Search for Related-key Differential Characteristics in DES-like ciphers

Alex Biryukov and Ivica Nikolić*

University of Luxembourg
{alex.biryukov,ivica.nikolic}uni.lu

Abstract. We present the first automatic search algorithms for the best related-key differential characteristics in DES-like ciphers. We show that instead of brute-forcing the space of all possible differences in the master key and the plaintext, it is computationally more efficient to try only a reduced set of input-output differences of three consecutive S-box layers. Based on this observation, we propose two search algorithms – the first explores Matsui's approach, while the second is divide-and-conquer technique. Using our algorithms, we find the probabilities (or the upper bounds on the probabilities) of the best related-key characteristics in DES, DESL, and s^2DES.

Keywords: Cryptanalysis tool, automatic search, differential characteristic, related-key attack, DES.

1 Introduction

The Data Encryption Standard (DES) [8], adopted by the U.S. National Bureau of Standards in 1977, was a block cipher standard for several decades. Some of the design principles of DES were fully understood by the public only after the first cryptanalysis presented by Biham and Shamir [2]. They introduced the idea of differential analysis and differential characteristics, and showed that if one encrypts with DES a pair of plaintexts with a specific XOR difference, then the pair of corresponding ciphertexts will have some predictable difference with a probability higher than expected.

In [7] Matsui showed that the differential characteristics found by Biham and Shamir were indeed the best, i.e. they have the highest probability among all characteristics. He was able to prove this fact by running a full search on the space of all possible characteristics, using a special algorithm that speeds up the search. Matsui's algorithm was adopted and applied for search of the best characteristics in LOKI and s^2DES [10], Twofish [9], FEAL [1], and others. In all of these cases, the search was targeting only single-key characteristics, i.e. the characteristics that have a difference in the plaintext, but not in the master

* This author is supported by the Fonds National de la Recherche Luxembourg grant TR-PHD-BFR07-031.

key. Biryukov and Nikolić in [4] showed that Matsui's idea indeed can be used to build a search algorithm that finds the best related-key characteristics (the difference can be in the key as well as in the ciphertext) for some classes of byte-oriented ciphers. To the best of our knowledge, there are no published results on a search for related-key characteristics in any bit-oriented cipher.

Our contribution. We present algorithms for finding the best (with the highest probability) round-reduced related-key differential characteristics in DES and DES-like ciphers. We show that instead of trying all differences in the key and in the plaintext, which would result in a search space of size 2^{120}, it is computationally more efficient to try only a reduced set of input-output differences of three consecutive S-boxes layers. Based on this observation, we are able to propose two algorithms for automatic search of related-key differential characteristics in DES-like ciphers – the first is based on Matsui's approach, while the second is in line with the technique of divide-and-conquer. We apply our algorithms to DES, DESL [6], and s^2DES [5] and find either the probabilities of the best round-reduced related-key differential characteristics, or the upper bounds on these probabilities. Interestingly, although for lower number of rounds these probabilities are much higher than in the case of single-key characteristics, for higher number of rounds, the best characteristics are single-key characteristics. We obtain an interesting result regarding DES. By providing the probability of the best related-key characteristic on 13 rounds, we show that Biham-Shamir attack cannot be improved if one uses related-key characteristic (instead of single-key). Moreover, the low probabilities of the best related-key characteristics on higher rounds indicate that NSA did not introduce any weakness (or trapdoor) in the key schedule of DES with regard to differential attacks. Although in this paper we apply our algorithms only to the DES-like ciphers, we believe that our approaches can be used as well to search for high probability related-key differential characteristics in any bit-oriented ciphers with linear key schedule.

2 Description of DES-like Block Ciphers

DES [8] is 64-bit block cipher with 56-bit key[1]. It is 16-round Feistel cipher with additional permutations IP, IP^{-1} at the beginning and at the end. The 64-bit plaintext, after the application of the initial permutation IP is divided into two halves L_0 and R_0 - each half has 32 bits. Then, the halves are updated 16 times with the round function:

$$L_i = R_{i-1}$$
$$R_i = L_{i-1} \oplus f(R_{i-1}, K_i),$$

where $i = 1, \ldots, 16$ and K_i are 48-bit round keys, obtained from the initial key K with some linear transforms (rotations that depend on the round number and bit selection function PC-2). The ciphertext is defined as $IP^{-1}(R_{16}||L_{16})$.

[1] Officially, the key has 64 bits, but 8 bits are only used to check the parity, and then discarded.

