

A note on the Edon80 S-box

T.E. Bjørstad

The Selmer Center, Department of Informatics,
University of Bergen, Pb. 7800, N-5020 Bergen, Norway.
Email : tor.bjorstad@ii.uib.no

Abstract. Edon80 [1] is one of the Phase 3 candidates in eSTREAM, the ECRYPT stream cipher project. This note examines the structure of the Edon80 quasigroup permutation, when viewed as an S-box or a pair of boolean functions. Although some interesting relations are found, we have *not* been able to apply these to attack the full cipher.

1 Introduction

The Edon80 stream cipher consists of 80 pipelined stages. Each stage has 4 bits of internal state: two fixed bits $k = (k_1, k_0)$ which are part of the secret key, and two dynamic state bits $s = (s_1, s_0)$ which are updated every clock cycle. The key bits are used to select one of four quasigroup permutations \star_i . During keystream generation, the state bits s of stage i at time t are updated by the rule

$$s_{i,t+1} = s_{i,t} \star_i s_{i-1,t}. \quad (1)$$

The initial stage takes as input the periodic string 01230123...0123..., while the last stage outputs its current s as keystream every other cycle. Key and IV setup is described in further detail in [1]. In the following, we denote the input

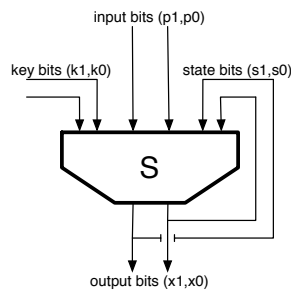


Fig. 1. The Edon80 S-box

coming from the previous stage as $p = (p_1, p_0)$, and the output of the S-box (the new state) as $x = (x_1, x_0)$. The most significant bit of each value is the first, so the notation $p = 2$ and $p_1 = 1, p_0 = 0$ will be used interchangeably. We shall refer to a single application of the quasigroup transformation as a *round*.

2 The Edon80 Quasigroups

The four Edon80 quasigroups are shown in Fig. 2. According to [1] the permutations have been chosen among 64 “good” choices among the 576 possible quasigroups.

\star_0	0 1 2 3	\star_1	0 1 2 3	\star_2	0 1 2 3	\star_3	0 1 2 3
0	0 2 1 3	0	1 3 0 2	0	2 1 0 3	0	3 2 1 0
1	2 1 3 0	1	0 1 2 3	1	1 2 3 0	1	1 0 3 2
2	1 3 0 2	2	2 0 3 1	2	3 0 2 1	2	0 3 2 1
3	3 0 2 1	3	3 2 1 0	3	0 3 1 2	3	2 1 0 3

Fig. 2. The Edon80 quasigroup permutations

Instead of taking a group-theoretical approach, we consider the quasigroup transformation as an S-box mapping a 6 bit input to 2 bits of output. We denote this map $S : \{0, 1\}^6 \rightarrow \{0, 1\}^2$, where the individual bits $(k_1 k_0 s_1 s_0 p_1 p_0) \mapsto (x_1 x_0)$. Alternately, S may be viewed as an ordered pair of boolean functions $f(x)$ and $g(x)$, which map 6 input bits to x_1 and x_0 respectively.

The best linear approximation to the entire Edon80 S-box transformation may be represented by the matrix

$$L = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}. \quad (2)$$

This approximation holds with probability 1/2 (i.e. for 32 of the 64 possible inputs). However, the approximation depends crucially on the key bits: if we set $k_0 = 1$, then the approximation holds with probability 3/4, whereas it only holds with probability 1/4 for the remaining quasigroups. In general, the action of each quasigroup may be approximated by a linear map that holds with probability 3/4, but it is somewhat odd that the same approximation holds optimally for two of the quasigroups.

2.1 Boolean functions

It may readily be confirmed that the functions $f(x)$ and $g(x)$ are of degree four.

$$\begin{aligned} f(x) &= p_0 + s_0 + k_1 \\ &+ k_0 s_0 + k_0 s_1 + k_1 p_1 \\ &+ k_0 s_0 p_0 + k_0 s_0 p_1 + k_1 s_1 p_1 + k_1 k_0 p_0 + k_1 k_0 s_0 \\ &+ k_1 k_0 s_0 p_0 + k_1 k_0 s_0 p_1 + k_1 k_0 s_1 p_0 + k_1 k_0 s_1 p_1 \end{aligned} \quad (3)$$

$$\begin{aligned} g(x) &= p_1 + s_1 + k_0 \\ &+ s_0 p_0 + k_0 s_0 + k_1 p_0 + k_1 p_1 + k_1 s_0 \\ &+ k_0 s_0 p_1 + k_1 s_0 p_0 + k_1 s_1 p_1 \\ &+ k_1 k_0 s_0 p_1 + k_1 k_0 s_1 p_1 \end{aligned} \quad (4)$$

The following are the best affine approximations to f and g . Both hold with probability $11/16$ (bias $3/16$). These approximations are not independent as they depend on the same variables; the probability that *both* these approximations hold simultaneously is only $7/16$.

