An Evaluation of the Security of CipherUnicorn-E Against Certain Attacks

David Wagner

December 17, 2001

Executive Summary

This report studies the security of the block cipher CipherUnicorn-E against two classes of attack: differential and linear cryptanalysis. I present the results of a limited time analysis of this cipher.

The analysis provides strong evidence that CipherUnicorn-E resists conventional differential and linear cryptanalysis attacks. In light of these results, it seems unlikely that a differential or linear cryptanalytic attack on the full CipherUnicorn-E construction will succeed without major new ideas.

1 Introduction

CipherUnicorn-E is a 64-bit block cipher proposed in the CRYPTREC standards effort. In this report, I evaluate the security of CipherUnicorn-E against differential cryptanalysis and linear cryptanalysis.

Section 2 examines whether differential cryptanalysis can be used to mount an attack on CipherUnicorn-E. The answer is no: CipherUnicorn-E appears to be secure against conventional differential cryptanalysis. The technical results can be summarized as follows: I argue that the probability of the best differential characteristic of the CipherUnicorn-E round function is upper bounded by 2^{-21}, and then this implies that the probability of the best differential characteristic is at most 2^{-126}, which effectively rules out the possibility of a successful differential cryptanalytic attack on CipherUnicorn-E. The conclusion is that CipherUnicorn-E has adequate security against conventional differential cryptanalysis.

Then, Section 3 examines whether linear cryptanalysis can be used to mount an attack on CipherUnicorn-E. This section provides strong evidence that CipherUnicorn-E is secure against conventional linear cryptanalysis. The technical results are as follows: no non-trivial linear characteristic for one round has bias larger than $2^{-12.3}$, so there is no linear characteristic for the full cipher with bias larger than 2^{-74}, and hence linear cryptanalytic attacks against CipherUnicorn-E can be expected to fail. The conclusion is that CipherUnicorn-E has adequate security against conventional linear cryptanalysis.

1.1 Caveats

The results of this evaluation should be interpreted with care.

First, this is a limited time evaluation by a single cryptographer. It is widely accepted in the cryptographic community that acquiring enough confidence in a new cipher usually requires years of scrutiny by the cryptographic community at large. While this report does describe significant evidence for the security of CipherUnicorn-E, it is no substitute for years of public study.

Second, this evaluation was limited in scope: I only studied security against conventional differential and linear cryptanalysis. While this report does provide strong evidence that these two classes of attacks will not work against CipherUnicorn-E, no promises can be made about other types of attacks. I have not attempted to evaluate the security of CipherUnicorn-E against generalizations of differential and linear cryptanalysis, such as truncated differential cryptanalysis, higher-order differential cryptanalysis, impossible differentials, boomerang attacks, and so on. Nor have I attempted to evaluate the security of CipherUnicorn-E against other attacks known in the literature, such as interpolation attacks, integrals, or slide attacks. I am not aware of any viable attacks on CipherUnicorn-E, but I have not studied other potential methods for attacking CipherUnicorn-E in any depth.

Third, on a technical note, I have focused primarily on differential characteristics and linear characteristics, rather than differentials or linear hulls. Recall that differential characteristics specify one way that differences can propagate through the cipher,

whereas differentials specify many ways: characteristics specify "one trail", while differentials specify "many trails." From the point of view of security evaluation, there are two important differences. First, differentials may have higher probability than characteristics, and the existence of a high-probability differential is sufficient to break a cipher, even if no high-probability characteristics exist. Second, characteristics are much more amenable to evaluation than differentials: while it is often feasible to bound the probability of the best characteristic, finding the highest-probability differential is typically an intractable task. Similar comments apply to linear characteristics vs. linear hulls.

