The Security of the RC6™ Block Cipher

Scott Contini
RSA Laboratories
scontini@rsa.com

Ronald L. Rivest
M.I.T. Laboratory for Computer Science
rivest@theory.lcs.mit.edu

M.J.B. Robshaw
RSA Laboratories
matt@rsa.com

Yiqun Lisa Yin
RSA Laboratories
yiqun@rsa.com

Version 1.0 - August 20, 1998

Abstract

This report presents a preliminary analysis of the security offered by the RC6™ block cipher. RC6 is an evolutionary improvement of RC5, designed to meet the requirements of the Advanced Encryption Standard (AES). Our analysis demonstrates that RC6 is highly resistant to differential and linear cryptanalytic attack, which are currently the two most effective analytical attacks on block ciphers. The data requirements to mount an attack using either of these two attacks exceeds the amount of available data with considerable margins for safety.

Contents

List of Tables

1 Introduction

RC6™ is a new block cipher submitted to NIST for consideration as the new Advanced Encryption Standard (AES). The design of RC6 began with a consideration of RC5 [28] as a potential candidate for an AES submission. Modifications were then made to meet the AES requirements, to increase security, and to improve performance. The inner loop, however, is based around the same "half-round" found in RC5.

RC5 was intentionally designed to be extremely simple, to invite analysis shedding light on the security provided by extensive use of data-dependent rotations. Since RC5 was proposed in 1994, various studies [2, 4, 8, 9, 14, 31] have provided a greater understanding of how RC5's structure and operations contribute to its security. While no practical attack on RC5 has been found, the studies provide some interesting theoretical attacks, generally based on the fact that the "rotation amounts" in RC5 do not depend on all of the bits in a register. RC6 was designed to thwart such attacks, and indeed to thwart all known attacks, providing a cipher that can offer the security required for the lifespan of the AES.

Since RC6 is closely built on RC5, many of the same security considerations come into play during analysis. We will concentrate most on the use of differential and linear cryptanalysis in attacking RC6, but we will also consider other styles of attack. First we review the design of RC6. A more complete description of the cipher design, the motivation and issues such as performance can be found in the companion paper *The RC6 Block Cipher* [29].

1.1 Description of RC6

RC6 is one of a fully parameterized family of encryption algorithms. A version of RC6 is more accurately specified as RC6-$w/r/b$ where the word size is w bits, encryption consists of a nonnegative number of rounds r, and b denotes the length of the encryption key in bytes. Note that in the description of RC6 the term "round" is somewhat analogous to the usual DES-like idea of a round: half of the data is updated by the other half; and the two are then swapped. Since the AES submission is targeted at $w = 32$ and $r = 20$ we shall use RC6 as shorthand to refer to such versions. When any other value of w or r is intended in the text, the parameter values will be specified as RC6-w/r. Of particular relevance to the AES effort will be the versions of RC6 with 16-, 24-, and 32-byte keys. For all variants, RC6-$w/r/b$ operates on units of four w-bit words using the following six basic operations. The base-two logarithm of w will be denoted by $\lg w$.

$a + b$	integer addition modulo 2^w
$a - b$	integer subtraction modulo 2^w
$a \oplus b$	bitwise exclusive-or of w-bit words
$a \times b$	integer multiplication modulo 2^w
$a \lll b$	rotate the w-bit word a to the left by the amount given by the least significant $\lg w$ bits of b
$a \ggg b$	rotate the w-bit word a to the right by the amount given by the least significant $\lg w$ bits of b

The user supplies a key of length k bytes and the 128-bit plaintext block is loaded into words A, B, C, and D starting with the low-order byte of A. These four w-bit words contain the output ciphertext at the end. The key schedule of RC6 is described in the document *The RC6 Block Cipher* [29] and in Part III of this report where it comes under closer examination. Here we describe encryption and decryption. RC6 works with four w-bit registers A, B, C, D which contain the initial input plaintext as well as the output ciphertext at the end of encryption. The first byte of plaintext or ciphertext is placed in the least-significant byte of A; the last byte of plaintext or ciphertext is placed into the most-significant byte of D. We use $(A, B, C, D) = (B, C, D, A)$ to mean the parallel assignment of values on the right to registers on the left.

Encryption with RC6-$w/r/b$

Input: Plaintext stored in four w-bit input registers A, B, C, D
Number r of rounds
w-bit round keys $S[0, \ldots, 2r + 3]$

Output: Ciphertext stored in A, B, C, D

Procedure:

$B = B + S[0]$
$D = D + S[1]$
for $i = 1$ **to** r **do**
{
$t = (B \times (2B + 1)) \lll \lg w$
$u = (D \times (2D + 1)) \lll \lg w$
$A = ((A \oplus t) \lll u) + S[2i]$
$C = ((C \oplus u) \lll t) + S[2i + 1]$
$(A, B, C, D) = (B, C, D, A)$
}
$A = A + S[2r + 2]$
$C = C + S[2r + 3]$

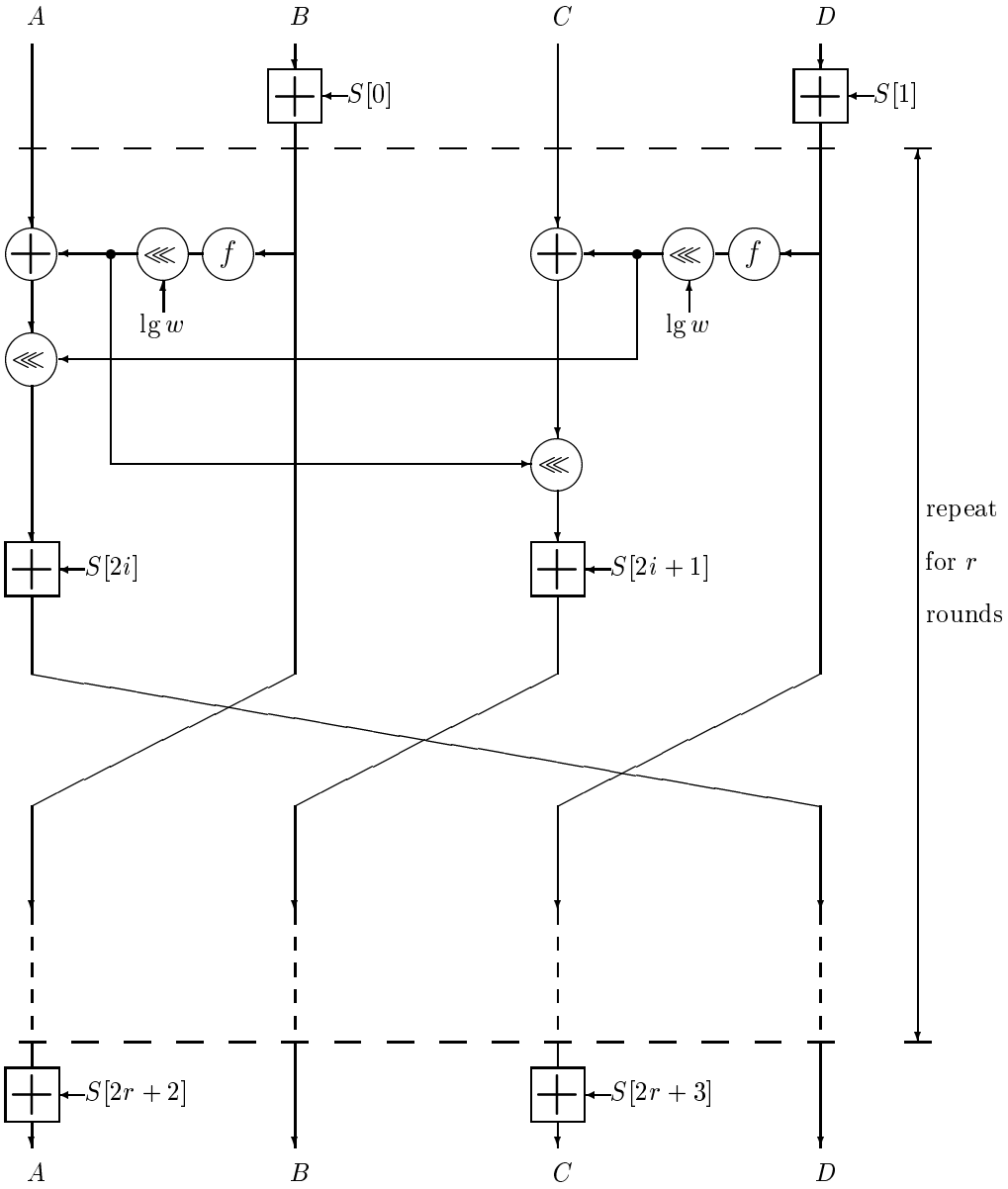


Figure 1: Encryption with RC6-$w/r/b$. Here $f(x) = x(2x + 1)$.

Decryption with RC6-$w/r/b$

Input: Ciphertext stored in four w-bit input registers A, B, C, D
Number r of rounds
w-bit round keys $S[0, \ldots, 2r+3]$

Output: Plaintext stored in A, B, C, D

Procedure:
$C = C - S[2r+3]$
$A = A - S[2r+2]$
for $i = r$ **downto** 1 **do**
{
$(A, B, C, D) = (D, A, B, C)$
$u = (D \times (2D+1)) \lll \lg w$
$t = (B \times (2B+1)) \lll \lg w$
$C = ((C - S[2i+1]) \ggg t) \oplus u$
$A = ((A - S[2i]) \ggg u) \oplus t$
}
$D = D - S[1]$
$B = B - S[0]$

1.2 Overview of this report

The remainder of this report is split into four parts. In the first part we consider the security of RC6 with regards to differential cryptanalysis [1]. Then we consider linear cryptanalysis [22] and its application to RC6. Finally in the last two parts we will briefly address issues related to the key schedule and the attack of differential-linear cryptanalysis [20].

To facilitate our analysis we make the simple observation that if we were to drop the fixed rotation by $\lg w$ bits (FR) from RC6 along with the quadratic function $f(x) = x(2x+1)$ (i.e. replacing it with the identity function $f(x) = x$), then the resulting cipher would be very similar to how we might imagine a four-strand version of RC5 would look. This will be the starting point for our analysis and we will gradually introduce additional complexity as we analyze increasingly close approximations to RC6. We introduce some of these simpler variants to RC6 here.

RC6-I-NFR A version of RC6 in which $f(x) = x(2x+1)$ is replaced by the identity function $f(x) = x$ and there are no fixed rotations.

Encryption with RC6-I-NFR

Procedure:
$B = B + S[0]$
$D = D + S[1]$
for $i = 1$ **to** r **do**
{
$t = B$
$u = D$
$A = ((A \oplus t) \lll u) + S[2i]$
$C = ((C \oplus u) \lll t) + S[2i+1]$
$(A, B, C, D) = (B, C, D, A)$
}
$A = A + S[2r+2]$
$C = C + S[2r+3]$

RC6-NFR A version of RC6 in which there are no fixed rotations.

Encryption with RC6-NFR

Procedure:
$B = B + S[0]$
$D = D + S[1]$
for $i = 1$ **to** r **do**
{
$t = B \times (2B+1)$
$u = D \times (2D+1)$
$A = ((A \oplus t) \lll u) + S[2i]$
$C = ((C \oplus u) \lll t) + S[2i+1]$
$(A, B, C, D) = (B, C, D, A)$
}
$A = A + S[2r+2]$
$C = C + S[2r+3]$

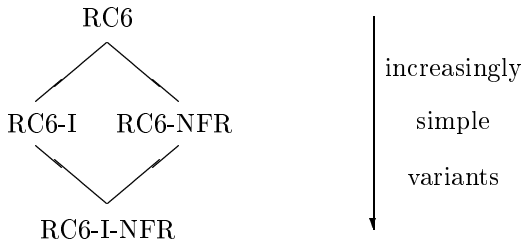
RC6-I A version of RC6 in which $f(x) = x(2x+1)$ is replaced by the identity function $f(x) = x$.

Encryption with RC6-I

Procedure:
$B = B + S[0]$
$D = D + S[1]$
for $i = 1$ **to** r **do**
{
$t = B \lll \lg \mathrm{w}$
$u = D \lll \lg \mathrm{w}$
$A = ((A \oplus t) \lll u) + S[2i]$
$C = ((C \oplus u) \lll t) + S[2i+1]$
$(A, B, C, D) = (B, C, D, A)$
}
$A = A + S[2r+2]$
$C = C + S[2r+3]$

RC6 The full version of RC6 as described in *The RC6 Block Cipher* [29] and Section 1.1 of this report.

It might be illustrative to view these different variants of RC6 as forming a lattice with the most sophisticated variant (that is RC6) uppermost.



1.3 Summary of current attacks against RC6

1.3.1 Brute force attacks

In [29] we conjecture that to attack RC6 the best approach is that of exhaustive search for the b-byte encryption key (or the expanded key array $S[0, \ldots, 43]$ when the user-supplied encryption key is particularly long). Such brute force attacks are applicable to any block cipher, and are only countered by using appropriately long keys.

Estimates of the Plaintext Requirements to Attack RC6					
	number of rounds				
attack	8	12	16	20	24
differential cryptanalysis	2^{56}	2^{117}	$\boxed{2^{190}}$	$\boxed{2^{238}}$	$\boxed{2^{299}}$
linear cryptanalysis	2^{47}	2^{83}	2^{119}	$\boxed{2^{155}}$	$\boxed{2^{191}}$

Table 1: A summary of the best differential and linear cryptanalytic attacks against RC6. RC6 with 20 rounds is recommended. A box denotes when the data requirements for a successful attack exceed 2^{128}, the total number of possible plaintexts.

With RC6 the work effort required for an on-line search for a b-byte key or for the expanded key array $S[0,\ldots,43]$ is $\min\{2^{8b}, 2^{1408}\}$ operations. In principal at least, with considerable memory (in excess of 2^{700} bytes!) and around 2^{704} off-line pre-computations, one could mount a meet-in-the-middle attack to recover the expanded key array $S[0,\ldots,43]$. This would require 2^{704} on-line computations so the work effort for exhaustive search might best be estimated by $\min\{2^{8b}, 2^{704}\}$ operations [29]. For the key sizes specified in the AES effort, a brute-force attack appears to be infeasible.

1.3.2 Analytical attacks

In Table 1 we give the data requirements for the basic differential and linear attacks we have considered against RC6. We have typically concerned ourselves in this report with an investigation of the encryption properties of RC6. A brief consideration of the decryption process suggests that the properties and structure of the decryption process are very similar to those of encryption. The attacks and considerations used in assessing the security of encryption appear to apply equally to decryption and we have not found any styles of attack where the security of the two processes would be widely different. The text of this report contains complete descriptions of a variety of attacks and an analysis of simplified variants of RC6. Table 1 provides the most advantageous figures for the cryptanalyst as a result of our analysis. A surrounding box denotes when the data requirements for a successful attack exceed 2^{128}, the total number of possible plaintexts. We conclude that RC6 with 20 rounds is secure against differential and linear cryptanalytic attacks.

Part I
Differential Cryptanalysis

There are a variety of approaches that can be taken in assessing the security of RC6. We attempt to do so by considering the strength of increasingly accurate approximations to RC6, observing how the security might be enhanced at successive steps. In Sections 3 and 4 we consider two variants of RC6 in which the quadratic function is replaced by the identity function and where the fixed rotation by five bits is either present (RC6-I) or not present (RC6-I-NFR). In Section 5 we study the characteristics of the quadratic function. Then in Section 6, we consider a variant of RC6 in which the fixed rotation is omitted (RC6-NFR). Finally, we assess the strength of the complete version of RC6 against differential-style attacks in Section 7.

Even though the report is written so as to guide the reader through an increasingly sophisticated set of differential attacks on increasingly complicated variants of RC6, the eager reader should feel at liberty to move to Section 7 which deals exclusively with the cryptanalysis of RC6. We also feel that Section 5 is rather technical and involved. For those readers interested in an overview of the security of RC6 but less interested in an in-depth analysis, Section 5 can be read over quickly without the loss of too much continuity.

2 Overview

In this section we review differential cryptanalysis, some of the advanced techniques that might be useful during our analysis, and an overview of our approach to the differential cryptanalysis of RC6. In particular we also consider existing differential attacks on RC5 and how they might relate to RC6 (Section 2.2.1).

2.1 Differential cryptanalysis

Differential cryptanalysis is a chosen plaintext attack on iterative block ciphers. By choosing the plaintext pairs that are encrypted, the difference between the inputs to the final round of the cipher can be predicted with a certain probability. The particular definition of "difference" depends on the block cipher under attack. The aim is that the predicted difference between the pair entering the final round can be used together with the difference in the ciphertext pair (which is observed and hence known) to deduce information about the subkey used in the final round of the cipher. Full details of this style of attack pioneered by Biham and Shamir can be found in [1].

We note that the probability that a chosen plaintext pair will provide the desired difference at the end of round $(r-1)$ will typically decrease as r increases.