$$f(x) \cong k_1 + s_0 + p_0 \tag{5}$$

$$\cong k_1 + k_0 + s_0 + p_0$$

$$g(x) \cong k_1 + s_1 + p_0 \tag{6}$$

$$\cong k_1 + k_0 + s_1 + p_0 + 1$$

The four approximations all depend on bits from all three inputs, k , s and p . A simple search reveals four additional relations between input- and output bits in the S-box that hold with probability $11/16$.

$$s_1 + p_1 + x_1 + x_0 + 1 = 0$$

$$k_0 + s_1 + p_1 + x_1 + x_0 + 1 = 0$$

$$k_1 + s_0 + p_1 + p_0 + x_1 + x_0 + 1 = 0$$

$$k_1 + k_0 + s_0 + p_1 + p_0 + x_1 + x_0 = 0$$

(7)

The most interesting of these relations is probably the first one, which does not depend on any of the key bits k at all.

2.2 Restrictions on the boolean functions

In some cases we have partial knowledge of the input to the S-box. Situations where this is always true include the final S-box (where the state s is output as keystream every other cycle and hence known), the first S-box (whose input p

Restriction	f-approximation	Prob.	g-approximation	Prob.
$k = 0$	$s_0 + p_0$	1	$s_1 + p_1$	$3/4$
$k = 1$	$s_1 + s_1 + p_0$	$3/4$	$s_1 + p_0$	$3/4$
$k = 2$	$s_0 + p_0 + 1$	$3/4$	$s_0 + p_0$	$3/4$
$k = 3$	$s_0 + p_1 + 1$	$3/4$	$s_1 + p_0 + 1$	1
$s = 0$	p_0	$3/4$	$k_0 + p_1$	$3/4$
$s = 1$	p_1	$3/4$	$k_1 + p_0$	$7/8$
$s = 2$	$k_1 + k_0 + p_0$	$7/8$	$k_0 + p_1 + 1$	$3/4$
$s = 3$	$k_1 + p_0 + 1$	$3/4$	$k_1 + p_0 + 1$	$3/4$
$p = 0$	$k_1 + s_0$	$3/4$	$k_0 + s_1$	$3/4$
$p = 1$	$k_1 + k_0 + s_0 + 1$	$7/8$	$k_1 + s_1 + 1$	$3/4$
$p = 2$	$s_1 + s_0$	$7/8$	$k_1 + k_0 + s_1 + 1$	$7/8$
$p = 3$	$k_0 + s_0 + 1$	$3/4$	$k_0 + s_1 + s_0 + 1$	$3/4$

Fig. 3. Restricted approximations to the S-box

is known during key generation), and during the IV setup (where the states s of stages 0-39 are initialised using the IV and a known string). In these cases, the

restrictions of the boolean functions f and g to the remaining free variables are of lower degree, and better affine approximations to f and g may be found. Table 3 gives a list of “good” affine approximations and their respective probabilities of being correct. Some of the approximations (particularly those with probability $3/4$) are not unique, though most of the alternate approximations only involve additional input terms. Of particular interest is the fact that $f(x)$ becomes linear when $k = 0$, and similarly that $g(x)$ is linear when restricted to $k = 3$.

It is also interesting to look at 1-bit restrictions on the key. A plausible attack scenario is that the adversary guesses 40 bits of the key (one bit of each S-box), and uses these restrictions to obtain a good affine approximation to the cipher. We find that the approximations obtained when guessing only one key bit are “almost” as good as when we guess both bits. The approximations given in Table 4 for $k_1 = 0$, $k_1 = 1$ and $k_1 + k_0 = 0$ are not unique.

Restriction	f-approximation	Prob.	g-approximation	Prob.
$k_0 = 0$	$k_1 + s_0 + p_0$	$7/8$	$k_1 + s_1 + s_0 + p_1 + p_0 + 1$	$3/4$
$k_0 = 1$	$s_1 + s_0 + p_1 + 1$	$3/4$	$k_1 + s_1 + p_0$	$7/8$
$k_1 = 0$	$s_0 + p_0$	$3/4$	$k_0 + s_1 + p_1$	$3/4$
$k_1 = 1$	$s_1 + s_0 + p_1 + p_0 + 1$	$3/4$	$s_1 + p_0 + 1$	$3/4$
$k_1 + k_0 = 0$	$s_0 + p_0$	$3/4$	$s_1 + p_1 + 1$	$3/4$
$k_1 + k_0 = 1$	$s_1 + s_0 + p_0$	$3/4$	$k_1 + s_1 + s_0 + p_0 + 1$	$3/4$

Fig. 4. Approximations with 1-bit restrictions on k

As we see from these tables, guessing a portion of the Edon80 secret key may be quite helpful in obtaining biased linear relations between bits of the cipher state.

2.3 Multiround approximations

A natural question when making linear approximations of a single S-box, is whether it is possible to find good linear approximations to multiple “rounds” of Edon80, preferably involving a low number of terms. If we consider each round as independent random functions that admit linear approximations of bias $3/16$, the accumulated bias of our linear relation after 80 rounds would be on the order of 2^{-114} , which is probably not very useful.