In this report, I study only characteristics, not differentials or hulls. Evaluating the differentials and hulls of CipherUnicorn-E appears to be beyond the state of the art, and I do not know how to find a provable bound on their probability. Therefore, I shall rely on a standard heuristic, which says that the probability of the best differential is not much larger than the probability of the best characteristic. This heuristic, while not necessarily justified in principle from a theoretical viewpoint, seems to be accurate in practice for many ciphers. While I do not expect that this feature of our analysis will cause any inaccuracy, and while this limitation seems to be more or less unavoidable with the current state of the art in cryptography, the reader should recognize that there is always the chance that the heuristic could turn out to be inaccurate, and in this case the security of CipherUnicorn-E would need to be revisited.

That said, the results presented in this report do present good reasons to be quite confident in the security of CipherUnicorn-E against differential and linear cryptanalysis.

1.2 Notation

I use $x \oplus y$ for the XOR of x and y, $x \boxplus y$ for their sum modulo 2^{32}, and $x \lll y$ for the left-shift of x by y positions. I will sometimes write a 32-bit value x in terms of its four bytes, $x = (a,b,c,d)$, where a denotes the most significant byte and d the least significant, i.e., $x = a \cdot 2^{24} + b \cdot 2^{16} + c \cdot 2^8 + d$.

2 Differential Cryptanalysis of CipherUnicorn-E

CipherUnicorn-E is a Feistel cipher with 16 rounds, using a 32-bit round function $F : \{0,1\}^{32} \to \{0,1\}^{32}$. The security of the cipher against differential cryptanalysis relies on the non-existence of high-probability differential characteristics $\Delta X \to \Delta Y$ for F. Consequently, we will work towards a upper bound on the probability of the best differential characteristic for F.

Recall that the probability of a characteristic $\Delta X \to \Delta Y$ for F is defined as

$$\Pr[\Delta X \to \Delta Y] \stackrel{\text{def}}{=} \Pr[F(X) \oplus F(X \oplus \Delta X) = \Delta Y],$$

where on the right-hand side X is a random variable uniformly distributed on $\{0,1\}^{32}$. We will calculate a upper bound on the probability of the best characteristic, i.e., an estimate for $\max_{\Delta X, \Delta Y \neq 0} \Pr[\Delta X \to \Delta Y]$.

In CipherUnicorn-E, the round function F has the structure

$$F(x) = f(x \boxplus k, g(x \boxplus k))$$

where f represents the "main stream", g the "temporary key generation function" (in the terminology of the designers), and k part of the round subkey. We can analyze the probability of the best differential characteristic for F in terms of the best differential characteristics of f and g separately.

I consider each of these two substructures in turn, next. First, I show an upper bound 2^{-14} for the best characteristic for f. Then, I show an upper bound 2^{-7} for the best characteristic for g. The conclusion is that there is there is no characteristic for the full round function F with probability better than 2^{-21}. (Most likely the true value is much smaller, but this will suffice as an upper bound.)

This bound suffices to show that CipherUnicorn-E is secure against conventional differential cryptanalysis. If we assume that the attacker can bypass the first round and mount a 2-R attack, the attacker will need to find a 13-round differential characteristic. Every such characteristic involves at least 6 active rounds, and since each active round has probability at most 2^{-21}, we see that the probability of the best 13-round differential characteristic will be at most $2^{-21\times 6} = 2^{-126}$. As such an attack would require vastly more chosen plaintexts than are available, we conclude that CipherUnicorn-E appears to be secure against differential cryptanalysis.

2.1 Differential Characteristics of f, the "Main Stream"

The function f is computed as below.

Algorithm $f(s,z)$:
1. Apply $T(1)$, $T(2)$, $T(3)$, $T(4)$ to s to obtain t.
2. Compute $u = t \boxplus k'$, where k' is some additional key material.
3. Apply $T(1)$, $T(2)$, $T(3)$, $T(4)$ (in a permuted order determined by four bits of z) to u to obtain v.
4. XOR a byte of y into v, then apply a $T(\cdot)$ round (which one is selected by z).
5. XOR another byte of z into the result, then apply another $T(\cdot)$ round, and call the result w.
6. Return w as the value of $f(s,z)$.