Thus attacks on reduced versions of iterated ciphers can be devastatingly effective whereas the full version of the same ciphers can have sufficiently many rounds to deter a cryptanalyst from embarking on such an attack.

The optimal choice for the difference in the plaintext pairs is calculated by investigating *characteristics*; these specify the expected differences for each round of the cipher. A characteristic has some associated probability which is based on the likelihood that the expected difference in the last round (specified by the characteristic) actually occurs given that the specified difference in the first round is used.

Interesting theoretic work by Lai, Massey and Murphy [19] on the applicability of differential attacks to certain types of iterated ciphers introduced the idea of *differentials* which are a broader version of characteristics; only the input and output differences are specified while the differences at intermediate rounds are not considered.

There are some variants to the basic differential attack. While the use of differentials has become a routine consideration in the analysis of block ciphers, other enhancements are more limited in their scope and applicability. As research into the security of RC6 continues a more complete picture of these more advanced techniques will undoubtedly become clear.

Knudsen has introduced the idea of *truncated differentials* (formally called *partial differentials*) [11]. Here the cryptanalyst attempts to predict the behavior of part of the difference but not the full 128-bit differential. Instead, the successful prediction of part of the difference can sometimes lead to the recovery of key material. Truncated differentials can allow the cryptanalyst additional flexibility in attacking a cipher, though very often it seems that their application can be somewhat limited. We have sometimes used a broad notion of truncated differentials as enhancements to some of our basic attacks. At times we have forgone the opportunity to predict the behavior of certain words of a characteristic or differential in an attempt to improve the probability of a recognizable difference occuring. This has helped provide a reduction in the plaintext requirements, but might best be viewed as an optimization trick rather than as a full attack using truncated differentials. Ciphers can be constructed that are vulnerable to analysis with truncated differentials [12] but it is only very occasionally that this attack, or close variants, is more useful against more serious proposals. Two notable examples would be Safer [21, 13] and Skipjack [26, 17].

As an analogy to differentiation in calculus, *higher-order differential attacks* have been considered by Lai [18]. Knudsen subsequently demonstrated that ciphers could be constructed that were vulnerable to high-order differential attack while being resistant to conventional differential attack [12]. Similarly, *interpolation attacks* [5] have been demonstrated as effective attacks on some constructed ciphers. However, both these attacks seem to be somewhat limited when used on more sophisticated ciphers with a reasonable number of rounds. This is our experience with RC6.

One final variation on the theme of differential cryptanalysis is to consider

the use of so-called *quasi-differentials* [13]. By changing the notion of difference part way through an attack, the cryptanalyst might be able to work around some of the incompatabilities that the designer intended to hinder cryptanalysis. The two most obvious notions of difference for RC6 are exclusive-or and integer subtraction. Much of the work we pursue during our analysis of differential cryptanalysis provides a duality between these two notions. However, we feel that there is little to be gained in switching between these two notions of difference while trying to build up a useful characteristic or differential.

Future work will allow us, and others, to expand on the applicability of these more advanced techniques to the cryptanalysis of RC6. One of the nicer features of RC6 is its close relation to RC5. Since RC5 has been available to the community since December 1994, and in that time there have been no reports of these more sophisticated methods being used to attack RC5, it might be tempting to conclude that applying such techniques to RC6 would be as hard. Of course there is no guarantee that this is indeed the case. However we take some comfort from the fact our early analysis seems to suggest that RC6 is as resistant to these advanced methods of analysis as is RC5.

2.1.1 Notation and basic assumptions

Here we establish some of the notation that we will use during this report as well as some of the basic assumptions that we have made.

Throughout we use e_i to denote the 32-bit word which is zero except for a single one in bit position i. In integer terms we have $e_i = 2^i$. At times we will find the need to denote a "generic" difference. This we represent as Δ. As our analysis of the different variants gets more advanced, we add more detail to define the form of Δ thereby yielding a difference pattern of specific interest in our analysis.

In general, we will be assuming in our attacks that in attacking an r-round cipher, a differential for $(r-2)$ rounds will be required. This is a very typical starting assumption, but it is one that is likely to change as analysis progresses. For some ciphers [15] there is no easy way to use an $(r-2)$-round characteristic during an attack, whereas there are improved attacks on other ciphers [28] that allow the cryptanalyst to use a shorter characteristic or differential. Subsequent research will undoubtedly reveal the most effective way of measuring the true security offered by RC6. But in anticipation of the discovery of some advanced techniques that allow even a few more rounds to be removed from the characteristic used by the attacker, we have built in a large margin of safety.

Very often, we will represent the evolution of characteristics in RC6 in a graphical way. In such circumstances we will use $\downarrow$ to represent the action of one round of encryption.

2.2 Differential cryptanalysis and RC6

Many readers are familiar with the fact that the notion of difference can be adapted depending on the cipher under attack. We believe that the two most natural measures of difference for attacking RC6 are those of *bitwise exclusive-or* and *integer subtraction modulo* 2^{32}. We will be addressing both measures of difference during our analysis of RC6.

Before starting, we stress that the security estimates we will be giving for the simple variants of RC6 are only intended to offer guidance. For example we will only consider the bitwise exclusive-or measure of difference for these variants. It is straightforward (and often trivial) to find more data-efficient attacks for these simple variants since our estimates are often derived by using simple characteristics or differentials in the most basic ways. Our purpose in considering these simple variants at all is to illustrate how the security of RC6 might be built up as different components are added. While we have not attempted to be exhaustive in the cryptanalysis of these simple variants, we believe that the estimates we derive for the full cipher will provide the reader with a reasonably accurate picture of the security provided by RC6.

2.2.1 Differential cryptanalysis of RC5

Since the publication of RC5, there have been several results on the strength of RC5 against differential attacks [2, 8, 14]. Analysis of RC5 [8, 9] has shown that the most advantageous strategy for a cryptanalyst is to use differences that do not affect the rotation amount. In fact, once there is a difference in the rotation amount, a very quick avalanche of change takes place that appears to thwart existing differential attacks. With this in mind, our strategy throughout our attempts to cryptanalyze RC6 will be to use differences that do not provide different rotation amounts for a given pair.

We also mention here that it is in general to the attacker's advantage to keep the Hamming weight of almost all the intermediate differences low so as to keep better control over the evolution of the characteristic or differential. Despite this, more recent work on RC5 has shown that heavier weight characteristics can also be beneficial in attacking a cipher [2]. However, due to the use of the quadratic function in RC6, we feel that this style of attack is unlikely to apply to RC6, even though it might apply to some simple variants of RC6 in which the quadratic function is not present.

3 Differential Cryptanalysis of RC6-I-NFR

In this section:

- *We show that RC6-I-NFR offers reasonable resistance to differential cryptanalysis but as many as 38 rounds might be necessary for the purposes of the AES*
- *We demonstrate that there is a very significant effect when considering differentials instead of characteristics*

RC6-I-NFR is a much simplified version of RC6. In structure it might be compared to two parallel versions of RC5 with the rotation amount for one copy of RC5 being taken from the second copy and *vice versa*. By looking at RC6-I-NFR we get a glimpse of the underlying structure of RC6 and we get a feel for the likely propagation of differences through the cipher. In this way we might identify the cryptographic components (addition, exclusive-or, rotation and multiplication) that provide the biggest contribution to security, either in isolation or in combination.

Following Section 2.1.1 we will use Δ to denote a "generic" difference, where we will only use the exclusive-or difference to analyze this variant. We will use α to denote the probability that a difference Δ remains unchanged across an integer addition unit. We note that α depends primarily on the Hamming weight of Δ and so we will denote this probability α_Δ. We let $\rho\Delta$ denote the probability that the difference Δ remains unchanged by the data-dependent rotation.

Starting with a particular difference $(\Delta\,\Delta\,0\,0)$ in the four input words, say, it is interesting to note the path the difference follows through the cipher. Often a word that has some non-zero difference in it will be used to provide the argument for a rotation amount. We will assume in such situations that any differences in the bits is not in the rotation amount, i.e., the five least significant bits of Δ are zero. All 15 non-zero difference patterns involving the difference Δ can be "factored" into three cycles, and the cycles are shown in Table 2.

It seems that differences following the patterns in cycle (a) or (b) in Table 2 will be most useful in a differential attack, since cycle (c) will give a lower probability when extended to cover the same number of rounds. Because of the symmetry between cycle (a) and cycle (b), we will be mainly using cycle (a) as the basic pattern for characteristics and differentials throughout our analysis of RC6 and its variants. This cycle seems to be the best even when we consider some of the other variants of RC6, but we note in Section 7.4 that there are short iterative differentials for RC6 that are more closely related to the (c) cycle.

Immediately from Table 2, there are numerous characteristics that can be identified. Among them, by setting $\Delta = 2^{31}$ we have $\alpha_\Delta = 1$ and so in cycle

(a)				(b)				(c)			
Δ	Δ	0	0	0	0	Δ	Δ	Δ	Δ	Δ	Δ
		$\downarrow$				$\downarrow$				$\downarrow$	
Δ	0	0	0	0	0	Δ	0	Δ	0	Δ	0
		$\downarrow$				$\downarrow$				$\downarrow$	
0	0	0	Δ	0	Δ	0	0	0	Δ	0	Δ
		$\downarrow$				$\downarrow$				$\downarrow$	
0	Δ	Δ	0	Δ	0	0	Δ	Δ	Δ	Δ	Δ
		$\downarrow$				$\downarrow$					
Δ	Δ	0	Δ	0	Δ	Δ	Δ				
		$\downarrow$				$\downarrow$					
Δ	Δ	Δ	0	Δ	0	Δ	Δ				
		$\downarrow$				$\downarrow$					
Δ	Δ	0	0	0	0	Δ	Δ				
$\alpha_\Delta^6 \rho_\Delta^6$				$\alpha_\Delta^6 \rho_\Delta^6$				$\alpha_\Delta^4 \rho_\Delta^4$			

Table 2: Basic characteristics for attacking RC6-I-NFR illustrating the contribution of the addition unit (probability α) and the rotation unit (probability ρ). Here Δ denotes a "generic" difference with specific choices for Δ being made in Table 3 to maximize the resultant probability and each row-to-row transition represents one round of encryption.

general					*a specific choice*			
e_t	e_t	0	0		e_{31}	e_{31}	0	0
		$\downarrow$					$\downarrow$	
e_t	0	0	0		e_{31}	0	0	0
		$\downarrow$					$\downarrow$	
0	0	0	e_s		0	0	0	e_{31}
		$\downarrow$					$\downarrow$	
0	e_u	e_s	0		0	e_{31}	e_{31}	0
		$\downarrow$					$\downarrow$	
e_u	e_u	0	e_v		e_{31}	e_{31}	0	e_{31}
		$\downarrow$					$\downarrow$	
e_u	e_u	e_v	0		e_{31}	e_{31}	e_{31}	0
		$\downarrow$					$\downarrow$	
e_u	e_u	0	0		e_{31}	e_{31}	0	0

Table 3: A generalized six-round characteristic and one particular embodiment for RC6-I-NFR. Note that the values to s and v are internal to the cipher thereby suggesting a likely role for differentials.

(a) we obtain a six-round characteristic that holds with probability $\rho^6 = 2^{-30}$ (see Table 3). Note that addition operations can be crossed with probability one when the single bit of difference is in the most significant bit of the word. For convenience we will drop the notation α_Δ and just use α where this is no ambiguity in the text. Note that for the variant RC6-I-NFR, most of the security against low Hamming weight differences is derived from the rotation unit.

More general characteristics can be constructed by allowing the difference to take different values while still maintaining the pattern in cycle (a). In Table 3 (left half) we illustrate such general characteristics. The only restriction on the values t, s, u, and v is that they lie between 5 and 31, i.e., the difference is not in the rotation amount. We see that any characteristic satisfying the above condition holds with probability $\alpha^6 \times \rho^6$.

The general characteristics presented in Table 3 allow us to form differentials for RC6-I-NFR. In particular, for a given starting t the variables s, u, and v can each take on one of $32 - 5 = 27$ choices. Hence, we obtain a six-round differential[1] which holds with probability

$$\alpha^6 \times \rho^6 \times 27^3 \approx 2^{-36} \times 27^3 \approx 2^{-22}.$$

Since this six-round differential is iterative, we can use it to construct r-round differentials for any value r. For example, to attack RC6-I-NFR with 20-rounds under our assumptions in Section 2.1.1, we will need an 18-round differential which can be obtained by concatenating three six-round differentials. This holds with probability $(2^{-22})^3 = 2^{-66}$.

Since the data requirements for a differential attack are proportional to the inverse of the probability of the differential, we can easily estimate the number of chosen plaintext pairs that are needed to attack RC6-I-NFR with a selected number of rounds (see Table 4). To derive these estimates for a different number of rounds we merely found the best window consisting of an $(r-2)$-round differential within the iterated differential given in Table 5. This is the technique we will use for all the estimates we derive for this simplified variants of RC6.

We remark that the most unimportant aspects of our results on RC6-I-NFR are the data requirements to mount an attack. It is trivial to find further improvements that will give substantial savings in the data required for an attack. Instead, this analysis is important because it points a way for developing attacks on the version of RC6 and it also highlights the effectiveness of differentials and how they might help to improve some attack beyond the level implied by the analysis of a single characteristic.

Differential Cryptanalysis of RC6-I-NFR

	number of rounds				
variant	8	12	16	20	24
RC6-I-NFR *using basic characteristic*	2^{30}	2^{50}	2^{65}	2^{90}	2^{110}
RC6-I-NFR *+ differential considerations*	2^{22}	2^{32}	2^{45}	2^{66}	2^{76}

Table 4: An estimate of the number of plaintexts needed to mount a differential attack on RC6-I-NFR with a varying number of rounds. The probability of the differential for eight rounds has been verified experimentally.

general					*a specific choice*			
e_{t+5}	e_t	0	0		e_{16}	e_{11}	0	0
	$\downarrow$					$\downarrow$		
e_t	0	0	0		e_{26}	0	0	0
	$\downarrow$					$\downarrow$		
0	0	0	e_s		0	0	0	e_{26}
	$\downarrow$					$\downarrow$		
0	e_u	e_s	0		0	e_{26}	e_{26}	0
	$\downarrow$					$\downarrow$		
e_u	e_{u-5}	0	e_v		e_{26}	e_{21}	0	e_{26}
	$\downarrow$					$\downarrow$		
e_{u-5}	e_{u-10}	e_v	0		e_{21}	e_{16}	e_{26}	0
	$\downarrow$					$\downarrow$		
e_{u-10}	e_{u-15}	0	0		e_{16}	e_{11}	0	0

Table 5: A generalized iterative six-round characteristic and one particular embodiment for RC6-I. The fixed rotation forces certain conditions on these internal variables so we have that the values of $t+5$, $s+5$, $v+5$, $u+5$, u, $u-5$, and $u-10$ all lie between 5 and 31.

	number of rounds				
variant	8	12	16	20	24
RC6-I *using basic characteristic*	2^{36}	2^{60}	2^{78}	2^{108}	2^{132}
RC6-I *+ differential considerations*	2^{23}	2^{34}	2^{47}	2^{69}	2^{80}

Table 6: An estimate of the number of plaintexts needed to mount a differential attack on RC6-I with a varying number of rounds.

4 Differential Cryptanalysis of RC6-I

In this section:

- *We show that RC6-I offers reasonable resistance to differential cryptanalysis but as many as 38 rounds might be necessary for the purposes of the AES*
- *We demonstrate that there is a very significant effect when considering differentials instead of characteristics*
- *We note that by adding the fixed rotation to RC6-I-NFR we change the course of the differentials, but not their effect*

The major difference between the variant RC6-I and RC6-I-NFR is the fixed rotation by five bits. Therefore, we will pay particular attention to how the fixed rotation affects the evolution of the characteristics and differentials in RC6-I.

Once again we will use exclusive-or as the measure for difference and use e_i to denote 2^i. Starting with the characteristics for RC6-I-NFR given in Table 3 we can construct characteristics for RC6-I. In Table 5, we demonstrate both a general and a specific six-round characteristic for RC6-I. We remark that even though these characteristics for RC6-I are similar to those for RC6-I-NFR it is harder to line up the bit differences within each word. As a result, there are more restrictions on the values to the variables t, s, v, and u. In particular, the fixed rotation provides some constraints to these variables and we need the values of $t+5$, $s+5$, $v+5$ and $u+5$, u, $u-5$, $u-10$ all to lie between 5 and

[1] We do not need to fix the value of u at the end of the sixth round.

31. This forces us to choose $0 \leq t, s, v \leq 26$ and $15 \leq u \leq 26$. It is easy to see that a general characteristic satisfying these constraints holds with probability $\alpha^6 \times \rho^6$. By setting $t = u - 15 = 11$ and $s = v = 26$ we obtain a specific iterative characteristic[2] that holds with probability 2^{-36}.