We have used exhaustive computer search to find linear relations with greater bias across 2, 3 and 4 S-box iterations. The results are somewhat encouraging, as we do obtain relations with larger biases than predicted using the Piling-up Lemma and repeated one-round approximations. However, the best 4-round relation (approximating $f[k^{(3)}, s^{(3)}, S(k^{(2)}, s^{(2)}), S[k^{(1)}, s^{(1)}, S(k^{(0)}, s^{(0)}, p)]]$) still only has a bias only $9/512$, and depends on 10 of the 20 different input, output, key and state variables. By the Piling-up Lemma, a bias of $9/512$ over 4 rounds yields an estimated bias of roughly 2^{-97} for the corresponding linear relation over

80 S-box iterations. Furthermore, most of the biased relations found depend on *more* than 10 state variables.

There are slight improvements to be made. The approximation of the first group of S-boxes can be treated as a special case, since we know the input seed at all times during key generation. This allows us to find better approximations for this particular transformation; however, the overall gain is not very large. As an example, the best 4-round approximation given the input value $p = 0$ has bias of $49/2048$. A more significant improvement is obtained by guessing part of the key. For example, if we know that the 4 round keys all obey the relation $(k_0^{(i)} + k_1^{(i)} = 0)$ (i.e. a 4-bit restriction on the 8 key bits involved) we are able to obtain relations with much greater biases. For example, we have found one relation over 4 rounds between the high order bit of the input and the high order bit of the output, which depends on 9 of the internal key and state variables, and with bias $23/256$.

2.4 Key cancellation

An interesting feature of Edon80 is the simple key schedule. The *KeySetup* algorithm splits the 80-bit secret key K into 40 2-bit values, $K_1|K_2|\dots|K_{40}$, and assigns $(k_1, k_0) = K_i$ in stages i and $i + 40$. The fact that each bit of the key is used twice, means that if we are trying to replace the cipher with a sequence of linear approximations, some key bits may appear twice and cancel out. In particular, if the same approximation is used for stages $i \dots j$ and $(i + 40) \dots (j + 40)$, *all* key bits should cancel from the equation.

As an example of this effect, consider the 4-round relation

$$p_1 + k_1^{(0)} + k_0^{(2)} + s_1^{(0)} + s_1^{(1)} + s_0^{(1)} + s_1^{(2)} + s_1^{(3)} + s_0^{(3)} + x_1 = 1, \quad (8)$$

with bias $9/512$, and the 8-round system shown in Fig. 5 where the key repeats after 4 rounds. If we apply the relation from Eq. 8, using (p, x) as the first in-

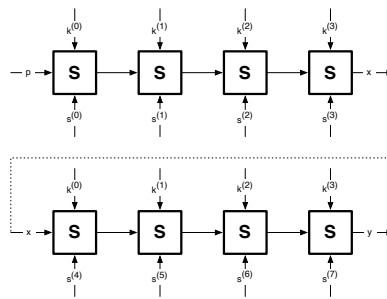


Fig. 5. The cyclical key schedule of Edon80, reduced to 2x4 rounds

put/output pair and (x, y) as the second, and combine the two, all key-dependent

terms as well as x will cancel. Thus we obtain the relation

$$\begin{aligned}
 p_1 + s_1^{(0)} + s_1^{(1)} + s_0^{(1)} + s_1^{(2)} + s_1^{(3)} + s_0^{(3)} \\
 + s_1^{(4)} + s_1^{(5)} + s_0^{(5)} + s_1^{(6)} + s_1^{(7)} + s_0^{(7)} + y_1 = 0.
 \end{aligned} \tag{9}$$

By the Piling-up Lemma, the bias of this linear approximation is estimated to be $81/131072$, or roughly $2^{-10.66}$. However, since the two relations combined are clearly not independent, it is important to verify that the bias behaves as predicted. Computer verification shows that the actual bias of the relation is in fact $78/131072$, which is quite close to our estimate. It is quite curious that we can find such multi-round approximations that do not depend on *any* of the fixed key bits $k^{(i)}$. Thus it does not seem unlikely that one may find a biased linear relation over the full 80 rounds of Edon80, depending only on some of the dynamic cipher states $s^{(i)}$, given a linear relation (of sufficient bias and low weight) over the first 40 rounds.

If we also may assume that we have a 4-bit restriction on the key, the bias of our relations increases dramatically. For the relation in 11 variables (7 state bits, 2 key bits, 1 input and 1 output bit) with bias $23/256$ referred to in the previous section, we find that the corresponding 8-round relation in 16 variables (14 state bits, 1 input and 1 output bit) with bias $39/2048$ ($2^{-5.71}$).

3 Conclusion

We have examined some biases and structure in the Edon80 S-box, but are not able to find linear relations of sufficiently high bias and low weight to admit an efficient attack against the full Edon80 stream cipher. It may be interesting to look further at scaled-down versions of Edon80, to learn more about the cipher's security margin and how many rounds are really needed to make linear distinguishing attacks infeasible.

References

1. D. Gligoroski, S. Markovski, L. Kocarev, and M. Gusev. Edon80. eSTREAM - ECRYPT Stream Cipher Project, Report 2005/007, 2005.