Any non-trivial differential characteristic for Step 1, above, must include at least one active $T(\cdot)$-round. An exhaustive computer search confirms that the probability of the best differential characteristic for a single active $T(\cdot)$-round is $1/2^7$. Consequently, we can conclude that the best differential characteristic through Step 1 has probability at most $1/2^7$.

Similarly, any non-trivial differential characteristic that goes through Step 3 also involves an active $T(\cdot)$-box in Step 3, and thereby contributes another factor of $1/2^7$ (or smaller) to the overall probability. This allows us to argue that the best differential characteristic for f, i.e., for the main stream, has probability at most $1/2^7 \times 1/2^7 = 2^{-14}$.

At this point, a caveat should be mentioned. The alert reader might ask: are these probabilities are truly independent? The answer is not entirely clear. Without the

insertion of the subkey k' in Step 2, the answer would be a definite 'no', and our bound of 2^{14} would be completely unjustified; fortunately, however, Step 2 *does* introduce key material, so our bound above may be ok. If Step 2 introduced key material using the XOR operation, instead of using addition modulo 2^{32}, then the two probabilities for Step 1 and Step 3 would indeed be independent, since XORing with key material erases everything except the XOR-difference; consequently, the upper bound of 2^{-14} would be fully justified in this case.

Unfortunately, the use of addition rather than XOR complicates matters. I know of no proof that Steps 1 and 3 should be independent, so in principle, multiplying the probabilities of Step 1 and Step 3 is not necessarily justified. The theoretical issue is that the values after Step 2 retain information not only on the XOR-difference before Step 2, but also some partial information on the values of both inputs to Step 2. In practice, though, there is every reason to believe that the use of addition is at least as good as, and possibly better than, XOR. I tried very hard to find even a contrived scenario where use of addition could be less secure than XOR, and I could not find one. In practice, addition reduces the differential probability even further, because the carry bits introduce a limited amount of non-linearity that makes it hard to predict the differences entering Step 3.

In summary, then, independence is not strictly 100% guaranteed from a theoretical standpoint, but I am confident that in practice the independence assumption is unlikely to cause inaccuracy in our analysis. This situation, by the way, is a consequence of the design philosophy of "adopting a structure that is difficult to analyze": such a structure may be difficult for the malicious attacker to analyze, but this philosophy comes with the disadvantage that it is also difficult for legitimate evaluators to develop reliable evidence for the security of CipherUnicorn-E. In any case, I fully expect that there is no differential characteristic for f with a probability noticeably exceeding 2^{-14}. In the rest of this report, I will use 2^{-14} as the upper bound on the probability of the best differential characteristic for f (the "main stream").

One might wonder whether this upper bound is very conservative. At first glance, it appears to complete ignore the extra security provided by the modular addition in Step 2 and by the final two K/T-rounds in Steps 4 and 5. However, I do not believe those features provide as much security as one might expect. In particular, I believe one can find a differential characteristic for f with probability fairly close to the upper bound. One possibility might look something like this:

$$\begin{aligned}
\Delta s &\xrightarrow{T(1)} (0,0,0,\Delta t_3) \xrightarrow{T(2,3,4)} (0,0,0,\Delta t_3) && \text{(Step 1, prob. } 2^{-7}\text{)} \\
&\xrightarrow{\boxplus k'} (0,0,0,\Delta u_3) && \text{(Step 2, prob. } \approx 1\text{)} \\
&\xrightarrow{T(2,4,3)} (0,0,0,\Delta u_3) \xrightarrow{T(1)} (\Delta v_1,\Delta v_2,\Delta v_3,\Delta v_4) && \text{(Step 3, prob. } 2^{-7}\times 2^{-4}\text{)} \\
&\xrightarrow{K(2)} (\Delta v_1,0,\Delta v_3,\Delta v_4) \xrightarrow{T(2)} (\Delta v_1,0,\Delta v_3,\Delta v_4) && \text{(Step 4, prob. 1)} \\
&\xrightarrow{K(4)} (\Delta v_1,0,0,\Delta v_4) \xrightarrow{T(4)} (\Delta v_1,0,0,\Delta v_4) && \text{(Step 5, prob. 1)}
\end{aligned}$$

where, e.g., $\Delta z = (0,*,\Delta v_2,\Delta v_3)$.