Taking into account the effect of differentials over these six rounds, we note that starting with a given value t, there are 27 possibilities for the values of s and v and 12 possibilities for the value u. So a useful six-round differential for attacking RC6-I holds with estimated probability

$$\alpha^6 \times \rho^6 \times 27^2 \times 12 = 2^{-36} \times 27^2 \times 12 \approx 2^{-23}.$$

Just as we did with RC6-I-NFR we can build $(r-2)$-round differentials using these six-round iterative differentials and mount a differential attack on r-round RC6-I. Table 6 gives an estimate for the resulting plaintext requirements. To derive these estimates for a different number of rounds we merely found the best window consisting of an $(r-2)$-round differential within the iterated differential given in Table 5.

5 The Quadratic Function

In this section:

- *We establish some of the technical tools for analysis of the quadratic function*
- *We conclude that integer subtraction is a better measure of difference for the analysis of RC6 than exclusive-or*

The introduction of the quadratic function is perhaps the major innovation that took place during the evolution from RC5 to RC6. The main security goal is to make the data-dependent rotation amount, which is derived from the output of the quadratic function, dependent on all bits of the input word. This should thwart existing differential attacks that apply to RC5.

To analyze the quadratic function under the differential framework, we need to choose an appropriate measure for difference. Since the operations involved are integer addition and multiplication, it seems quite natural to start with integer subtraction as the measure for difference. As we will see in the analysis of RC6 this measure turns out to be very useful, though for completeness and comparison we will also be considering the use of exclusive-or.

[2] If we set $t = 26$ instead, we could get a non-iterative characteristic that holds with probability 2^{-35}. As stated at the beginning of the chapter, however, we do not consider such optimizations for RC6 variants, but they will be taken into account during the analysis of the full RC6.

5.1 Basic properties of the quadratic function

We first show that the quadratic function $f(x) = x(2x+1) \bmod 2^{32}$ is a permutation by proving the following more general result.

Lemma 1 *Let $f(x) = x(ax+b) \bmod 2^w$ where a is even and b is odd. Then $f(x)$ is a one-to-one mapping from $\{0,1\}^w$ to $\{0,1\}^w$.*

Proof. By contradiction. Suppose for some $x_1 \neq x_2$ that $f(x_1) = f(x_2)$. Then, $ax_1^2 + bx_1 = ax_2^2 + bx_2 \bmod 2^w$. Combining terms we get

$$(x_1 - x_2)(ax_1 + ax_2 + b) \bmod 2^w = 0.$$

Since $(ax_1 + ax_2 + b)$ is odd $(x_1 - x_2)$ must be a multiple of 2^w. This is a contradiction. □

We now consider some basic differential properties of the quadratic function. For two inputs x_1 and x_2, let $y_1 = f(x_1)$ and $y_2 = f(x_2)$ and define

$$\begin{aligned} \delta_x &= x_2 - x_1, \\ \delta_y &= y_2 - y_1. \end{aligned}$$

Lemma 2 *For inputs x_1 and x_2, let $y_1 = f(x_1)$ and $y_2 = f(x_2)$. Define $\delta_x = x_2 - x_1$ and $\delta_y = y_2 - y_1$. Then*

$$\delta_y = (4x_1\delta_x + \delta_x + 2\delta_x^2) \bmod 2^{32}.$$

Proof. Straightforward evaluation of $f(x_1)$ and $f(x_2)$. □

Lemma 3 *Given the notation in Lemma 2, $\delta_y = \delta_x$ if, and only if, $4x_1\delta_x + 2\delta_x^2 = 0 \bmod 2^{32}$.*

Proof. Follows immediately from Lemma 2. □

While we will not use the following lemma directly, we feel that it is useful in providing some justification for the use of the high-order bits of the output from the quadratic function as a rotation amount, instead of any others.

Lemma 4 *Given an input x_1 chosen uniformly at random from $\{0,1\}^{32}$, let $g_{i,j}$ denote the probability that flipping bit i of x_1 will flip bit j of $y_1 = f(x_1)$. Then,*

$$\begin{aligned} g_{i,j} &= \begin{cases} 0 & \text{for } j < i, \\ 1 & \text{for } j = i, \\ 1 & \text{for } j = 1 \text{ and } i = 0, \text{ and} \end{cases} \\ g_{i,j} &\in [1/4, 3/4] \text{ for } j > i \geq 1 \text{ or } j \geq 2 \text{ and } i = 0. \end{aligned}$$

For the last case, $g_{i,j}$ is close to $3/4$ if $j = 2i+2$, and for most of the other i,j pairs $g_{i,j}$ is close to $1/2$.

Proof. Let $x_2 = x_1 \oplus 2^i$ and let $y_2 = f(x_2)$. Without loss of generality, we assume that $x_1[i] = 0$, i.e., the i^{th} bit of x_1 is zero. From Lemma 2, we obtain

$$\begin{aligned} \delta_y = y_2 - y_1 &= 2^{i+2}x_1 + 2^i + 2^{2i+1} \bmod 2^{32} && (1) \\ &\approx 2^{i+2}x_1 + 2^{2i+1} \bmod 2^{32} && (2) \\ &\approx 2^{i+2}x_1 \bmod 2^{32}. && (3) \end{aligned}$$

Now consider each case of the lemma separately.

Case 1: $j < i$. From Equation 1, we see that bit j of δ_y is zero. So flipping bit i of x_1 does not flip bit j of $f(x_1)$.

Case 2: $j = i$. Again using Equation 1, we know that bit i of δ_y is one. Hence flipping bit i of x_1 always flips bit i of $f(x_1)$.

Case 3: $j = 1$ and $i = 0$. Using Equation 1, $\delta_y = 4x_1 + 1 + 2$. Since $x_1[0] = 0$, $y_1[0] = 0$. Therefore, when computing $y_2 = y_1 + \delta_y$, there is no carry from bit 0 into bit 1. Since bit 1 of δ_y is one, flipping bit 0 of x_1 always flips bit 1 of $f(x_1)$.

Case 4: $j > i \geq 1$ or $j \geq 2, i = 0$. We first consider the case where $j = i + 1$. From Equation 1, bit $i + 1$ of δ_y is always zero and bit i of δ_y is always one. When computing $y_2 = y_1 + \delta_y$, there is a carry into bit $i + 1$ with probability approximately $1/2$, and hence $g_{i,i+1} \approx 1/2$.

We next consider the case where $j = 2i + 2$. Since $x_1[i] = 0$, bit $2i + 2$ of $2^{2i+1}x$ is zero. We analyze the carry effect when computing $y_2 = y_1 + \delta_y$ using approximation 2. If $x_1[i - 1] = 1$, then there is always a carry into bit $2i + 2$ by due to the term 2^{2i+1}. If $x_1[i - 1] = 0$, then there is a carry with probability about $1/2$. Overall, there is a carry with probability about $3/4$, and so $g_{i,j} \approx 3/4$.

We finally consider the case where $j \geq i + 2$ and $j \neq 2i + 2$. Using Approximation 3, we obtain that bit j of δ_y is random since x_1 is random. So flipping bit i of x_1 flips bit j of $f(x_1)$ with probability about $1/2$. The inexact value is due to the presence of a carry effect.

Experimental results showed that $g_{i,j}$ ranges between $1/4$ and $3/4$, and most of the probabilities are very close to $1/2$ especially when $i \geq 16$. □

Lemma 4 demonstrates that even a change of a single bit is likely to change at least one of the high-order bits that will be used as a rotation amount. This will help provide a very fast avalanche of change. This lemma is also interesting because it is closely related to the results presented in Lemma 6 and Lemma 7.

5.2 Using integer subtraction as a measure of difference

In this section, we will study characteristics of the quadratic function that have the form $\delta_y = \delta_x$. These characteristics are quite useful since they can easily be joined to give characteristics with a similar form as those derived in Section 3. We will call these characteristics *static characteristics* for the quadratic function. Other forms of characteristic will be discussed in Section 5.5.

Lemma 5 *If δ_x is odd then the characteristic $\delta_y = \delta_x$ holds with probability zero.*

Proof. Since δ_x is odd, $2\delta_x^2 = 2 \bmod 4$. Hence $4x_1\delta_x + 2\delta_x^2$ is never equal to zero modulo 2^{32}. □

Lemma 6 *If $\delta_x = v2^i$ for some odd integer v and $1 \leq i \leq 30$, then the characteristic $\delta_y = \delta_x$ holds with probability 2^{i-30}. If $\delta_x = 2^{31}$, then the characteristic $\delta_y = \delta_x$ holds with probability one.*

Proof. From Lemma 3, $\delta_y = \delta_x$ if, and only if, $4x_1v2^i + 2(v2^i)^2 = 0 \bmod 2^{32}$. For $1 \leq i \leq 15$ this is equivalent to $x_1 + v2^{i-1} = 0 \bmod 2^{30-i}$ and for $16 \leq i \leq 30$ it is equivalent to $x_1 = 0 \bmod 2^{30-i}$. Note that in either case, the most significant $(i+2)$ bits of x_1 can take any value. The only constraint imposed is on the least $(30-i)$ bits of x_1. Therefore, averaging over all possible values of x_1, the probability that $\delta_y = \delta_x$ is $\frac{2^{i+2}}{2^{32}} = 2^{i-30}$. □

The two lemmas just given cover all possible static characteristics, and they show that the probability of such characteristics depends on the input difference δ_x. The characteristics that we will use to analyze RC6 and RC6-NFR are those in which the integer difference has Hamming weight one (i.e., $\delta_y = \delta_x = 2^i$). It is worth observing that for these characteristics, the probability the characteristic holds drops quickly as i, the bit position involved in the approximation, moves to less significant bit positions. As we will see, this property of the quadratic function is critical in reducing the threat of differential cryptanalysis.

5.3 Using exclusive-or as a measure of difference

In this section we will consider counterparts to the single-bit, static characteristics $\delta_y = \delta_x = 2^i$ that we studied in Section 5.2. These counterparts are formed by using exclusive-or as the measure of difference. Following the notation established in Section 5.1 we define

$$\begin{aligned} \delta_x^{\oplus} &= x_2 \oplus x_1, \text{ and} \\ \delta_y^{\oplus} &= y_2 \oplus y_1. \end{aligned}$$

Lemma 7 *Let p_i be the probability of the characteristic $\delta_y^{\oplus} = \delta_x^{\oplus} = 2^i$. Then*

$$p_i = \begin{cases} 1 & \text{for } i = 31 \\ 2^{i-30} & \text{for } 15 \leq i \leq 20 \\ 0 & \text{for } 0 \leq i \leq 14 \end{cases} \tag{4}$$

$$\text{and } p_i \approx 2^{i-31} \text{ for } 21 \leq i \leq 30. \tag{5}$$

Proof. There are two situations to consider when we have $\delta_x^{\oplus} = 2^i$. Either (a) $x_2 = x_1 + 2^i$ and $x_1[i] = 0$, or (b) $x_1 = x_2 + 2^i$ and $x_2[i] = 0$. Since x_1 is

uniformly distributed, each case happens with probability 1/2. By symmetry arguments we only need to consider (a). We will divide the proof into several cases according to the value of i.

Case 1: $16 \leq i \leq 30$. Based on Lemma 2, we have

$$\begin{aligned} y_2 - y_1 &= 2^{i+2}x_1 + 2^i + 2^{2i+1} \bmod 2^{32} \\ &= 2^{i+2}x_1 + 2^i \bmod 2^{32} \text{ (since } i \geq 16). \end{aligned}$$

Note that $\delta_y^{\oplus} = 2^i$ if, and only if, both of the following events hold.

$$\begin{aligned} &\text{Event A} \quad 2^{i+2}x_1 = 0 \bmod 2^{32}. \\ &\text{Event B} \quad y_1[i] = 0. \end{aligned}$$

Following the argument in Lemma 6, we obtain $prob(A) = 2^{i-30}$. Therefore, the probability of the characteristic $\delta_y^{\oplus} = \delta_x^{\oplus} = 2^i$ is equal to

$$p_i = prob(A) \times prob(B|A) = 2^{i-30} \times prob(B|A).$$

Let $q_i = prob(B|A)$. We will show that $q_i = 1$ when $16 \leq i \leq 20$ and $q_i \approx 1/2$ when $21 \leq i \leq 30$. This will be sufficient to prove this case.

If $16 \leq i \leq 20$, then Event A implies that the lower $(30-i)$ bits of x_1 are 0. Hence, the lower $2(30-i)+1 = 61-2i$ bits of $2x_1^2$ are zero. Since $y_1 = 2x_1^2 + x_1$ we have $y_1[i] = x_1[i] = 0$.

If $21 \leq i \leq 30$, then Event A implies that the upper $(i+2)$ bits of x_1 are random. Hence, bits $(61-2i)$ through 31 of $2x_1^2$ are approximately "random", and bit i falls into this range when $21 \leq i \leq 30$. Since $y_1 = 2x_1^2 + x_1$ we have that with probability approximately 1/2, $y_1[i] = x_1[i] = 0$.

This counter-intuitive approximation is due to the fact that each bit of x^2 is not uniformly distributed even when x is uniformly distributed. The values of p_i for $21 \leq i \leq 30$ are given by the following table.

i	2^{31-i}	$1/p_i$
21	1024	819.200000
22	512	455.111111
23	256	248.242424
24	128	126.025089
25	64	63.750731
26	32	31.938838
27	16	15.992672
28	8	7.998251
29	4	3.999823
30	2	1.999954

Case 2: $i = 15$. This case follows essentially the same argument as presented for Case 1 but the details are slightly different. Setting $\delta_x = 2^{15}$ in

Lemma 2, we have

$$y_2 - y_1 \quad = \quad 2^{17}x_1 + 2^{15} + 2^{31} \bmod 2^{32}.$$

Note that $\delta_y^{\oplus} = 2^{15}$ if, and only if, both of the following events hold.

$$\begin{array}{ll} \text{Event A} & 2^{17}x_1 + 2^{31} = 0 \bmod 2^{32}. \\ \text{Event B} & y_1[15] = 0. \end{array}$$

Following the argument in Lemma 6, we obtain

$$prob(A) = 2^{-15}.$$

Therefore, the probability of the characteristic $\delta_y^{\oplus} = \delta_x^{\oplus} = 2^{15}$ is equal to

$$p_{15} = prob(A) \times prob(B|A) = 2^{-15} \times prob(B|A).$$

Note that Event A implies that the lower 14 bits of x_1 are 0 and that bit 14 is 1. Hence, the lower $(2 \times 14) + 1 = 29$ bits of $2x_1^2$ are zero. Since $y_1 = 2x_1^2 + x_1$, we have $y_1[15] = x_1[15] = 0$. Therefore, $p_{15} = 2^{-15} = 2^{i-30}$.

Case 3: $1 \leq i \leq 14$. In this case, we have

$$y_2 - y_1 = 2^{i+2}x_1 + 2^{2i+1} + 2^i \bmod 2^{32}.$$

Now, $y_2 - y_1$ can be 2^i if, and only if,

$$x_1 = -2^{i-1} \bmod 2^{30-i}.$$

This condition says that bits $(i-1)$ through $(29-i)$ of x_1 must be 1. But this contradicts the initial assumption that $x_1[i] = 0$.

Case 4: $i = 0$ or $i = 31$. For $i = 0$, the result follows from Lemma 5. For $i = 31$, the result is straightforward since integer subtraction and exclusive-or are the same when the difference is in the most significant bit. □

We suspect the reader will find part of the result in Lemma 7 surprising. As the next section will show, it does however give us the reason why we prefer to use integer subtraction as the measure of difference when analyzing RC6 instead of exclusive-or.

5.4 Comparing integer subtraction and exclusive-or

In the preceeding sections we studied static characteristics for the quadratic function using both integer subtraction and exclusive-or as the measure of difference. Based on Lemma 6 and Lemma 7 we can now compare the probabilities of these characteristics under both difference measures in Table 7.

From this table it is clear that the probabilities for characteristics using integer subtraction are slightly higher than when using exclusive-or. More importantly, for small values of i (that is, with $i \leq 14$), we can only use integer

bit position i	*probability* $\delta_y = \delta_x = 2^i$	*probability* $\delta_y^{\oplus} = \delta_x^{\oplus} = 2^i$
31	1	1
$21 \leq i \leq 30$	2^{i-30}	$\approx 2^{i-31}$
$15 \leq i \leq 20$	2^{i-30}	2^{i-30}
$1 \leq i \leq 14$	2^{i-30}	0
0	0	0

Table 7: The probabilities of static single-bit characteristics for the quadratic function. Probabilities are given using both integer subtraction and exclusive-or as the measure of difference.

subtraction since the probabilities for the relevant characteristics become zero when exclusive-or is used.

Before concluding that integer subtraction should be used to analyze RC6, however, there is one more factor that we need to take into account. We need to decide whether the chosen measure of difference will also work well with the other operations present in RC6.

