight and does not by itself eliminate the possibility of a differential attack (either in MISTY1 or the elastic version). However, the state transition analysis does reduce the number of state sequences that need to be investigated to tighten the bound over r' rounds. The bound of $2^{-14(r'-1)}$ also allows the potential contribution needed from the initial and final key-dependent mixing steps in preventing differential attacks to be determined.

5 Conclusions

We showed how to convert a linear, or more generally any algebraic, attack on an elastic block cipher into such an attack on the fixed-length version of the block cipher to prove that if the fixed-length version is immune to such an attack then so is the elastic version. This was accomplished by proving that any set of linear or algebraic equations used in an

attack on the elastic version can be converted in polynomial time and memory into equations for the fixed-length version. We also devised a method for bounding the probability of a differential characteristic on the elastic version of a block cipher using the differential bounds for the cycle of the fixed-length version of the cipher. When performing differential cryptanalysis on an elastic block cipher, the differential bound for the round function is the bound from the cycle of the original version of the cipher. The swapping of bits between rounds in the elastic version impacts the sequence of differentials entering the series of rounds by altering the output of the i^{th} application of the round function before it is input to the $(i+1)^{st}$ application of the round function. The bound for the round function and the impact of the swap step can be combined to bound the probability a differential characteristic occurs in the elastic version of a block cipher. This is accomplished by defining states representing whether or not there is a non-zero differential in the b-bit portion and/or y-bit portion of the round's input, then determining what states may potentially occur as input to each round. The possible state sequences in the elastic version of the cipher are combined with the probabilities a differential characteristic occurs in one cycle of the original cipher to bound the probability of a differential characteristic across all rounds of the elastic version of the cipher.

Acknowledgments

This work was partially supported by NSF Grants ITR CNS-04-26623 and CPA CCF-05-41093. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF or the U.S Government.

References

1. E. Biham and A. Shamir, *Differential Cryptanalysis of the Data Encryption Standard*, Springer-Verlag, New York, 1993.
2. D. Cook, *Elastic Block Ciphers*, Ph.D. Thesis, Columbia University, 2006.
3. D. Cook, M. Yung and A. Keromytis, Elastic Block Ciphers: The Basic Design, *Proceedings of ASIACCS*, ACM, pages 350-355, 2007.
4. D. Cook, M. Yung, and A. Keromytis, The Security of Elastic Block Ciphers Against Key-Recovery Attacks. *Proceedings of ISC*, LNCS 4779, pages 89-103, 2007.
5. J. Daemen and V. Rijmen. The Design of Rijndael: AES the Advanced Encryption Standard. Springer-Verlag, Berlin, 2002.
6. M. Matsui, Linear Cryptanalysis Method for DES Cipher, *Proceedings of Advances in Cryptology - Eurocrypt 1993*, LNCS 0765, Springer-Verlag, pages 386-397, 1994.
7. M. Matsui, New Block Encryption Algorithm MISTY, *Proceedings of Fast Software Encryption 1997*, LNCS 1267, Springer-Verlag, pages 54-68, 1997.
8. M. Matsui, New Structure of Block Ciphers with Provable Security Against Differential and Linear Cryptanalysis, *Proceedings of Fast Software Encryption 1996*, LNCS 1039, Springer-Verlag, pages 205-218, 1996.
9. NIST, FIPS 197 Advanced Encryption Standard (AES), 2001.