We explain the ideas behind the above characteristic in detail next. It is straightforward to find a differential characteristic $\Delta s \to \Delta t$ for the first four T-rounds that holds with probability 2^{-7} so that $\Delta t = (0,0,0,\Delta t_3)$, i.e., Δt is zero in all but the low byte: the active round will be the first T-round, and I have verified that one can find many characteristics for the first T-round that hold with probability 2^{-7}. We can additionally choose t so that Δt has low Hamming weight, and then the difference will pass through the modular addition (Step 2) with high probability (I assume this probability can be chosen to be very close to 1). We will hope that, through blind luck, Step 3 will use an ordering of the T-boxes that uses $T(1)$ last in both pairs; in this case, the first three T-rounds will have probability 1, and we can arrange the final $T(1)$-round to have probabilty 2^{-7}. If we can choose Δz appropriately, we can ensure that both pairs use the same T-ordering in Step 3, and then the probability that $T(1)$ appears last is 2^{-2}. Finally, we can choose Δz so that the differences entering the last two T-rounds (Steps 4 and 5) are cancelled out by the K-round before passing through the S-boxes, thereby ensuring that Steps 4 and 5 are passed through with probability 1. We may need another factor of 2^{-2} to control the order of the other T-rounds in Steps 4 and 5.

To summarize the implications of this lengthy discussion, we obtain a carefully chosen differential characteristic $(\Delta s, \Delta z) \to \Delta w$ for f that holds with probability close to 2^{-18}. This is not too far from the simple upper bound of 2^{-14} derived above, so I suspect the 2^{-14} upper bound probably cannot be significantly improved without a lot more work. This completes the analysis of f, the "main stream."

2.2 Differential Characteristics of g, the "Temporary Key Generation Mechanism"

The analysis of g is a bit more involved. The function g incorporates many operations that exploit the additive structure (rather than the XOR structure) of 32-bit words.

In particular, the Y-rounds use addition modulo 2^{32} as well as left-shift: e.g., computations such as $x' = x \boxplus (x \lll 3)$. A left-shift by k positions can be perhaps best viewed as multiplication by 2^k, and so we find that the Y-rounds correspond to multiplication modulo 2^{32} of the input by some constant. For instance, the above-mentioned example $x' = x \boxplus (x \lll 3)$ can be equivalently expressed as $x' = (1+2^3)x \bmod 2^{32}$. We see that the $Y(3,8,16)$ round corresponds to multiplication by $(1+2^3)(1+2^8)(1+2^{16})$ modulo 2^{32}, and the $Y(7,9,13)$ round to multiplication by $(1+2^7)(1+2^9)(1+2^{13})$. Fortunately, both of these constants are odd, so the Y rounds are bijective.

The additive structure of the Y-rounds and presence the subkey additions suggest that XOR-differentials may not do so well: additive differentials modulo 2^{32} may be more appropriate. However, the T-rounds are best approximated using XOR-differentials, not additive differentials, so it is clear that we need to consider both viewpoints.

We recall the generalization of XOR-differentials to other groups. If (x, x^*) represents a pair of encryptions, then their XOR-difference and additive-difference are defined as follows:

$$\Delta_{\oplus} x \stackrel{\text{def}}{=} x^* \oplus x \qquad \Delta_{\boxplus} x \stackrel{\text{def}}{=} x^* \boxminus x = x^* - x \bmod 2^{32}.$$

We note that it is possible to convert between the two types of differences: for instance,

if $\Delta_{\oplus}x = 1$, then $\Delta_{\boxplus}x = 1$ holds with probability about 1/2, or in other words,

$$\Pr[\Delta_{\boxplus}x = 1 \mid \Delta_{\oplus}x = 1] = 1/2 \qquad \Pr[\Delta_{\boxplus}x = 3 \mid \Delta_{\oplus}x = 1] = 1/4 \qquad \ldots$$

The conversion works this way because of the carry bits in addition. One finds that the conversion tends to hold with high probability only for differences with low Hamming weight, or more generally, with a small number of runs of 1's in their binary representation.