In particular, we will now consider how a single-bit difference under integer subtraction propagates through the exclusive-or. At first glance, one might think that to compute the probability of a characteristic across the function $w = z \oplus f(x)$, one could just compute the probability across "$\oplus$" and then join it with the probability for the characteristic across $f(x)$ (which was derived in Lemma 6). However, the quadratic function $f(x)$ has some complicated differential behavior and it is better to study the function $w = z \oplus f(x)$ as a single component.

For two sets of inputs x_1, z_1 and x_2, z_2 we define $w_1 = z_1 \oplus f(x_1)$ and $w_2 = z_2 \oplus f(x_2)$. Now set $\delta_x = x_2 - x_1$, $\delta_z = z_2 - z_1$, and $\delta_w = w_2 - w_1$.

Lemma 8 *Let p_i be the probability of the characteristic $(\delta_x, \delta_z) \rightarrow \delta_w$ where $(\delta_x, \delta_z) = (2^i, 0)$ and $\delta_w = 2^i$. Similarly, let q_i be the probability of the characteristic $(\delta_x, \delta_z) \rightarrow \delta_w$ where $(\delta_x, \delta_z) = (2^i, 2^i)$ and $\delta_w = 0$. Then we always have that $p_i = q_i$ and further that*

$$p_i = q_i = \begin{cases} 2^{i-31} & \text{for } 15 \leq i \leq 31, \\ x \in [2^{i-35}, 2^{i-30}] & \text{for } 0 \leq i \leq 14. \end{cases}$$

Proof. We will describe experiments and related arguments needed for computing the probability p_i. The probability q_i can be computed in a similar way. In our experiments, for each value of i between 0 and 31 we count the number of times that $\delta_w = 2^i$ given that $(\delta_x, \delta_z) = (2^i, 0)$. This is for both x_1 and z_1 ranging over all possible 32-bit words (a total of 2^{64} words). This initially appears to be impractical, but the following useful observations make it possible:

1. $\delta_w = 2^i$ only if $f(x_2) \oplus f(x_1)$ contains exactly one block of consecutive 1-bits. This block will always begin at bit i and end at some bit k ($k \geq i$). So the length of the block is $(k - i + 1)$ bits.

2. Given $f(x_1)$ and $f(x_2)$ satisfying the first condition, the number of words $z_1 = z_2$ for which $\delta_w = 2^i$ is

$$\begin{array}{rl} 2^{32-(k-i+1)} & \text{if } k < 31, \text{ and} \\ 2^{32-(k-i)} & \text{if } k = 31. \end{array}$$

Experiments show that p_i is exactly 2^{i-31} for $15 \leq i \leq 31$. For smaller values of i, however, experimental results are given in the following table.

i	2^{31-i}	$1/p_i$	i	2^{31-i}	$1/p_i$
0	2^{31}	$2^{30.10}$	1	2^{30}	$2^{29.15}$
2	2^{29}	$2^{28.57}$	3	2^{28}	$2^{28.17}$
4	2^{27}	$2^{27.71}$	5	2^{26}	$2^{27.21}$
6	2^{25}	$2^{26.89}$	7	2^{24}	$2^{26.42}$
8	2^{23}	$2^{25.92}$	9	2^{22}	$2^{25.44}$
10	2^{21}	$2^{24.98}$	11	2^{20}	$2^{23.09}$
12	2^{19}	$2^{19.58}$	13	2^{18}	$2^{18.10}$
14	2^{17}	$2^{17.01}$			

□

We remark that the probability p_i decreases monotonically as i decreases. However, the rate of decrease is not constant and there is a big drop in the probability from p_{12} to p_{11}. Once again the reader might be surprised at the "non-uniform" probabilities given in Lemma 8 and wonder why the results in Lemma 6 and Lemma 7 do not imply or at least provide a basis for proving the results in Lemma 8. One main reason appears to be that given $f(x_2) - f(x_1) = 2^i$, it is not clear what the distribution of $f(x_2) \oplus f(x_1)$ is. Indeed, Lemma 7 suggests that the distribution is not uniform for all i. Moreover, the exclusive-or of different values of z might further alter the distribution.

Despite the irregularities highlighted by Lemma 8 it is still a very useful lemma. In fact, from Lemma 8 it is reasonable to conclude that integer subtraction is a better notion of difference than exclusive-or during the differential cryptanalysis of both RC6 and RC6-NFR.

5.5 Other characteristics for the quadratic function

In this section, we will study some non-static characteristics for the quadratic function, and we will use integer subtraction as the measure of difference. These characteristics will illustrate some interesting properties of the quadratic function with respect to differential cryptanalysis.

The single-bit static characteristics that we studied in earlier sections have the general form of $\delta_x = \delta_y = 2^i$. Below, we consider characteristics in which a two-bit difference in the input will only yield a single-bit difference in the output. In some sense, the quadratic function is being used to simplify the differences in certain situations.

We extend the notation previously established and let $e_{i,j}$ denote $2^i + 2^j$ for some $i \neq j$.

Lemma 9 *Let $m_{i,j}$ be the probability of the characteristic $\delta_x \rightarrow \delta_y$ where $\delta_x = e_{i,j}$ and $\delta_y = 2^i$ for some $i \neq j$. Then we have*

$$m_{i,j} = \begin{cases} 2^{i-30} & \text{for } 29 \geq i \geq 1 \text{ and } j \geq i+2, \\ 0 & \text{for } i = 0 \text{ and } j \geq i+2, \\ 0 & \text{for } 30 \geq i \geq 1 \text{ and } j = i+1, \\ 2^{i-30} & \text{for } i = 0 \text{ and } j = i+1, \\ 0 & \text{for } 31 \geq i \geq 1 \text{ and } j < i. \end{cases}$$

Proof. From Lemma 2 we have

$$\delta_y = x_1(2^{i+2} + 2^{j+2}) + (2^i + 2^j) + 2(2^i + 2^j)^2 \bmod 2^{32}. \tag{6}$$

We will divide the proof into three cases according to the value of j.

Case 1. Suppose $j \geq i + 2$. From Equation 6, $\delta_y = 2^i$ if, and only if,

$$2^{i+2}x_1(1 + 2^{j-i}) + [2^j + 2(2^i + 2^j)^2] = 0 \bmod 2^{32}.$$

Let $a = 1 + 2^{j-i}$ and $b = 2^j + 2(2^i + 2^j)^2$. Since a is odd, $\delta_y = 2^i$ if, and only if,

$$2^{i+2}x_1 + a^{-1}b = 0 \bmod 2^{32}. \tag{7}$$

If $i \geq 1$, then $b \geq 2^{i+2}$. So $\delta_y = 2^i$ if, and only if, $x_1 = a^{-1}b \bmod 2^{30-i}$. The $(i+2)$ most significant bits of x_1 can take any value and the least $(30 - i)$ bits of x_1 are fixed. Averaging over all possible x_1, we have that the probability $m_{i,j} = 2^{i-30}$. If $i = 0$, then $b = 2 \bmod 4$, so Equation 7 never holds and $m_{i,j} = 0$ in this case.

Case 2. Suppose $j = i + 1$. Then based on Equation 6, we have for some integer c that

$$\begin{aligned} \delta_y &= x_1(2^{i+2} + 2^{i+3}) + 2^{i+1} + 2(2^i + 2^{i+1})^2 \bmod 2^{32} \\ &= 2^i + 2^{i+1} + 2^{2i+1} + 2^{i+2}c. \end{aligned}$$

If $i \geq 1$, then bit $(i+1)$ of δ_y is always one, and so δ_y is never equal to 2^i. In this case, $m_{i,j} = 0$. If $i = 0$, then the two terms 2^{i+1} and 2^{2i+1} "cancel" each other. A simple calculation shows that $\delta_y = 1 + 4(3x + 5) \bmod 2^{32}$. Similar to Case 1, we obtain that $m_{i,j} = 2^{-30} = 2^{i-30}$.

Differential Cryptanalysis of RC6-NFR					
	number of rounds				
variant	8	12	16	20	24
RC6-NFR *using basic characteristic*	2^{30}	2^{50}	2^{65}	2^{90}	2^{110}
RC6-NFR *+ differential considerations*	2^{28}	2^{47}	2^{61}	2^{84}	2^{103}

Table 8: An estimate of the number of plaintexts needed to mount a differential attack on RC6-NFR with a varying number of rounds.

Case 3. Suppose $31 \geq i \geq 1$ and $j < i$. In this case, it is easy to see from Equation 6 that bit j of δ_y is always one. So the probability $m_{i,j} = 0$. □

We will briefly discuss in Section 7.5 how the characteristics derived in this section can be used to attack RC6. Note that the number of bits in the difference is reduced by the action of these characteristics and so they might potentially be useful in controlling the avalanche of change as one moves from one round to another. However we have not been able to use them in a way that yields better attacks than those based on static characteristics.

6 Differential Cryptanalysis of RC6-NFR

In this section:

- *We show that RC6-NFR displays reasonable resistance to differential cryptanalysis but as many as 30 rounds may be necessary for the purposes of the AES*
- *We demonstrate that introduction of the quadratic function greatly reduces the effect of differentials*

As established in Section 5.4, we will now use integer subtraction as the measure of difference in attacking RC6-NFR.

We begin with the two characteristics derived in Lemma 8 and adopt the notation used there. That is, we set p_i to denote the probability of the characteristic $(\delta_x, \delta_z) \rightarrow \delta_w$ where $w = z \oplus f(x)$, $\delta_x = \delta_w = 2^i$, and $\delta_z = 0$. We also use q_i to denote the probability of the characteristic $(\delta_x, \delta_z) \rightarrow \delta_w$ where

$\delta_x = \delta_z = 2^i$, and $\delta_w = 0$. It is clear that we can combine these two characteristics in a manner reminiscent of Section 3 since the characteristics for the quadratic function component are static.

Let us consider the general characteristic from Table 3. The probability of this six-round characteristic for RC6-NFR is

$$\rho^6 \times (p_t \times p_s \times p_v \times p_u \times q_u^2).$$

Note that the addition unit can be crossed with probability one since we are using integer subtraction as measure of difference. The only restriction on the values of t, s, u, and v is that they lie between 5 and 31. We obtain a six-round characteristic that holds with probability $\rho^6 = 2^{-30}$ by setting $t = s = v = u = 31$ since $p_{31} = q_{31} = 1$.

It is worth mentioning that the probability of this characteristic for RC6-NFR is the same as the corresponding characteristic for RC6-I-NFR. It might therefore be tempting to question the relevance of the quadratic function since its inclusion appears to have no relevance for this particular six-round characteristic. In response we will show that the effect of the quadratic function on the performance of differentials is particularly significant.

Consider the differential effect over these six rounds. Starting with a given value t there are $32-5 = 27$ values for s, u, and v. However the probabilities p_s, p_v, p_u, and q_u are not the same for these different permissible values. In fact, Lemma 8 shows that the probabilities drop quickly as the indexing variable gets smaller. In order to maximize the probability of the differential we set $t = u = 31$ and we let s and v vary. Based on Lemma 8, the probability of this six-round differential is given by

$$\rho^6 \times \left(\sum_{i=5}^{31} p_i \right)^2 \approx 2^{-28}.$$

We can see that the differential effect only boosts the probability by a factor of about $2^2 = 4$ due to the presence of the quadratic function. For RC6-I-NFR, however, the factor of increase was $27^3 \approx 2^{14}$. With RC6-I-NFR we have a great many equally viable paths through the cipher but the introduction of the quadratic function has ensured that the differential is dominated by the action of a single characteristic.

As with the earlier variants, we give estimates for the plaintext requirements in attacking RC6-NFR using the six-round iterative differential we previously described. The results are summarized in Table 8.

7 Differential Cryptanalysis of RC6

In this section:

- *We show that RC6 offers good resistance to differential cryptanalysis and that 20 rounds are sufficient for the purposes of the AES*
- *We demonstrate that the combination of the quadratic function and the fixed rotation greatly hinders the construction of effective differentials*
- *We note that the use of customized differentials helps reduce the plaintext requirements in an attack*

We saw in Section 5 that the use of the quadratic function alone is not sufficient to hinder differential cryptanalysis since there remain good characteristics across the quadratic function. In this section, we will demonstrate that it is the combination of the quadratic function and the fixed rotation by five bit positions that ensures that an attacker is hindered from finding good characteristics with which to efficiently attack RC6.

In the following analysis we will again focus on one-bit characteristics and differentials. As previously we will derive iterative characteristics and differentials for RC6 though we will also derive some non-iterative, customized, differentials with which to refine our analysis. Finally, we will consider the possibility of multiple-bit characteristics and differentials.

7.1 Iterative characteristics and differentials for RC6

Based on Lemma 8 we can easily construct the following one-bit characteristics for one round of RC6.

$$\begin{array}{cccc} 0 & 0 & e_{t+5} & e_t \\ & & \downarrow & \\ 0 & 0 & e_t & 0 \end{array} \qquad \text{and} \qquad \begin{array}{cccc} 0 & 0 & 0 & e_t \\ & & \downarrow & \\ 0 & e_s & e_t & 0 \end{array}$$

The left-hand characteristic holds with probability q_t while the right-hand characteristic holds with probability $\rho \times p_t$ where q_t and p_t are given by Lemma 8. Note that the data-dependent rotation allows the choice of bit position s to be made independently of the choice of t.

Using these two one-round characteristics for RC6 as components in the six-round characteristic for RC6-I from Table 5, we readily obtain a six-round iterative characteristic for RC6. This characteristic is given in Table 9. The constraints on the variables due to the fixed rotation are the same as for the analysis of RC6-I so we have that $0 \leq t, s, v \leq 26$ and $15 \leq u \leq 26$. Any characteristic satisfying these constraints holds with probability

$$\rho^6 \times (q_t \times p_s \times p_v \times p_u \times q_{u-5} \times q_{u-10}).$$

general					*a specific choice*			
e_{t+5}	e_t	0	0		e_{16}	e_{11}	0	0
	$\downarrow$					$\downarrow$		
e_t	0	0	0		e_{11}	0	0	0
	$\downarrow$					$\downarrow$		
0	0	0	e_s		0	0	0	e_{26}
	$\downarrow$					$\downarrow$		
0	e_u	e_s	0		0	e_{26}	e_{26}	0
	$\downarrow$					$\downarrow$		
e_u	e_{u-5}	0	e_v		e_{26}	e_{21}	0	e_{26}
	$\downarrow$					$\downarrow$		
e_{u-5}	e_{u-10}	e_v	0		e_{21}	e_{16}	e_{26}	0
	$\downarrow$					$\downarrow$		
e_{u-10}	e_{u-15}	0	0		e_{16}	e_{11}	0	0

Table 9: A generalized iterative six-round characteristic for RC6. The corresponding differential is denoted I_6 in the text. The fixed rotation restricts the values of internal variables, the measure of difference is integer subtraction, and the probability is much less than for the characteristic given in Table 5.

Now consider the six-round differentials based on this family of six-round characteristics. Starting with a given value t, there are 27 possible values (0 through 26) for each of s and v. To maximize the probability of the differential, we set $t = 11$ and $u - 15 = 11$ though s and v will still be free. Based on the probabilities derived in Lemma 8, we can compute the probability of this six-round differential as

$$\begin{aligned} & \rho^6 \times q_{11} \times \left(\sum_{i=0}^{26} p_i \right)^2 \times p_{26} \times q_{21} \times q_{16} \\ \approx \quad & 2^{-30} \times 2^{-23} \times (2^{-4})^2 \times 2^{-5} \times 2^{-10} \times 2^{-15} \\ = \quad & 2^{-91}. \end{aligned}$$

Note that the effect of differentials has only provided an increase by a factor of $2^2 = 4$ over the probability of the characteristic. That is, by considering the 27^2 additional paths generated by the choices to s and v the probability of the differential is only a factor of 4 better than consideration of the single best characteristic. Nevertheless, we will find that this six-round iterative differential is very useful and we denote it I_6 for ease of reference in the rest of this section.

general					*a specific choice* (E'_6)			
e_{t+5}	e_t	0	0		e_{31}	e_{26}	0	0
		$\downarrow$					$\downarrow$	
e_t	0	0	0		e_{26}	0	0	0
		$\downarrow$					$\downarrow$	
0	0	0	e_s		0	0	0	e_{26}
		$\downarrow$					$\downarrow$	
0	e_u	e_s	0		0	e_{26}	e_{26}	0
		$\downarrow$					$\downarrow$	
e_u	e_{u-5}	0	e_v		e_{26}	e_{21}	0	e_{26}
		$\downarrow$					$\downarrow$	
e_{u-5}	e_w	e_v	0		e_{21}	e_{26}	e_{26}	0
		$\downarrow$					$\downarrow$	
e_w	e_x	0	$e_{u-5+y} \oplus e_{w+5+y}$		e_{26}	e_{31}	0	$e_{21} \oplus e_{31}$

Table 10: A generalized non-iterative six-round characteristic for RC6 and one particular embodiment. Note that by an appropriate choice of the parameter t (namely $t = 11$) we can join the associated differential to this characteristic to the end of I_6 in Table 9. The differential is denoted in the text by E'_6.