For convenience in writing long differentials combining both additive and XOR-differentials, we will sometimes subscript the difference itself with the appropriate symbol to indicate which type of difference is being considered. For instance, $1_{\oplus} \to 7_{\boxplus}$ represents the differential holding when $F(X \oplus 1) \boxminus F(X) = 7$. I will consider differentials where we are allowed to use either additive or XOR-differentials at essentially every position, and I will look for an upper bound on the probability of the best differential characteristic for g.

First, note that the Y-rounds have many probability-1 additive differential characteristics.

$$\delta_{\boxplus} \xrightarrow{Y(i,j,k)} \gamma_{\boxplus} \qquad \text{(prob. 1)}$$

$$\text{when } \gamma = (1+2^i)(1+2^j)(1+2^k)\delta \bmod 2^{32}.$$

Note that if we know the input difference δ, we can compute the output difference γ using a single multiplication, and similarly we can go backwards by using the inverse modulo 2^{32}.

Subkey additions also leave additive differences unchanged (with probability 1). This makes additive differential characteristics a very effective tool for attacking the subkey additions and the Y-rounds.

Second, note that we can obtain probability-1 characteristics for the T-rounds if we use XOR-differences. For instance, we have the following

$$(0,a,b,c)_{\oplus} \xrightarrow{T(1)} (0,a,b,c)_{\oplus} \qquad \text{(prob. 1)}$$

$$(0,0,b,c)_{\oplus} \xrightarrow{T(1),T(2)} (0,0,b,c)_{\oplus} \qquad \text{(prob. 1)}$$

We can see that the only difficulty in constructing high-probability differential characteristics for g (the "temporary key generation mechanism") will be the need to convert between additive and XOR-differences. We have probability-1 characteristics for all components of g, so if we could convert freely between the two without restriction, we would have a probability-1 characteristic for g by combining the characteristics for the components. In real life, though, things don't work this way: we must convert between additive and XOR-differences to get the characteristics to line up, and this comes at a price of reducing the probability of the characteristic.

It is not hard to see that we need to switch between additive and XOR-differences in at least four positions. The natural places are before and after each of the two segments of two T-rounds. Because the interaction of additive and XOR-differences are difficult to analyze theoretically, due to the complications created by the carry bits, I carried out

an computer search for the best differential characteristic. In particular, I searched for all differential characteristics in which the T-rounds are passed through with probability 1. This places enough restrictions on the structure of a differential characteristic to exhaustively search all possibilities: there are essentially only 2^{16} possible differences to consider just before the pair of $T(1), T(2)$-rounds, and working backwards through g we can enumerate all possible high-probability characteristics.

The results of this exhaustive computer search are as follows. No characteristics with probability exceeding 2^{-13} were found. One of the best characteristics was the following (with values reported in hexadecimal):

$$\texttt{0xB787C072}_{\boxplus} \xrightarrow{g} \texttt{0x00007F02}_{\boxplus} \qquad (\text{prob.} \approx 2^{-13})$$

There were a number of other characteristic with similar and smaller probabilities. This particular differential characteristic can be expanded, showing intermediate values, as follows:

$$\begin{aligned}
\texttt{0xB787C072}_{\boxplus} &\xrightarrow{\boxplus k'} \texttt{0xB787C072}_{\boxplus} && (\text{prob. } 1)\\
&\xrightarrow{Y(3,8,16)} \texttt{0xFF8BC602}_{\boxplus} = \texttt{0x008...}_{\oplus} && (\text{prob. } 1)\\
&\xrightarrow{T(1)} \texttt{0x008..}_{\oplus} = \texttt{0x0073C5FE}_{\boxplus} && (\text{prob.} \approx 2^{-10})\\
&\xrightarrow{\boxplus k''} \texttt{0x0073C5FE}_{\boxplus} && (\text{prob. } 1)\\
&\xrightarrow{Y(7,9,13)} \texttt{0x000080FE}_{\boxplus} = \texttt{0x0000...}_{\oplus} && (\text{prob. } 1)\\
&\xrightarrow{T(1),T(2)} \texttt{0x0000...}_{\oplus} = \texttt{0x00007F02}_{\boxplus} && (\text{prob.} \approx 2^{-3})
\end{aligned}$$

A word of warning: These probabilities are very approximate. Because I was searching primarily for an upper bound on the probability of the best differential characteristic, it is possible that the true probability of this differential characteristic is considerably lower. I did not try to confirm whether the full characteristic for g is even possible; instead, I focused on upper bounds for characteristics for each of the components. Also, I did not spend a lot of effort trying to assess this characteristic from the point of view of an attacker—rather, I focused on finding an upper bound on the highest probability characteristic, with the goal of showing that no ordinary differential cryptanalytic attack on CipherUnicorn-E is likely to succeed.

If we try to maximize the probability only of passing through $T(1)$ (and converting between additive and XOR-differences around this T-round), ignoring the probability of the final two T-rounds, the best probability is still fairly low: at most about 2^{-8}. One example of such a characteristic, with probability around 2^{-8}, is

$$\begin{aligned}
\texttt{0x563BC537}_{\boxplus} &\xrightarrow{\boxplus,Y} \texttt{0xFFF7DDEF}_{\boxplus} \xrightarrow{T(1),\boxplus} \texttt{0x00082211}_{\boxplus} && (\text{prob.} \approx 2^{-8})\\
&\xrightarrow{Y} \texttt{0x00006C91}_{\boxplus} \xrightarrow{T(1),T(2)} \cdots && (\text{prob. } 1)
\end{aligned}$$

This suggests that the conversion between additive and XOR-differences is the main barrier to high-probability differential characteristics.

Returning to the study of the full g function, the results of the computer search can be summarized as follows. The computer search enumerated all characteristics where the three T-rounds are passed through with probability 1, i.e., where the S-boxes are inactive. Assuming there are no bugs in my computer program, this shows there are no characteristics for g from this class with probability better than 2^{-13}. We can see that the conversion between additive and XOR-differences has imposes serious degradation on the probability of characteristics.

We can also consider all characteristics where some T-round is active. I verified exhaustively that there is no characteristic for any active T-round with probability larger than 2^{-7}. This shows that, among the characteristics with an active T-round, none have probability higher than 2^{-7}. In fact, I suspect that all such characteristics have a much lower probability, because if a T-round is active we obtain intermediate differences with very high Hamming weight, which makes the probability of passing through the additions and Y-rounds quite low. As a result, I suspect that there is no characteristic for g from this class with probability greater than 2^{-13}, but I have no proof, and I do not know how to verify this conjecture. In any case, the weaker (and fully-justified) bound of 2^{-7} will be sufficient for our purposes.

Note that these two classes of characteristics cover all possibilities. In every characteristic, either some T-round is active (and we have a characteristic of the second class), or all T-rounds are inactive (and we have a characteristic of the first class). Therefore, an upper bound on the probability of the best characteristic is $\max(2^{-13}, 2^{-7}) = 2^{-7}$.

The conclusion is that there is no differential characteristic for g (the "temporary key generation mechanism") that holds with probability greater than 2^{-7}. Most likely, the true maximum differential probability is much smaller—I conjecture that it is probably at most 2^{-13} or so—but despite the existence of partial evidence, I do not know of any proof for this conjecture. Therefore, I will use only the fully-justified and more conservative figure of 2^{-7}.