7.2 Non-iterative customized differentials for RC6

It is easy to construct r-round differentials for any value of r once we have an iterative differential. For instance, $I_6 - I_6 - I_6$ is an 18-round differential with which we can attack 20-round RC6. However, we will try to refine our analysis by constructing non-iterative differentials that have higher probabilities than their iterative counterparts. These will then be used at the beginning and the end of a differential instead of the less efficient iterative differentials[3]. We will refer to these as customized differentials.

We will now study two-round, four-round, and six-round non-iterative differentials that can be joined to the differential I_6 along with a six-round differential that can be used at the beginning of the differential I_6. In other words, we will relax the conditions on the differentials at the beginning and at the end. We will focus here on the six-round non-iterative differential that appends to I_6 since it is the one that is the most illustrative of our approach. The other useful differentials are given in Table 11.

In Table 10 we present some non-iterative six-round characteristics for RC6. Once again the first four rounds follow the same pattern as the iterative characteristic shown in Table 5. However the last two rounds are slightly different and

[3]We did not consider this line of analysis for the simplified RC6 variants since our focus there was on an analysis of the basic structure and the constituent operations.

B_2 (2^{-40})			
e_{26}	e_{21}	0	e_{26}
		$\downarrow$	
e_{21}	e_{16}	e_{26}	0
		$\downarrow$	
e_{16}	e_{11}	0	e_s

E_2 (2^{-23})			
e_{16}	e_{11}	0	0
		$\downarrow$	
e_{26}	0	0	0
		$\downarrow$	
0	0	0	e_{26}

B_4 (2^{-63})			
0	0	0	e_{26}
		$\downarrow$	
0	e_u	e_{26}	0
		$\downarrow$	
e_u	e_{21}	0	e_v
		$\downarrow$	
e_{21}	e_{16}	e_v	0
		$\downarrow$	
e_{16}	e_{11}	0	0

E_4 (2^{-41})			
e_{16}	e_{11}	0	0
		$\downarrow$	
e_{26}	0	0	0
		$\downarrow$	
0	0	0	e_s
		$\downarrow$	
0	e_{26}	e_s	0
		$\downarrow$	
e_{26}	e_{21}	0	e_{26}

B_6 (2^{-76})			
e_{31}	e_{26}	0	0
		$\downarrow$	
e_{26}	0	0	0
		$\downarrow$	
0	0	0	e_s
		$\downarrow$	
0	e_{26}	e_s	0
		$\downarrow$	
e_{26}	e_{21}	0	e_v
		$\downarrow$	
e_{21}	e_{16}	e_v	0
		$\downarrow$	
e_{16}	e_{11}	0	0

E_6 (2^{-71})			
e_{16}	e_{11}	0	0
		$\downarrow$	
e_{11}	0	0	0
		$\downarrow$	
0	0	0	e_s
		$\downarrow$	
0	e_{26}	e_s	0
		$\downarrow$	
e_{26}	e_{21}	0	e_v
		$\downarrow$	
e_{21}	e_{26}	e_v	0
		$\downarrow$	
e_{26}	e_{31}	0	$e_{21} \oplus e_{31}$

Table 11: Some useful two-, four-, and six-round differentials for RC6. Those denoted by B_r can be used to attach to the beginning of the differential I_6 (Table 9) and those denoted by E_r can be attached to the end of I_6. The unspecified arguments illustrate a limited differential effect. The probability of the differential is given in parentheses.

allow a more complicated pattern in the output difference (with an improved probability). The constraints on the relevant variables are $0 \leq t, s, v, w \leq 26$, $5 \leq u \leq 26$, and $0 \leq x, y \leq 31$. The probability of this characteristic is given by

$$\rho^7 \times (q_t \times p_s \times p_u \times p_v \times q_{u-5} \times p_w).$$

Let us consider some specific choices for the values of these variables.

1. By setting $t = 11$, $s = u = v = w = 26$ we obtain a six-round characteristic with probability $2^{-35} \times 2^{-20-(4\times 5)-10} = 2^{-85}$. The values of x and y are free. Note that the input difference of the associated differential is the same as the output difference of differential I_6 and so we will be able to append it to I_6.

2. If we set $t = s = u = v = w = 26$ then we obtain a better six-round characteristic with probability $2^{-35} \times 2^{-(5\times 5)-10} = 2^{-70}$. Again, x and y can take any value. This characteristic is shown in the right half of Table 10. Since $t \neq 11$ this six-round characteristic and associated differential cannot be appended to I_6. However, as a differential it will be useful on its own as a basis for an attack on an eight-round version of RC6.

In moving from the characteristics given above to differentials we allow the variables s, u, v, and w to range between 0 and 26. As we saw in the derivation of the differential I_6, by allowing a single variable to cover the range between 0 and 26 this increases the probability of the differential over that of the characteristic by a factor of 2. Since there are in fact four free variables, the factor increase becomes 2^4.

In providing estimates for the security of RC6 we go further and try to anticipate some of the tricks that might be played by a cryptanalyst in reducing the plaintext requirements for attacking the cipher. One approach we have considered is to relax the conditions on the output difference by looking for a difference of Hamming weight one in strand B and weight two in strand C instead of looking explicitly for a difference in certain bit positions as predicted by the differential. By doing this, we are no longer concerned about predicting the rotation amount used in the last round and this immediately increases the probability by a factor of 2^{10}.

Combining both the differential and optimization considerations we therefore suggest that the following two differentials will be useful when joined to the differential I_6 or standing alone in an attack on eight-round RC6.

1. Set $t = 11$ and allow other internal variables to range over all possible values. The differential (with optimizing trick) holds with probability $2^{-85} \times 2^4 \times 2^{10} = 2^{-71}$. We will denote this differential by E_6. This differential can be joined with I_6 to give $I_6 - E_6$.

	number of rounds				
variant	8	12	16	20	24
RC6 *iterative characteristic*	2^{93}	2^{151}	2^{214}	2^{279}	2^{337}
RC6 *iterative differential*	2^{91}	2^{147}	2^{210}	2^{273}	2^{329}
RC6 *using customized differentials*	2^{56} E'_6	2^{117} $B_6 - E_4$	2^{190} $B_6 - I_6$ $-E_2$	2^{238} $B_6 - I_6$ $-E_6$	2^{299} $B_6 - I_6$ $-I_6 - E_4$

Table 12: An estimate of the number of plaintexts needed to mount a differential attack on RC6 with a varying number of rounds. These attacks are based around a six-round iterative differential I_6 joined at the beginning and/or end to a choice of customized differential.

2. Set $t = 31$ and allow other internal variables to range over all possible values. The differential (with optimizing trick) holds with probability $2^{-70} \times 2^4 \times 2^{10} = 2^{-56}$. We will denote this differential by E'_6. This differential cannot be joined with I_6 so it is only useful in an attack on eight-round RC6.

In a similar fashion it is possible to construct two- and four-round non-iterative differentials (see Table 11) that can be joined to the end of I_6. Analysis has shown that the two-round differential (denoted by E_2) holds with probability 2^{-23} and the four-round differential (denoted by E_4) holds with probability 2^{-41}.

Finally, we construct a six-round non-iterative differential that can be used at the beginning of an r-round differential attack. This differential replaces the first invocation of I_6. By setting $t = 26$ instead of $t = 11$, we boost the probability of I_6 by a factor of 2^{15}. The resulting six-round differential (denoted by B_6) therefore holds with probability 2^{-76}. We have also considered the action of the differentials B_2 and B_4 presented in the Table 11, though we find that they are less useful.

7.3 Attacking r-round RC6

Given the iterative and non-iterative differentials derived in the preceeding sections, we are now ready to construct r-round differentials for RC6. In particular, to attack 20-round RC6 we will use the 18-round differential $B_6 - I_6 - N_6$ which holds with probability $2^{-76} \times 2^{-91} \times 2^{-71} = 2^{-238}$. Estimates for the plaintext requirements in attacking a different number of rounds of RC6 are given in Table 12.

7.4 Other interesting differentials

We have found some other interesting differentials during the analysis of RC6. The general three-round iterative characteristic shown in Table 13 holds with probability $r^4 \times q_t \times q_u \times p_s \times p_v$. By setting $t = u = 21$ and $s = v = 26$, we obtain a specific characteristic that holds with probability 2^{-50}. Unlike the six-round iterative characteristic described in Section 7.1, there is essentially no accompanying differential effect for this characteristic. This is because there are no free variables after setting the values of t, u, s and v. Extending this characteristic to more rounds, we easily obtain a six-round characteristic that holds with probability 2^{-100}. This is about a factor of 2^9 smaller than the probability of six-round iterative characteristic that we used to attack RC6. While interesting, we feel that using this differential isn't as useful in the analysis of RC6 as the differentials we have already described.

7.5 Multiple-bit differential cryptanalysis

Here we extend the results from Section 5.5 and briefly discuss how to use some non-static characteristics in the analysis of RC6. Again, we use integer subtraction as our measure of difference.

We follow the notation in Section 5.5 and let $e_{s,t} = 2^s + 2^t$. Based on Lemma 9, we obtain the following two-bit characteristics for one round of RC6 which both hold with probability $r \times m_{t,t+5}$. Recall that we defined $m_{i,j}$ to be the probability of the characteristic $\delta_x \rightarrow \delta_y$ where $\delta_x = e_{i,j}$ and $\delta_y = 2^i$ for some $i \neq j$.

$$\begin{array}{cccc} 0 & 0 & e_t & e_{t+5,t} \\ & & \downarrow & \\ 0 & e_{s+5,s} & e_{t+5,t} & 0 \end{array} \quad \text{and} \quad \begin{array}{cccc} 0 & 0 & e_{t+5,t} & e_{t+5,t} \\ & & \downarrow & \\ 0 & e_s & e_{t+5,t} & 0 \end{array}$$

Even though these characteristics are "heavier" than the one-bit characteristics for one-round RC6, they appear to be useful in constructing iterative characteristics for RC6. The four-round iterative characteristic given in Table 14 is such an example and holds with probability

$$\rho^8 \times q_t^4 \times m_{t,t+5}^4.$$

general					*a specific choice*			
e_{t+5}	e_t	e_{u+5}	e_u		e_{26}	e_{21}	e_{26}	e_{21}
	$\downarrow$					$\downarrow$		
e_t	0	e_u	0		e_{26}	0	e_{26}	0
	$\downarrow$					$\downarrow$		
0	e_s	0	e_v		0	e_{26}	0	e_{26}
	$\downarrow$					$\downarrow$		
e_s	e_{s-5}	e_v	e_{v-5}		e_{26}	e_{21}	e_{26}	e_{21}

Table 13: A generalized iterative three-round characteristic and one particular embodiment for RC6. While involving more rotation units this characteristic and associated differential is nearly, but not quite, as effective as the others we have been considering.

e_t	e_t	e_t	$e_{t+5,t}$
	$\downarrow$		
e_t	$e_{t+5,t}$	$e_{t+5,t}$	$e_{t+5,t}$
	$\downarrow$		
$e_{t+5,t}$	e_t	$e_{t+5,t}$	$e_{t+5,t}$
	$\downarrow$		
e_t	e_t	$e_{t+5,t}$	e_t
	$\downarrow$		
e_t	e_t	e_t	$e_{t+5,t}$

Table 14: A useful four-round multiple-bit iterative characteristic for RC6. While consideration of multiple-bit characteristics allows for shorter iterative differentials, their reduced probability means that they are less useful.

This probability is maximized at $t = 26$, giving a probability of about 2^{-80}. Extending the result to six rounds, we obtain a six-round iterative characteristic that holds with probability of 2^{-120}.

We remark that there is no significant differential effect related to the characteristics given in this section. This is due to constraints on how the multiple bit positions in a difference line up. Hence, the single-bit differentials presented in Section 7.1 have much higher probabilities than the characteristics derived in this section.

While the existence of heavier characteristics and differentials is interesting to note, they seem to be more difficult to find and use than the lighter ones we have already identified. This, however, will remain the object of future work.

Part II
Linear Cryptanalysis

In assessing the security of RC6 with regards to linear cryptanalysis [22, 23] we have described our findings on the topic in two different ways depending on the style of the analysis.

The first style of analysis yields a set of linear cryptanalytic attacks that are suitable for attacking all the simple variants of RC6 that we consider (see Section 1.2) along with RC6 itself. The second style of analysis is less effective than the first, and the different variants of RC6 need to be considered separately. However we have included this analysis for the sake of completeness.

We assume throughout that the reader is already familiar with much of linear cryptanalysis, its application, and its extensions.

8 Overview

In this section we review linear cryptanalysis, some of the advanced techniques that might be useful during our analysis, and an overview of our approach to using linear cryptanalysis in the assessment of RC6.

8.1 Linear cryptanalysis

Linear cryptanalysis is an attack on iterated ciphers that bears more than a passing resemblance to the differential cryptanalytic attacks of Biham and Shamir. Indeed many commentators make use of this duality in providing a parallel and simultaneous analysis of the resistance of a cipher to both differential and linear cryptanalysis. However the structure of RC6 does not lend itself to this kind of analysis and we have considered the two issues separately in this report.

The aim of a linear cryptanalytic attack is to find an effective linear expression connecting some bits of the plaintext, some bits of the output at round r and some key bits. This linear expression is valid over, say, r rounds. Suppose that the probability p that the approximation holds is anything but $1/2$, that is, there is some non-zero bias where the bias ϵ is given by $\epsilon = p - 1/2$. Then by taking sufficiently many plaintext/ciphertext pairs the correct value of the exclusive-or of the key bits can be identified.

The greater the bias, the fewer the number of plaintext/ciphertext pairs that need to be taken before the correct key bit value can be deduced. It is important to note that the data requirements for a linear cryptanalytic attack are inversely proportional to the square of the bias of the approximation [22]. This can be contrasted with the situation for differential cryptanalysis where the data requirements are inversely proportional to the probability of the differential.

Some additional techniques can be used to help in a linear cryptanalytic attack. Analyzing other expressions will yield information about other key bits, and counters can be used to search over a small subset of key bits; this potentially allows the cryptanalyst to predict the action of some round and to use a shorter linear approximation with an improved, exploitable bias [23]. Other techniques [6, 16] have also proved to be useful in certain situations.

8.1.1 Notation and basic assumptions

Suppose that we have a linear approximation that involves four starting words A, B, C, and D, four ending words A', B', C', and D' and a bit of key material k. Then we will use a second set of 32-bit words to indicate which bits of these eight words are to be exclusive-ored together in evaluating the linear approximation.

By way of example, suppose that the linear approximation we are interested in consists of the least significant bits of A, C' and D'. As when considering differential cryptanalysis, we will use e_t to denote 2^t, the 32-bit word with a single one in the t^{th} least significant bit position.

In a slight abuse of notation we will consider a 32-bit word x as a vector in Z_2^{32} and we will use some 32-bit quantity, τ say, to indicate the bits of x that are to be used in a linear approximation. This is most conveniently described by means of the *scalar product* of two vectors. Thus the $\{0,1\}$-vector τ will be used to denote the specific bits of x to be used in an approximation and $x \cdot \tau$ is the value of these bits combined using exclusive-or. An example linear approximation might be written in the following way

$$(A \cdot e_0) \oplus (C' \cdot e_0) \oplus (D' \cdot e_0) = k.$$

Clearly we can use other words than e_i to pick out single, or multiple bits, from a word of text.

Suppose the bias of an approximation A_0 is given by ϵ_0. Then Matsui [22] demonstrates that the amount of plaintext required to exploit this bias with a high success rate is $c \times \epsilon_0^{-2}$ where c is some constant dependent on the style of attack mounted. Typically, the more key material we try to recover using some approximation, the greater the value of c. For the more sophisticated attacks on DES, for instance, in which 13 bits of key material are recovered, experimental evidence suggests that c should be equal to eight. As designers, however, we have taken a pessimistic stance and set $c = 1$.

To compute the bias of a combination of approximations, say A_0 and A_1, we will assume that the so-called *piling-up lemma* [22] can be used. This indicates that the effective bias of the combination of A_0 and A_1 is given by $\epsilon_0 \times \epsilon_1 \times 2$. This can be continued for any number of approximations. There are some interesting technical questions in the cryptographic literature about how widely one can call upon the use of the *piling-up* lemma, particularly when one needs to take into account the issue of key-dependency. But in the absence of this

lemma there appear to be no other tools available to assess the resistance of a cipher to linear cryptanalysis for either the cryptographer or the cryptanalyst.

Finally, throughout our analysis we have taken the line that it is reasonable to assume that in attacking an r-round cipher a linear approximation for $(r-2)$ rounds will typically be required. To give a basic starting point for our analysis this seems to be a reasonable assumption [16, 23]. As analysis of the algorithms in the AES effort continues, there will be tricks that apply to one cipher but not to others which reduce the length of the required approximation. Alternatively, some ciphers will resist attempts to use even an $(r-2)$-round linear approximation, requiring instead one that runs over $(r-1)$ rounds. At this stage we feel that it is not useful in speculating as to the exact nature of future linear cryptanalytic attacks on RC6. We quote the data requirements for mounting a linear cryptanalytic attack on RC6 under the assumption that an $(r-2)$-round linear approximation will be required but we provide enough information for the reader to adapt our estimates as required.