3 Linear Cryptanalysis of CipherUnicorn-E

Linear cryptanalysis seems to be a much less effective method for attacking CipherUnicorn-E than differential cryptanalysis. This is perhaps not surprising, because this sort of construction (an "unbalanced, target-heavy generalized Feistel network" [2]) tends to much more resistant to linear attacks than to differential attacks. As a consequence, analyzing the security of CipherUnicorn-E against linear cryptanalysis seems much easier than analyzing its security against differential attacks.

Recall that the probability of a linear approximation $\Gamma \rightarrow \Gamma'$ is defined as

$$\Pr[\Gamma \rightarrow \Gamma'] \stackrel{\text{def}}{=} \Pr[(X \cdot \Gamma) \oplus (F(X) \cdot \Gamma') = 1],$$

where on the right-hand side X is a random variable uniformly distributed on $\{0,1\}^{32}$ and where $x \cdot y$ denotes the dot-product of two 32-bit vectors. The bias of the linear approximation $\Gamma \rightarrow \Gamma'$ is then defined as

$$\text{Bias}[\Gamma \rightarrow \Gamma'] \stackrel{\text{def}}{=} 2|\Pr[\Gamma \rightarrow \Gamma'] - 1/2|.$$

Note that biases multiply when we concatenate independently-keyed functions (this is the content of Matsui's piling-up lemma), and that a linear characteristic of bias b can be used in an attack using on the order of $1/b^2$ known texts (Matsui's rule of thumb). We will calculate a upper bound on the bias of the best linear characteristic, i.e., an estimate for $\max_{\Gamma,\Gamma' \neq 0} \text{Bias}[\Gamma \rightarrow \Gamma']$.

We start by analyzing the bias of the best linear characteristic for the CipherUnicorn-E Feistel function. Note that any non-trivial linear approximation for a T-round (inside the Feistel function) must include at least one active S-box. Moreover, the CipherUnicorn-E Feistel function includes at least 8 T-rounds. If we let b_T denote the bias of the best linear characteristic for a single T-round, then the bias of the best linear characteristic for the Feistel function will be at most b_T^8 (again, by the piling-up lemma). This then gives us a bound for the bias of the best 13-round linear characteristic: it can be no larger than $(b_T^8)^6 = b_T^{48}$.

I do not know know what the maximum bias b_T for a single T-round is. I have verified computationally that, if we restrict to the case where at most two S-boxes are active, the best probability is $1/2 \pm 44/256$, and hence the best bias is $88/256 \approx 2^{-1.54}$. This was checked by an exhaustive search over all the 2^{24} such linear characteristics, a computation with workfactor approximately 2^{32}. Note that every linear characteristic for a T-round can be specified by a 40-bit quantity: the 8-bit mask Γ entering the S-box inputs, and a 32-bit mask covering the entire output of the T-round.

If this result also were to extend to the case where three or four S-boxes are active, we would have a bound $2^{-1.54 \times 48} \approx 2^{-74}$ on the best bias for the whole cipher. A linear attack using a characteristic of bias 2^{-74} would need at least 2^{148} known texts, which vastly exceeds the number of texts available to any attacker. Consequently, if this partial bound extends to the unanalyzed cases, this would be compelling evidence that CipherUnicorn-E is secure against linear cryptanalysis.

Unfortunately, the task of finding the best linear characteristic for a single T-round appears challenging. I do not know of any easy way to compute the exact value of b_T. The problem is that the naive approach—namely, enumerating all 2^{40} of the possible characteristics—requires a very large workfactor (about 2^{48} operations). This means that an exhaustive enumeration strategy appears to be out of the reach of today's computers.

I am forced to leave it as an open problem to estimate the true value of b_T. This seems like an interesting algorithm problem, and there is some hope for algorithmic improvements in this regard. We note that it is equivalent to the following problem (which is in turn closely connected to coding theory):

> Given a 256×41 matrix M over $GF(2)$ and a value $n \in \mathbb{R}$, check whether there exists a 41-bit column vector $v \neq 0$ so that Mv is a vector of Hamming weight at most n.