8.1.2 Multiple linear approximations

The use of multiple linear approximations [6, 7] can sometimes enhance a basic linear cryptanalytic attack. The basic idea is to reuse the data one already possesses in a different way. This is done by using a different linear approximation and it is possible to extract more information about the key by using a variety of linear approximations together.

The most important feature of this technique is the bias of the different approximations we might use. The usefulness of multiple approximations can sometimes be limited if the bias of the additional approximations are much less than the bias of one dominating approximation.

Given n approximations A_i with biases ϵ_i for $0 \leq i \leq n-1$, we already have that the amount of plaintext required to successfully use A_0 alone is proportional to ϵ_0^{-2}. Assuming the most advantageous conditions for the cryptanalyst, the amount of plaintext required to successfully use all n approximations A_i for $0 \leq i \leq n-1$ is proportional to $(\sum_{i=0}^{n-1} \epsilon_i^2)^{-1}$. For example by using n linear approximations, which all hold with the same bias, one might obtain a reduction in the plaintext requirements for a linear cryptanalytic attack by a factor of n. More generally, the reduction in plaintext is by a factor of $(\sum_{i=0}^{n-1} \epsilon_i^2)/\epsilon_0^2$.

There is of course some penalty to pay in using multiple approximations, and that is in an increased work load. However this is rarely an issue when considering security at this broad level. It does however become an issue when one is serious about implementing such an attack. There are also technical issues in mounting such an attack which can greatly hinder the use of multiple linear approximations. In this report, however, we have taken the most pessimistic view (as designers!) and assumed that all our attacks and the more esoteric enhancements we consider can be used without any practical difficulty.

8.1.3 Linear hulls

It is sometimes said that just as the notion of differentials generalizes that of the characteristic in differential cryptanalysis, the notion of linear hulls generalizes the idea of a single linear approximation.

As with multiple linear approximations, we consider the idea of many different linear approximation coexisting within a cipher. With multiple linear approximations we assumed that these were different approximations in as much as they used different bits at the start of the linear approximation and at the end. In [27] it is observed that even when considering one linear approximation in isolation, the practical exploitable bias of the approximation might be different to that obtained by analysis of a single path through the cipher. There are also other approximations through the cipher involving the same starting and ending bits, but which take a different set of internal key bits. Consideration of these paths might allow for a more accurate estimate of the number of plaintexts required in an attack.

As with using multiple approximations, in some practical situations the bias of one approximation dominates the bias of the associated linear hull. As a result, consideration of the single most prominent approximation can often be sufficiently accurate in practice. We have taken account of linear hulls in our analysis of RC6 and its variants and we mention in the text when the effect is likely to be significant.

8.2 Linear cryptanalysis and RC6

In considering the range of linear cryptanalytic attacks that can be mounted on RC6, we have drawn a distinction between two different types of approximation.

The first type, which we will call *Type I*, uses one particular type of linear approximation across the data-dependent rotation. These attacks lead to the most efficient linear cryptanalytic attacks on all simplified variants of RC6 and also on RC6 itself. It is consideration of this style of attack that led us to choose 20 rounds as offering adequate security for RC6.

The second type, *Type II*, involve a different style of approximation across the data-dependent rotation. Such approximations in turn lead to approximations across the fixed rotation and the quadratic function. Our analysis presented in this report suggests that the Type II approximations offer less efficient attacks than those dependent solely on Type I approximations. We note that it is possible to consider a hybrid style of attack consisting of both Type I and Type II approximations. Once again, however, it is the Type I approximations that appear to be the most useful. While we have found an intriguing role for both multiple approximations and linear hulls in attacks involving Type I approximations, they seem to be less useful where Type II approximations are concerned.

The advanced technique of using non-linear approximations [16] is unlikely

to be of substantial use in the linear cryptanalytic attacks we consider here. This technique currently offers only slight improvements in attacks and can be viewed as one that might finesse the effectiveness of an existing attack across a notional boundary that separates the barely impractical attacks from those that are barely practical. Currently we have not observed any significant opportunity for the use of non-linear approximations in attacking RC6.

8.2.1 Type I and Type II approximations

Here we draw the distinction between Type I and Type II approximations. This distinction revolves around the way we form linear approximations across the data-dependent rotation.

For $,_a, ,_b, ,_c \in \{0,1\}^{32}$ a *Type I* linear approximation to RC6 uses approximations to the data-dependent rotation $A = B \lll C$ of the form

$$A \cdot ,_a = B \cdot ,_b \oplus C \cdot ,_c.$$

For $,_a, ,_b \in \{0,1\}^{32}$ a *Type II* linear approximation to RC6 uses approximations to the data-dependent rotation $A = B \lll C$ of the form

$$A \cdot ,_a = B \cdot ,_b.$$

Type I Approximations. For these approximations to be useful the forms of $,_a$, $,_b$, and $,_c$ are severely restricted. In particular $,_c$ can only consist of bits from the least significant $\lg w$ bit positions. Otherwise the bias of the approximation is zero. $,_a$ can consist of any subset of all w bits, but to be useful as an approximation, $,_b$ must consist of the same set of w bits as $,_a$ but rotated cyclically (possibly by zero positions). If we assume that $,_a$ can be rotated onto $,_b$ in t different ways, then the bias of the approximation is given by $\rho = \frac{t}{32} + \frac{32-t}{32} \times \frac{1}{2}$.

When we come to use Type I approximations, the structure of RC6 forces the bits of $,_a$ to consist of those in the least significant $\lg w$ bit positions. This ensures that $,_a$ can only be mapped onto $,_b$ in one way ($t = 1$) and so the bias of the approximation $A \cdot ,_a = B \cdot ,_b \oplus C \cdot ,_c$ has bias 2^{-6}. Since there is no drop in the bias across the data-dependent rotation irrespective of the weight of $,_i$, and yet there is a drop in the bias across the integer addition as the weight increases, our analysis will concentrate exclusively on the use of single-bit Type I approximations.

Type II Approximations. For these approximations $,_a$ can consist of any subset of all w bits. To be useful as an approximation, $,_b$ must consist of the

same set of w bits as $, _a$ but rotated cyclically (possibly by zero positions). If we assume that $, _a$ can be rotated onto $, _b$ in t different ways, then the bias of the approximation is given by $\rho = \frac{t}{32} + \frac{32-t}{32} \times \frac{1}{2}$.

It is possible to use heavier masks $, _a$ so that $, _a$ can be rotated on to $, _b$ in several different ways ($t > 1$) ensuring an increased bias. However, studies on RC5 [24, 9] demonstrate that the more bits there are in $, _a$, the harder it is to use the approximation effectively across the integer addition. With RC6 we have another problem since it is difficult to use such multiple-bit approximations across the quadratic function. The best option for the cryptanalyst appears to be to use single-bit approximations in $, _a$ (and therefore in $, _b$ and $, _c$ as well). This then ensures that $, _a$ can only be mapped onto $, _b$ in one way ($t = 1$) and so the bias of the approximation $A \cdot e_a = B \cdot e_b \oplus C \cdot e_c$ has bias 2^{-6}.

To Summarize: For both Type I and Type II approximations we will only consider single-bit approximations. Multiple-bit approximations appear unlikely to provide any advantage over the single-bit case. While a hybrid attack using both Type I and Type II approximations is certainly conceivable, we will see that attacks using Type I approximations exclusively seem to be the most effective.

8.2.2 Some basic tools

As well as the data-dependent and fixed rotations, the other basic operations used in RC6 are the addition of key material using integer addition, the transformation produced by the quadratic function, and the bitwise exclusive-or.

Bitwise exclusive-or need not concern us with regards to linear cryptanalysis. With regards to the use of integer addition, we have the following lemma that describes the bias of a single-bit linear approximation across this operation.

Lemma 10 *Given $y = x + a$ for some fixed 32-bit word a and variable input x, let α_i denote the bias of the linear approximation $y \cdot e_i = x \cdot e_i$ for $0 \leq i \leq 31$ averaged over all possible values of a. Then*

$$\alpha_i = \begin{cases} 1/2 & \text{if } i = 0, \\ 1/4 & \text{otherwise.} \end{cases}$$

We won't need the following result on single-bit approximations across the quadratic function $f(x) = x(2x + 1)$ until we look at Type II approximations, but we give a description of the biases of these approximations here.

Lemma 11 *Let $y = x(2x + 1)$. Then*

$$\text{Probability } (x \cdot e_i = y \cdot e_j) = \begin{cases} 1/2 & \text{if } i \neq j, \\ 1/2 + b_i & \text{otherwise} \end{cases}$$

where the values of b_i for $0 \leq i \leq 31$ are given below:

bit i	0	1	2	3
bias b_i	0.5	0	0.25	0.125
bit i	4	5	6	7
bias b_i	0.0625	0.03125	0.046875	−0.07812
bit i	8	9	10	11
bias b_i	0.011719	0.005859	0.012695	0.004395
bit i	12	13	14	15
bias b_i	0.004639	−0.002563	0.003479	0.002106
bit	16	17	18	19
bias b_i	0.000351	−0.000008	0.000637	0.000025
bit	20	21	22	23
bias b_i	0.000365	0.000157	0.000242	0.000108
bit	24	25	26	27
bias b_i	0.000081	−0.000032	0.000070	0.000021
bit	28	29	30	31
bias b_i	0.000003	0.000001	0.000011	0.000001

Proof. Suppose $y = x(2x+1)$. Partition all possible inputs x into two sets S_0 and S_1 according to the value of $x \cdot e_j$. Consider an input $x \in S_0$ and suppose that $y \cdot e_j = x \cdot e_i$ for $i \neq j$. Now consider $x' = x + 2^j$ which is clearly in S_1. Let $y' = x' \times (2x'+1)$ so $y' = (x+2^j) \times (2x+2^{j+1}+1) = 2x^2 + 2^{j+2}x + x + 2^{2j+1} + 2^j$. We can rewrite this as $y' = y + 2^{j+2}x + 2^j + 2^{2j+1}$ and we see that $y' \cdot e_j \neq y \cdot e_j$. Thus, if $y \cdot e_j = x \cdot e_i$ for some input $x \in S_0$ then there is some input $x \in S_1$ for which $y \cdot e_j \neq x \cdot e_i$ and *vice versa*. That is, over all inputs $y \cdot e_j = x \cdot e_i$ with probability $1/2$. The values for the bias b_i when $i = j$ were derived experimentally. □

Intuitively we might expect a cryptanalyst to attempt to use the approximation involving bit 0 across the function $f(x)$ as much as possible in an attack. After all, this has the biggest bias. However the fixed rotation has an important role and instead of considering approximations to the function $f(x)$ we will need to consider approximations to the function $y = g(x) = (f(x) \lll 5)$. Immediately we see that the best linear approximation across the combined unit is provided by $y[5] = x[0]$ which holds with probability one. Lemma 11 is also useful since it shows that it is not possible to try and use a linear approximation across the quadratic function in such a way so as to "anticipate" the action of the fixed rotation and thereby to cancel it out.

8.2.3 Linear cryptanalysis of RC5

RC5 appears to be remarkably resistant to linear cryptanalysis. The most comprehensive study of how an attack might be carried out is due to Kaliski and Yin [8]. There it is demonstrated that after 12 half rounds of the 32-bit version of RC5, the amount of data that might be required to mount an attack exceeds the amount of data that is available.

More recent work by Selcuk [31] has made a more accurate assessment of the finer points of the attacks outlined by Kaliski and Yin. While observing that the more involved attacks do not work quite as claimed, Selcuk remarks that the basis of any linear cryptanalytic attack will still likely be the linear approximation outlined in [8]. As a result, linear cryptanalysis of even a moderate number of rounds of RC5 remains an extremely remote threat.

Much, if not all, of the resistance of RC5 to linear cryptanalysis stems from the data-dependent rotation. We would expect many of the same problems that are encountered in trying to launch an attack on RC5 to also feature in attempts to attack RC6.

Much of the analysis on RC5 has concentrated on using linear approximation with very few bits involved from each word of the text. Analysis there [9] suggests that the use of heavier linear approximations in a linear cryptanalytic attack is unlikely to be as fruitful as the single-bit case. While the bias of an approximation across the rotation unit can be increased by using more bits in the linear approximation, this is countered by a drop in the bias across the integer addition. For sufficiently large word size ($w > 8$) this more than makes up for any gain across the data-dependent rotation. A similar phenomenon can be expected with RC6, except if anything, it will be more pronounced since the heavier approximation also has to cross the quadratic function and the fixed rotation.

Obviously, as more work is completed on the linear cryptanalysis of RC6 we will have an increasingly accurate assessment of the security of the cipher. This will, undoubtedly, include more analysis of the use of multi-bit linear approximations.

9 Using Type I Approximations

In this section:

- *We show that a simple two-round iterative linear approximations can be used with RC6 and the other variants considered in this report*
- *We consider the possible effects of multiple approximations and linear hulls*
- *We present what are currently the most effective attacks on RC6*
- *We use these considerations to set the number of rounds for RC6*

It is straightforward to mount an attack using Type I approximations that applies to all of the simplified variants of RC6 given in this report and to RC6 itself. The basis for this analysis is the following two-round linear approximation

for RC6

$$(A \cdot e_t) \oplus (C \cdot e_s) = (A'' \cdot e_u) \oplus (C'' \cdot e_v).$$

Here A and C are the first and third words of some intermediate data, A'' and C'' are the first and third words of the intermediate data after a further two rounds of encryption using RC6 (or its simplified variants). This might be represented graphically as follows, with a general approximation given on the left, and a specific choice for the variables t, s, u and v given on the right.

general					*a specific choice*			
e_t	$-$	e_s	$-$		e_0	$-$	e_0	$-$
	$\downarrow$					$\downarrow$		
$-$	e_u	$-$	e_v		$-$	e_0	$-$	e_0
	$\downarrow$					$\downarrow$		
e_u	$-$	e_v	$-$		e_0	$-$	e_0	$-$

When averaged over all keys, the piling-up lemma suggests that this approximation would hold with bias

$$\alpha_u \times \rho \times \alpha_v \times \rho \times 2^3$$

where α_u and α_v are the biases of the single-bit linear approximations across the integer addition in positions u and v, and ρ is the contributing bias of the single-bit linear approximation across the data-dependent rotation. From Lemma 10 this evaluates to 2^{-11} when $u = v = 0$, 2^{-12} if one of u or v is equal to zero and the other is non-zero but less than five, and bias 2^{-13} when $1 \leq u, v \leq 4$.

If we were to consider the case of $t = s = u = v = 0$ alone, this implies that there is a six-round linear approximation[4] to RC6 (and RC6-I, RC6-NFR, RC6-I-NFR) which holds with bias $2^{-33} \times 2^2 = 2^{-31}$. Generalizing this to run over r rounds we can estimate the bias of this basic linear approximation as

$$(2^{-11})^{\lfloor \frac{r}{2} \rfloor} \times 2^{\lfloor \frac{r}{2} \rfloor - 1}$$

which evaluates to 2^{-101} for $r = 20$.

Next we consider some potential enhancements to this basic analysis and we will assume that there are little, or no, practical difficulties in using these enhancements. It is clear to those that have implemented such attacks, however, that there are often very involved practical and technical difficulties to be overcome when using these techniques. It is therefore very unlikely that enhancements of the type we are going to describe now could be used so readily and so effectively in practice. But as we have said before, as designers we are taking a pessimistic stance on such things.

First we note that we can generate $5^2 \times 32^2$ different approximations as we look over the 5^2 possible starting values for t and s at the start of the cipher

[4] In fact this approximation can be extended another round backwards at no extra cost hence the term $\lfloor \frac{r}{2} \rfloor$ in many of the expressions we derive.

and the possible 32^2 different sets of values to the output bits at the end of the linear approximation. Each of these possibilities defines a different linear approximation and we observe that this freedom in the values of the external variables might potentially allow the use of multiple approximations.

Taking account of the potential effect of multiple linear approximations is straightforward. The two input bits s and t can take any of 5^2 values without affecting the bias of the associated linear approximation. The two output bits can take one set of values (i.e. that corresponding to the least significant bit) which gives the full effect of the bias, 31×2 sets of values that give a bias dropping by a factor of two, and 31^2 sets of values that give a bias dropping by a factor of four. An estimate for the factor change in the plaintext requirements using multiple approximations will be given by

$$25 + (25 \times 62 \times 2^{-2}) + (25 \times 961 \times 2^{-4}) \approx 2^{11}.$$

Note that this is independent of the number of rounds.

We also note that at each stage the two-round approximation repeats (i.e. after every two rounds) the value of u and v can each take one of five values. These values to u and v remain internal to the cipher. As a consequence, for the six-round approximation there will be $5^4 \approx 2^9$ different approximations through the cipher with the same starting and ending bits involved. The biases of these different approximations will vary somewhat. Following [27], one basic and pessimistic estimate for the amount of plaintext needed for a linear cryptanalytic attack can be derived as $c \times 2^{18\lfloor \frac{r}{2} \rfloor + 4}$ plaintexts instead of the $c \times 2^{20\lfloor \frac{r}{2} \rfloor + 2}$ known plaintexts predicted by considering the best linear approximation in isolation.

In Table 15 we summarize these different considerations and make a conservative estimate for the amount of plaintext needed to attack a version of RC6 with 8, 12, 16, 20, and 24 rounds. While we have restricted our attention to single-bit Type I approximations (as motivated by Section 8.2.1), it is worth observing that there could be some very limited opportunity to use heavier approximations. While on their own they will not be effective as the single-bit approximations , it is possible that situations arise where multiple-bit Type I approximations provide an additional contribution to the more advanced techniques. However we believe that the biases of these alternative paths will be somewhat less than the ones we have been considering, and that the number of alternatives will not be that great. Therefore any effect they have on the estimates given in Table 15 will be very slight.

While the use of the advanced linear techniques in this section is somewhat speculative in terms of practicality, we believe that to build a sufficient margin of safety in RC6 we need to take them into account. While RC6 with 16 rounds offers very good security against differential cryptanalysis, the potential effects of multiple approximations and linear hulls might bring the number of known plaintexts for linear cryptanalysis under the threshold of 2^{128} known plaintexts. With this in mind, we chose 20 rounds as being a suitable number of rounds

Using Type I Approximations					
	number of rounds				
variant	8	12	16	20	24
RC6 *basic linear attack*	2^{62}	2^{102}	2^{142}	2^{182}	2^{222}
RC6 *+ multiple approximations*	2^{51}	2^{91}	2^{131}	2^{171}	2^{211}
RC6 *+ multiple approximations + linear hulls*	2^{47}	2^{83}	2^{119}	2^{155}	2^{191}

Table 15: An estimate of the number of plaintext needed to mount a linear cryptanalytic attack on RC6 with a varying number of rounds. Additional consideration of *multiple approximations* and potential improvements using *linear hulls* helped to set a suitable number of rounds for RC6. The attacks referred to in this table apply equally to the simplified variants we have considered in this document, namely RC6-I-NFR, RC6-NFR, and RC6-I.

for RC6, thereby offering good security without unnecessarily compromising performance.

10 Using Type II Approximations

In this section:

- *We show that RC6 displays good resistance to linear cryptanalysis using Type II approximations*
- *We demonstrate that the combination of the quadratic function and the fixed rotation greatly hinders the construction of effective approximations*
- *We observe that RC6-I-NFR appears to be equally vulnerable to Type II approximations as it was to Type I approximations*
- *We note that the other variants and RC6 itself are much less vulnerable to Type II approximations than they are to Type I approximations*

Recall from Section 8.2.1 that we drew a distinction between the way we make our linear approximations across the data-dependent rotation. In this section we will be interested in linear approximations that use what we called Type II approximations.

Despite the interesting results we obtain in this section our efforts are somewhat unproductive. By the close we will be concluding that the additional attacks we uncover are not as useful in mounting a linear cryptanalytic attack on RC6 as were the attacks in Section 9.

10.1 Linear cryptanalysis of RC6-I-NFR and RC6-NFR

When using Type II approximations it can be illustrative to consider a breakdown of the cycles found when tracing a linear approximation through RC6-I-NFR. We did something similar in Section 3 for the case of differential cryptanalysis. These cycles are presented in Table 16. We will denote the unspecified bits involved in the approximation by , and we will use α to denote the bias of a linear approximation across integer addition. We let ρ denote the bias in the linear approximation across the data-dependent rotation.

This is clearly a very simplified form of analysis. It assumes a great deal about the approximations we might actually wish to use. However we have found it useful in illustrating on a network-level how linear approximations might evolve from round to round. It seems that some derivative of either cycle (a) or (b) in Table 16 will be most useful to the cryptanalyst. The third cycle introduces more non-trivial approximations when extended over the same number of rounds than the other two.

If we assume that , $= e_0$ so that the linear approximation involves the least significant bit of a 32-bit word, we can immediately estimate the security of RC6-I-NFR with regards to a basic linear cryptanalytic attack. The linear approximation across the operations of exclusive-or and integer addition in the least significant bit always holds with probability one ($\alpha = 2^{-1}$). So there is a linear approximation to six rounds of RC6-I-NFR that holds with bias $(2^{-6})^6 \times 2^5 = 2^{-31}$. The approximation continues in the obvious way to other numbers of rounds and the approximations are presented graphically in Table 17.

Next we turn our attention to the simplified variant RC6-NFR. It is interesting to note that for the function $f(x) = x(2x + 1)$ there is a linear approximation that holds with probability one. Given $y = x(2x + 1)$ we always have that $y \cdot e_0 = x \cdot e_0$. Consequently we can consider exactly the same cycle breakdown as for RC6-I-NFR and there immediately follows a six-round, iterative linear approximation that can be used to attack RC6-NFR. The different approximations are presented in a generalized form in Table 17. We combine the biases of these approximations to give a linear approximation over six rounds of RC6-NFR and RC6-I-NFR with a bias of around $2^{-6\times 6} \times 2^5 = 2^{-31}$.

We present the results of our analysis in Table 18. There we give the data requirements to attack these simple variants. We have not considered the issue

(a)	(b)	(c)
, , − − ↓ − − − , ↓ − − , − ↓ − , , − ↓ , , , − ↓ − , , , ↓ , , − −	− − , , ↓ − , − − ↓ , − − − ↓ , − − , ↓ , − , , ↓ , , − , ↓ − − , ,	, , , , ↓ − , − , ↓ , − , − ↓ , , , ,
$\alpha^6 \times \rho^6 \times 2^{11}$	$\alpha^6 \times \rho^6 \times 2^{11}$	$\alpha^4 \times \rho^4 \times 2^7$

Table 16: Basic approximations for attacking RC6-I-NFR which illustrate the contribution of the addition unit (contributing bias α) and the rotation unit (contributing bias ρ). Here , denotes a "generic" set of bits for the approximation with specific choices being made to maximize the resultant bias.

general	*a specific choice*
e_t e_t − − ↓ − − − e_s ↓ − − e_s − ↓ − e_u e_s − ↓ e_u e_u e_s − ↓ − e_v e_s e_s ↓ e_v e_w − −	e_0 e_0 − − ↓ − − − e_0 ↓ − − e_0 − ↓ − e_0 e_0 − ↓ e_0 e_0 e_0 − ↓ − e_0 e_0 e_0 ↓ e_0 e_0 − −

Table 17: Generalized six-round, iterative linear approximation for RC6-I-NFR and RC6-NFR. A specific example useful in attacking eight-round RC6-NFR is presented on the right.

Using Type II Approximations					
	number of rounds				
variant	8	12	16	20	24
RC6-I-NFR *basic linear attack*	2^{62}	2^{92}	2^{132}	2^{182}	2^{212}
RC6-NFR *basic linear attack*	2^{62}	2^{92}	2^{132}	2^{182}	2^{212}

Table 18: An estimate of the number of plaintext needed to mount a linear cryptanalytic attack on two simple variants of RC6 with a varying number of rounds. The effect of advanced phenomena has not been presented, though their effectiveness is readily quantified. For RC6-I-NFR, plaintext requirements using enhanced techniques such as those obtained in Table 15 are likely to result. For RC6-NFR, their impact and effect is expected to be much reduced since the quadratic function has a substantial (detrimental) impact on the resultant bias.

of more advanced techniques for these variants since we are only considering them for illustrative purposes. We observe that for the variant RC6-I-NFR we would expect to see plaintext savings similar to those we observed with the Type I approximations, though perhaps not quite so pronounced. The reason for this is that there is a great deal of freedom in both the internal and external variables and there are likely to be a variety of alternative approximation possibilities. We expect the effect to be a little less pronounced because for particular choices of the internal variables, the approximations become somewhat heavier than was the case when using Type I approximations. These heavier alternatives have a reduced bias, and hence contribute less.

The situation with RC6-NFR is more interesting and gives us some valuable clues as to the behavior of Type II approximations. From Lemma 11 in Section 8.2.2 we saw that the bias of the single-bit approximations fell when considering any other bit but the least significant in approximations of $f(x) = x(2x+1)$. This is likely to ensure that any alternative approximations are likely to have a very slight effect. In addition, there appear to be very few opportunities for substantial plaintext gains in using multiple approximations for very much the same reasons. We expect the consideration of more enhanced linear techniques for RC6-NFR to have very little beneficial effect.

Clearly our analysis of these simplified variants is merely an indication of how we think they will perform in practice. More analysis, given more time, is likely to offer refinements. However we have illustrated that even though the

function $f(x) = x(2x + 1)$ allows very effective single linear approximations, it does seem to be very beneficial in thwarting more advanced techniques of analysis.

To Summarize: Type II approximations appear to offer some improvements over Type I approximations in mounting a basic attack on RC6-NFR and RC6-I-NFR. When we consider more advanced techniques in an attack on RC6-I-NFR, we expect the two types of approximations to offer similar benefit. However in attacking RC6-NFR we expect that the enhanced attacks using Type I approximations will be much more effective than the enhanced attacks using Type II approximations.

10.2 Linear cryptanalysis of RC6-I and RC6

As we move to consider RC6-I and RC6 we need to include consideration of the fixed rotation. We have already seen (Section 10.1) that there are good linear approximations across the function $f(x) = x(2x + 1)$. This meant that good basic linear approximations were not prevented by the use of the quadratic function. However, the effect of advanced linear cryptanalytic techniques for Type II approximations became severely restricted. In this section we will show that the fixed rotation helps to ensure that the attacker is unable to use good approximations at each round of a linear approximation. It is only in combining the linear approximations from one round to the next that a proper picture of the resistance of the cipher can be built up. In using Type II approximations we will essentially need to consider approximations of the form $y[5] = x[0]$ where $y = f(x) \lll 5$. This means that even in the most advantageous cases, we are unable to keep our attention focused on the approximations in the least significant bit of some word. This results in greatly reducing the effectiveness of even the basic linear approximations.

In Table 19 we present a good six-round linear approximation suitable for attacking eight rounds of RC6-I and RC6. We present both a general form, and also a specific choice where we choose values to some of the variables to provide what appears to be the optimum single choice for the attacker.

Recall from Lemma 11 the following notation. Let $y = x(2x+1)$. Then the probability that $(x \cdot e_i = y \cdot e_i)$ is equal to $1/2 + b_i$ where the values of b_i are given in Lemma 11. Then the bias for the general linear approximation given in Table 19 can be estimated by

$$\begin{aligned} \rho^6 \quad &\times \quad (\alpha_t \times \alpha_u \times \alpha_{u+5} \times \alpha_{v+5} \times \alpha_w \times \alpha_x) \\ &\times \quad (b_s \times b_t^2 \times b_u \times b_v \times b_{v-5}) \times 2^{17}. \end{aligned}$$

Considering Lemma 10 and Lemma 11 this is maximized for RC6 when $v = 5$ and $s = t = u = w = x = 0$ giving a bias of $2^{-36} \times 2^{-8} \times 2^{-10} \times 2^{17} = 2^{-37}$.

It is straightforward to extend these results to RC6-I reduced to eight rounds. Recalling that $\alpha_0 = 1/2$ and $\alpha_5 = 1/4$ and that for the identity function we

general					*a specific choice*			
e_{s+5}	e_s	e_{t+5}	$e_{t+5} \oplus e_t$		e_5	e_0	e_5	$e_5 \oplus e_0$
		$\downarrow$					$\downarrow$	
$-$	e_{u+5}	e_{t+5}	e_t		$-$	e_5	e_5	e_0
		$\downarrow$					$\downarrow$	
e_{u+5}	e_u	$-$	$-$		e_5	e_0	$-$	$-$
		$\downarrow$					$\downarrow$	
$-$	$-$	$-$	e_{v+5}		$-$	$-$	$-$	e_{10}
		$\downarrow$					$\downarrow$	
$-$	$-$	e_{v+5}	$-$		$-$	$-$	e_{10}	$-$
		$\downarrow$					$\downarrow$	
$-$	e_w	e_v	$-$		$-$	e_0	e_5	$-$
		$\downarrow$					$\downarrow$	
e_w	e_x	e_{v-5}	$-$		e_0	e_0	e_0	$-$

Table 19: A useful non-iterative six-round linear approximation for attacking eight-round RC6 and RC6-I. The approximation on the left is in a general form, illustrating the flexibility provided by the data-dependent rotation, while the approximation on the right is one that maximizes the bias.

general					*a specific choice*			
e_{s_1+5}	e_{s_1}	$-$	$-$		e_5	e_0	$-$	$-$
	$\downarrow$					$\downarrow$		
$-$	$-$	$-$	e_{t+5}		$-$	$-$	$-$	e_{20}
	$\downarrow$					$\downarrow$		
$-$	$-$	e_{t+5}	$-$		$-$	$-$	e_{20}	$-$
	$\downarrow$					$\downarrow$		
$-$	e_{u+5}	e_t	$-$		$-$	e_5	e_{15}	$-$
	$\downarrow$					$\downarrow$		
e_{u+5}	e_u	e_{t-5}	$-$		e_5	e_0	e_{10}	$-$
	$\downarrow$					$\downarrow$		
$-$	e_{s_2+5}	e_{t-10}	e_{t-15}		$-$	e_5	e_5	e_0
	$\downarrow$					$\downarrow$		
e_{s_2+5}	e_{s_2}	$-$	$-$		e_5	e_0	$-$	$-$

Table 20: A useful iterative six-round linear approximation for attacking eight-round RC6. The approximation on the left is in a general form, illustrating the possible effect of linear hulls and multiple linear approximations, while the approximation on the right is one that maximizes the bias.

general					*a specific choice*			
e_{s_1+5}	e_{s_1}	$-$	$-$		e_5	e_0	$-$	$-$
		$\downarrow$					$\downarrow$	
$-$	$-$	$-$	e_{t_1+5}		$-$	$-$	$-$	e_{20}
		$\downarrow$					$\downarrow$	
$-$	$-$	e_{t_1+5}	$-$		$-$	$-$	e_{20}	$-$
		$\downarrow$					$\downarrow$	
$-$	e_{u_1+5}	e_{t_1}	$-$		$-$	e_5	e_{15}	$-$
		$\downarrow$					$\downarrow$	
e_{u_1+5}	e_{u_1}	e_{t_1-5}	$-$		e_5	e_0	e_{10}	$-$
		$\downarrow$					$\downarrow$	
$-$	e_{s_2+5}	e_{t_1-10}	e_{t_1-15}		$-$	e_5	e_5	e_0
		$\downarrow$					$\downarrow$	
e_{s_2+5}	e_{s_2}	$-$	$-$		e_5	e_0	$-$	$-$
		$\downarrow$					$\downarrow$	
$-$	$-$	$-$	e_{t_2+5}		$-$	$-$	$-$	e_{20}
		$\downarrow$					$\downarrow$	
$-$	$-$	e_{t_2+5}	$-$		$-$	$-$	e_{20}	$-$
		$\downarrow$					$\downarrow$	
$-$	e_{u_2+5}	e_{t_2}	$-$		$-$	e_5	e_{15}	$-$
		$\downarrow$					$\downarrow$	
e_{u_2+5}	e_{u_2}	e_{t_2-5}	$-$		e_5	e_0	e_{10}	$-$
		$\downarrow$					$\downarrow$	
$-$	e_{s_3+5}	e_{t_2-10}	e_{t_2-15}		$-$	e_5	e_5	e_0
		$\downarrow$					$\downarrow$	
e_{s_3+5}	e_{s_3}	$-$	$-$		e_5	e_0	$-$	$-$
		$\downarrow$					$\downarrow$	
$-$	$-$	$-$	e_{t_3+5}		$-$	$-$	$-$	e_0
		$\downarrow$					$\downarrow$	
$-$	$-$	e_{t_3+5}	$-$		$-$	$-$	e_0	$-$

Table 21: A useful 14-round single-bit linear approximation for attacking 16-round RC6 and RC6-I. The approximation on the left is in a general form while the approximation on the right is one that maximizes the bias of the single approximation.

have $b_i = 1/2$ for $0 \leq i \leq 31$ we can make the appropriate substitutions into Equation 8. This yields an effective bias for the basic linear approximation of around 2^{-32}.

In trying to attack a greater number of rounds the six-round characteristic in Table 19 isn't that useful. We immediately see that it is non-iterative, i.e. it cannot be concatenated with itself to cover more rounds. Furthermore, we see that additional bits are being introduced, and the attack will begin to involve multiple bits from the same word. This reduces the effective bias and makes it difficult to connect the one-round approximations together.