The connection is as follows. Fill in each row of the matrix M with the 8-bit input to the S-boxes, the 32-bit output formed of the T function, and a final column that always contains a 1 bit. Then there is a one-to-one correspondence between vectors v such that Mv has low weight and linear characteristics $\Gamma \rightarrow \Gamma'$ of the T-round of large bias: the correspondence is given by $v = (\Gamma, \Gamma', v_{41})$, and Mv has weight n if and only if $\text{Bias}[\Gamma \rightarrow \Gamma'] = |n/128 - 1|$.

There is some hope for solving this algorithmic problem. For instance, if we want to check whether there exists a vector Mv with weight at most n, then we could use the following randomized algorithm:

1. Do the following $c \times (1 - n/256)^{-41}$ times:
2. Pick 41 random columns of M, and erase the rest. We obtain a 41×41 matrix M'.
3. For all solutions v to the equation $M'v = 0$ with $v \neq 0$, do:
4. If Mv has weight at most n, output v.
5. Output "no solution found."

One can expect that this algorithm will output an incorrect answer with probability at most e^{-c}. For an error probability of 2^{-20} and with $n = 64$, this algorithm will require about 2^{21} inversions of a 41×41 matrix, corresponding to about $2^{21} \times 41^3 \approx 2^{37}$ operations. The output would tell us, with high confidence, whether or not there is any linear characteristic for the CipherUnicorn-E T-round with bias exceeding $1/2$. If the answer is that there is no such characteristic (as we expect), then this would imply (except for a 2^{-20} probability of error) that there is no linear characteristic with bias exceeding 2^{48} for the full CipherUnicorn-E cipher, and this would be sufficient to rule out the possibility of linear cryptanalytic attacks on CipherUnicorn-E.

Unfortunately, due to time constraints, I have not been able to develop this direction further. I am forced to leave it as an open problem to determine the probability of the best linear characteristics for the CipherUnicorn-E S-boxes and T-rounds.

I will note in passing that my brute-force search for good linear characteristics found a number of linear approximations that seem to contradict the designer's security claims for the CipherUnicorn-E S-boxes. For example, here are four linear characteristics of the T-function. I specify the 8-bit mask Γ covering the input byte to the S-boxes and the 32-bit mask $\Gamma' = (\Gamma'_0, \Gamma'_1, \Gamma'_2, \Gamma'_3)$ covering the output of the T-function (where Γ'_i is the mask for the S-box S_i). All masks are given in hexadecimal.

$$\begin{aligned} \Pr[D6 \rightarrow (16, CB, 00, 00)] &= 1/2 \pm 44/256 \\ \Pr[46 \rightarrow (7A, 00, D1, 00)] &= 1/2 \pm 44/256 \\ \Pr[90 \rightarrow (00, 32, 73, 00)] &= 1/2 \pm 44/256 \end{aligned}$$

In contrast, the designers' self-evaluation report seems to claim a maximum probability of $1/2 \pm 2^{-3.08}$ for these approximations [1], which is a puzzling discrepancy. There seem to be at least three possibilities: I could be misunderstanding the specification of the cipher or their analysis; my program could be in error; or the designers' calculations could be erroneous. I do not know what the cause of this discrepancy is, and perhaps it should be investigated further. In any case, the conclusion that CipherUnicorn-E appears to be secure against linear cryptanalysis seems to remain intact given our current knowledge.

References

[1] NEC Corporation, "Self Evaluation Report: CIPHERUNICORN-E," CRYPTREC submission, 26 pages, Ver. 3.

[2] Bruce Schneier and John Kelsey, “Unbalanced Feistel Networks and Block Cipher Design,” *Fast Software Encryption ’96*, LNCS 1039, pp.121–141, 1996.