Instead, the cryptanalyst might prefer to use a single-bit six-round iterative linear characteristic. A general form of this characteristic, along with a specific choice that appears to optimize the bias for the attacker, is given in Table 20. While there might be other examples of such iterative characteristics, the approximation given there appears to be among the most useful. An estimate for the bias for the general approximation in Table 20 is

$$\begin{aligned} \rho^6 \quad &\times \quad (\alpha_{t+5} \times \alpha_u \times \alpha_{u+5} \times \alpha_{t-15} \times \alpha_{s_2+5} \times \alpha_{s_2}) \\ &\times \quad (b_{s_1} \times b_t \times b_{t-5} \times b_{t-10} \times b_{t-15} \times b_u) \times 2^{17}. \end{aligned}$$

This is maximized for RC6 when $s_1 = u = 0$ and $t = 15$ (s_2 can take any value) giving a bias of $2^{-36} \times 2^{-9} \times 2^{-23.2} \times 2^{17} \approx 2^{-51}$. The bias of the corresponding six-round iterative linear approximation for RC6-I is 2^{-33}.

We can iterate these six-round linear approximations to get ones covering 12 rounds of RC6 or RC6-I respectively, adding another two rounds to get a 14-round linear approximation. The bias of the 14-round linear approximation for RC6 obtained in this way and presented in Table 21 can be estimated as

$$\begin{aligned} \rho^{12} \quad &\times \quad (\alpha_{t_1+5} \times \alpha_{u_1} \times \alpha_{u_1+5} \times \alpha_{t_1-15} \times \alpha_{s_2+5} \times \alpha_{s_2}) \\ &\times \quad (\alpha_{t_2+5} \times \alpha_{u_2} \times \alpha_{u_2+5} \times \alpha_{t_2-15} \times \alpha_{s_3+5} \times \alpha_{s_3}) \\ &\times \quad (b_{s_1} \times b_{t_1} \times b_{t_1-5} \times b_{t_1-10} \times b_{t_1-15} \times b_{u_1}) \\ &\times \quad (b_{s_2} \times b_{t_2} \times b_{t_2-5} \times b_{t_2-10} \times b_{t_2-15} \times b_{u_2}) \\ &\times \quad (r \times b_{s_3} \times \alpha_{t_3+5}) \times 2^{38}. \end{aligned}$$

Set $s_1 = s_2 = u_1 = u_2 = 0$. Following our earlier work on the eight round version of this iterative linear approximation, set $t_1 = t_2 = 15$ and for the extension of the additional two rounds set $s_3 = 0$ and $t_3 = 27$. This yields a linear approximation over 14 rounds of RC6 with a bias given by

$$\rho^{13} \times (b_{t_1} \times b_{t_1-5} \times b_{t_1-10}) \times (b_{t_2} \times b_{t_2-5} \times b_{t_2-10}) \times 2^{12} \approx 2^{-106}.$$

Just as we added another two rounds as needed to the iterative linear approximation, we can also add the first four rounds of the iterative six-round approximation. This simple four-round extension has bias 2^{-21} and for RC6-I the effective bias would be around 2^{-16}. The results of our work on Type II approximations are given in Table 22.

Using Type II Approximations					
	number of rounds				
variant	8	12	16	20	24
RC6 *basic linear attack*	2^{74}	2^{142}	2^{212}	2^{302}	2^{342}
RC6-I *basic linear attack*	2^{66}	2^{96}	2^{140}	2^{192}	2^{224}

Table 22: An estimate of the number of plaintext needed to mount a linear cryptanalytic attack on RC6 and a simplified variant RC6-I using Type II approximations. See the text for the additional consideration of customized linear approximations. Due to the effect of the quadratic function, more advanced linear cryptanalytic techniques are unlikely to offer much reduction to these figures. Attacks based solely on Type I approximations (Table 15) are almost certainly going to be of more benefit to the cryptanalyst.

The careful reader will have observed that we haven't made the same kind of allowances for customization at the beginning and end of a long approximation as we did in the case of differential cryptanalysis (Section 7.2). This is primarily because even with such customizations, our work on Type I approximations still gives us the best attacks on RC6. However, to get some idea of the limits even of customization, we might assume that when the six-round iterative characteristic is used first, we can replace it with the six-round approximation for RC6-I which holds with bias 2^{-33}. Similarly, we might assume that we can replace the last invocation of the six-round linear approximation with the same approximation. Even with such assumptions on the level of customization, which ignore the presence of the quadratic function, the estimated data requirements to attack RC6 with 16, 20 and 24 rounds are 2^{176}, 2^{230}, and 2^{304} known plaintexts respectively. These still exceed the data requirements in Table 15 by a substantial margin.

A simple comparison with the results in Table 15 shows that for the basic linear approximations, the use of Type II approximations is not as good as the use of Type I approximations. When we include the potential advantages of advanced techniques in our considerations, the substantial gains for Type I approximations and the meagre gains expected for Type II approximations, the distinction between the two forms of analysis becomes even more pronounced.

Part III
The Key Schedule

11 Description of the Key Schedule

The key schedule of RC6-$w/r/b$ is practically identical to the key schedule of RC5-$w/r/b$ the only difference being the number of words ($2r + 4$ instead of $2r + 2$) derived from the user-supplied key for encryption and decryption. The user supplies a key of b bytes. Extra zero bytes are appended to the key if necessary to make the length of the key a non-zero multiple of four bytes. This is then stored as a sequence of c w-bit words $L[0], \ldots, L[c-1]$, with first byte of key stored as low-order byte of $L[0]$, etc., and $L[c-1]$ padded with high-order zero bytes if necessary. (Note that if $b = 0$ then $c = 1$ and $L[0] = 0$.) The additive round keys are stored in the array $S[0, ..., 2r + 3]$.

The constants $P_{32} =$ `B7E15163` and $Q_{32} =$ `9E3779B9` (hexadecimal) are the same "magic constants" as used in the RC5 key schedule. The value of P_{32} is derived from the binary expansion of $e - 2$, where e is the base of the natural logarithm function. The value of Q_{32} is derived from the binary expansion of $\phi - 1$, where ϕ is the Golden Ratio.

12 Security of the Key Schedule

Since the key schedule of RC6 is identical to that of RC5 we can make claims for the security of the RC6 key schedule based on the results of over three years of scrutiny of the key schedule for RC5 by the research community.

12.1 Weak keys

The term *weak keys* is used widely but it is not always applied in the same way. The most famous examples of weak keys are perhaps those for DES [25] which exploit a structural property of the cipher. In the case of DES their existence has only a very limited security implication.

Other ciphers with classes of weak keys are IDEA [19, 3] and Blowfish [30, 32] where the weakness might allow cryptanalysis of the cipher in some very limited, and typically unlikely, situations.

Since the publication of RC5 there have been no reported examples of weak keys. This includes ones demonstrating a structural weakness or ones allowing a limited form of cryptanalytic attack. Since the key schedule for RC6 is in all significant ways identical to that of RC5 it might be expected that the same lack of weak keys will also apply here.

12.2 Related-key attacks

A class of attacks that has gained some attention in the community is that of related-key attacks [10]. The essential premise of the attack is that keys related in some known way are used during encryption, and by observing the relations between the plaintexts and ciphertexts, it might be possible to deduce information about the two unknown keys.

Again we might argue that since the RC5 key schedule has been available for scrutiny with regards to these attacks for several years, and that no such attacks have been reported, then we would not expect such attacks to be applicable to RC6. In addition however, we might argue than any development in this direction is highly unlikely since the key schedule is quite complicated and more importantly has a design that might be viewed as being somewhat incompatible with the structure of the encryption process. It is hard to imagine a way in which changes in the used-defined key can be readily translated into known and useful changes in the subkeys that are used during encryption. Of course continued research will establish whether or not this is indeed the case.

Key schedule for RC6-$w/r/b$

Input: User-supplied b byte key preloaded into the c-word array $L[0, \ldots, c-1]$
Number r of rounds

Output: w-bit round keys $S[0, \ldots, 2r+3]$

Procedure:

$S[0] = P_w$

for $i = 1$ **to** $2r+3$ **do**
    $S[i] = S[i-1] + Q_w$

$A = B = i = j = 0$

$v = 3 \times \max\{c, 2r+4\}$
for $s = 1$ **to** v **do**
    {
        $A = S[i] = (S[i] + A + B) \lll 3$
        $B = L[j] = (L[j] + A + B) \lll (A + B)$
        $i = (i+1) \bmod (2r+4)$
        $j = (j+1) \bmod c$
    }

Part IV
Other Attacks

13 Differential-Linear Cryptanalysis

Differential-linear cryptanalysis was introduced by Langford and Hellman at Crypto'94 [20]. This very elegant attack uses a differential to predict the difference between two texts part way through the encryption. From knowledge of this difference, it is possible to use a linear approximation starting at this later stage during the encryption. Sometimes, with a sufficiently good linear approximation, more of the cipher can be covered than could be achieved by a good differential alone. The hope is that this leads to an attack over a greater number of rounds with a reduced data cost.

While providing the best existing attack on an eight-round version of DES [20], this style of attack is not very widely applicable. It depends on the existence of a good differential for the start of the attack, and the existence of a good linear approximation for the later stages of the attack.

Our work on the differential and linear cryptanalysis of RC6 has demonstrated that both of these requirements are exceptionally unlikely to be fulfilled, particularly if the aim is to mount a differential-linear attack that threatens the full 20 rounds of RC6.

Acknowledgments

We would particularly like to thank Burt Kaliski, Ray Sidney, and Jeff Ylvisaker for many useful suggestions during the design and analysis of RC6. We would also like to thank Kaisa Nyberg and Ali Selcuk for kindly answering our queries.

References

[1] E. Biham and A. Shamir. *Differential Cryptanalysis of the Data Encryption Standard*. Springer-Verlag, New York, 1993.

[2] A. Biryukov and E. Kushilevitz. Improved cryptanalysis of RC5. In K. Nyberg, editor, *Advances in Cryptology — Eurocrypt '98*, volume 1403 *Lecture Notes in Computer Science*, pages 85–99, 1998. Springer Verlag.

[3] J. Daemen, R. Govaerts and J. Vandewalle. Weak keys for IDEA. In D. Stinson, editor, *Advances in Cryptology — Crypto '93*, volume 773 of *Lecture Notes in Computer Science*, pages 224–231, New York, 1994. Springer Verlag.

[4] M.H. Heys. Linearly weak keys of RC5. *IEE Electronic Letters*, Vol. 33, pages 836–838, 1997.

[5] T. Jakobsen and L.R. Knudsen. The interpolation attacks on block ciphers. In E. Biham, editor, *Fast Software Encryption*, volume 1267 of *Lecture Notes in Computer Science*, pages 28–40, 1997. Springer Verlag.

[6] B.S. Kaliski and M.J.B. Robshaw. Linear cryptanalysis using multiple approximations. In Y.G. Desmedt, editor, *Advances in Cryptology — Crypto '94*, volume 839 of *Lecture Notes in Computer Science*, pages 26–39, New York, 1994. Springer Verlag.

[7] B.S. Kaliski and M.J.B. Robshaw. Linear cryptanalysis using multiple approximations and FEAL. In B. Preneel, editor, *Fast Software Encryption*, volume 1008 of *Lecture Notes in Computer Science*, pages 249–264, 1995. Springer Verlag.

[8] B.S. Kaliski and Y.L. Yin. On differential and linear cryptanalysis of the RC5 encryption algorithm. In D. Coppersmith, editor, *Advances in Cryptology — Crypto '95*, volume 963 of *Lecture Notes in Computer Science*, pages 171–184, 1995. Springer Verlag.

[9] B.S. Kaliski and Y.L. Yin. On the Security of the RC5 Encryption Algorithm. RSA Laboratories Technical Report TR-602. Available at `www.rsa.com/rsalabs/aes/`.

[10] J. Kelsey, B. Schneier and D. Wagner. In N. Koblitz, editor, *Advances in Cryptology — Crypto '96*, volume 1109 of *Lecture Notes in Computer Science*, pages 237–251, New York, 1996. Springer-Verlag.

[11] L. Knudsen. *Block Ciphers - Analysis, Design and Applications.* PhD thesis, Aarhus University, 1994.

[12] L.R. Knudsen. Applications of higher order differentials and partial differentials. In B. Preneel, editor, *Fast Software Encryption*, volume 1008 of *Lecture Notes in Computer Science*, pages 196–211, 1995. Springer Verlag.

[13] L.R. Knudsen and T. Berson. Truncated differentials of SAFER. In D. Gollmann, editor, *Fast Software Encryption*, volume 1039 of *Lecture Notes in Computer Science*, pages 15–25, 1996. Springer Verlag.

[14] L.R. Knudsen and W. Meier. Improved differential attacks on RC5. In N. Koblitz, editor, *Advances in Cryptology — Crypto '96*, volume 1109 of *Lecture Notes in Computer Science*, pages 216–228, 1996. Springer Verlag.

[15] L.R. Knudsen, V. Rijmen, R.L. Rivest and M.J.B. Robshaw. On the design and security of RC2. In S. Vaudenay, editor, *Fast Software Encryption*, volume 1372 of *Lecture Notes in Computer Science*, pages 206–221, 1998. Springer-Verlag.

[16] L.R. Knudsen and M.J.B. Robshaw. Non-linear approximations in linear cryptanalysis. In U. Maurer, editor, *Advances in Cryptology — Eurocrypt '96*, volume 1070 of *Lecture Notes in Computer Science*, pages 224-236, 1996. Springer Verlag.

[17] L.R. Knudsen, M.J.B. Robshaw and D. Wagner. Truncated differentials of Skipjack. In preparation.

[18] X. Lai. *On the Design and Security of Block Ciphers.* PhD thesis, ETH, Zurich, 1992.

[19] X. Lai, J.L. Massey, and S. Murphy. Markov ciphers and differential cryptanalysis. In D.W. Davies, editor, *Advances in Cryptology — Eurocrypt '91*, volume 547 of *Lecture Notes in Computer Science*, pages 17–38, Berlin, 1992. Springer-Verlag.

[20] S.K. Langford and M.E. Hellman. Differential-linear cryptanalysis. In Y.G. Desmedt, editor, *Advances in Cryptology — Crypto '94*, volume 839 of *Lecture Notes in Computer Science*, pages 17–25, 1994. Springer Verlag.

[21] J. Massey. SAFER K-64: A byte-oriented block-ciphering algorithm. In R. Anderson, editor, *Fast Software Encryption*, volume 809 of *Lecture Notes in Computer Science*, pages 1–17, 1994. Springer Verlag.

[22] M. Matsui. Linear cryptanalysis method for DES cipher. In T. Helleseth, editor, *Advances in Cryptology — Eurocrypt '93*, volume 765 of *Lecture Notes in Computer Science*, pages 386–397, 1994. Springer-Verlag.

[23] M. Matsui. The first experimental cryptanalysis of the Data Encryption Standard. In Y. G. Desmedt, editor, *Advances in Cryptology — Crypto '94*, volume 839 of *Lecture Notes in Computer Science*, pages 1–11, New York, 1994. Springer-Verlag.

[24] S. Moriai, K. Aoki, and K. Ohta. Key-dependency of linear probability of RC5. March 1996. To appear in *IEICE Trans. Fundamentals*.

[25] National Institute of Standards and Technology (NIST). *FIPS Publication 46-2: Data Encryption Standard*, December 30, 1993.

[26] National Security Agency. Skipjack and KEA algorithm specifications. May 1998. Available at `csrc.ncsl.nist.gov/encryption/skipjack-1.pdf`.

[27] K. Nyberg. Linear approximation of block ciphers. In A.D. Santis, editor, *Advances in Cryptology — Eurocrypt '94*, volume 950 of *Lecture Notes in Computer Science*, pages 439–444, 1994. Springer-Verlag.

[28] R.L. Rivest. The RC5 encryption algorithm. In B. Preneel, editor, *Fast Software Encryption*, volume 1008 of *Lecture Notes in Computer Science*, pages 86–96, 1995. Springer Verlag.

[29] R.L. Rivest, M.J.B. Robshaw R. Sidney and Y.L. Yin. The RC6 Block Cipher. v1.1, August 20, 1998. Available at `www.rsa.com/rsalabs/aes/`

[30] B. Schneier. Description of a new variable-length key, 64-bit block cipher (Blowfish). In R. Anderson, editor, *Fast Software Encryption*, volume 809 of *Lecture Notes in Computer Science*, pages 191–204, 1994. Springer Verlag.

[31] A. A. Selcuk. New results in linear cryptanalysis of RC5. In S. Vaudenay, editor, *Fast Software Encryption*, volume 1372 of *Lecture Notes in Computer Science*, pages 1–16, 1998, Springer-Verlag.

[32] S. Vaudenay. On the weak keys of Blowfish. In D. Gollmann, editor, *Fast Software Encryption*, volume 1039 of *Lecture Notes in Computer Science*, pages 27–32, 1996. Springer Verlag.