The round function $f(R, K_i)$ takes 32-bit state R and 48-bit round key K_i and produces 32-bit output. First it expands the 32-bit value of R to 48 bits with the linear function E and then it XORs the values of $E(R)$ and K_i to produce some intermediate result, which we further denote as f_i. This 48-bit value is divided into 8 six-bit values, and each of these values goes through a separate 6x4 S-box. Finally, the 32-bit output of the S-boxes goes through a bit permutation P and the output $\tilde{f}_i$ of the round function is produced.

The DES-like block ciphers DESL [6] and s^2DES [5], differ from DES only in the definition of the S-boxes and the initial and final permutations. Since these permutations have no cryptographic values, we can assume that the only difference among the ciphers of the DES-family is in the S-boxes.

3 Automatic Search for Related-key Differential Characteristics in DES-like ciphers

The best characteristic on r rounds, i.e. the best r round-reduced characteristic, is the one that has the highest probability among all characteristics on r rounds of the cipher. In this section we propose two methods for building efficient automatic search algorithms for finding the best round-reduced related-key differential characteristics in DES-like ciphers. When constructing these algorithms, the main problem that has to be tackled is how to deal with the enormous search space. There are 64 bits in the state and 56 bits in the key, hence in total there are 2^{120} starting values for differential characteristics. However, in general, this number can be reduced significantly. Our first method is based on Matsui's search tool applied for finding the best single-key round-reduced characteristics in DES. The second method, which we call *the split* approach, can be used when Matsui's approach fails – to find characteristics on high number of rounds when not all the characteristics on lower number of rounds are known.

Due to the complementation property of DES, there are related-key characteristics (including round-reduced) that hold with probability 1. Further, we do not consider these characteristics.

Considering the different rotation amounts in the key-schedule, the probability of the best round-reduced related-key characteristic depends on the rounds covered by the characteristic. For example, the best 5-round related-key characteristic covering rounds 0-4, can have different probability from the best characteristic that covers rounds 1-5. The best related-key characteristics in our paper, always cover the last rounds, e.g. the characteristic on 7 rounds, covers the rounds 9-15.

3.1 Matsui's Approach for Single-key Characteristics

The search for the best single-key differential characteristics in DES was successfully performed by Matsui in [7]. Note that even in this case, when there is no difference in the key, the search space is rather large – 2^{64} starting differences.

However, Matsui presented several useful approaches how to deal with a large number of starting differences and how to significantly reduce the search space.

A naive approach to search for the best n-round characteristic would be to try all possible starting differences in the plaintext and try to extend each of them to n rounds. The non-linearity of the S-boxes will introduce branching, and a k-round characteristic ($k < n$) is extended for an additional round only if its probability is higher than the probability P_n^* of some known characteristic on n rounds.

Matsui's approach on the other hand, cuts out a large number of round-reduced characteristics in the early stage. Given the probabilities $\overline{P_1}, \ldots, \overline{P_{n-1}}$ of the best characteristics on the first $n-1$ rounds, and some estimate[2] P_n^* for the probability of the characteristic on n rounds, the algorithm produces the best characteristic on n rounds. Hence, the attacker can sequentially produce, starting from 1 or 2-round reduced, characteristics on all rounds of DES. In short, the attacker, as in the naive approach, tries all possible starting differences[3]. For each of them he produces 1-round characteristic (there can be many one-round characteristics, and the following procedure is repeated for each of them) that holds with probability P_1. Then, he tries to extend it to two rounds only if $P_1 \cdot \overline{P_{n-1}} > P_n^*$. This is because in order to extend 1-round characteristic to n rounds, one should use an additional $(n-1)$-round characteristic. Since the best one has probability $\overline{P_{n-1}}$, the total probability of the n-round characteristic will be at most $P_1 \cdot \overline{P_{n-1}}$ and this value should be better than the probability P_n^* of the best known characteristic on n-rounds. Similarly, if the attacker has built k-round characteristic with probability P_k than he tries to extended for an additional round only if $P_k \cdot \overline{P_{n-k}} > P_n^*$. Note that in the naive approach, the attacker only checks if $P_k > P_n^*$. Therefore, Matsui's approach stops the extension of many round-reduced characteristics and that way speeds up the search.

Now let us take a closer look how to reduce the number of possible starting differences. Interestingly, the same approach as above can be used. First note that a characteristic on the first two rounds (assuming this 2-round characteristic is part of the best n-round characteristic) has a probability P_2 such that $P_2 < P_n^*/\overline{P_{n-2}}$. The following observation is used to explore this property of 2-round single-key characteristics.

Observation 1 *Given the input and the output differences* $(\Delta f_1, \Delta\tilde{f}_1), (\Delta f_2, \Delta\tilde{f}_2)$ *of the S-boxes layers in the first two rounds, one can find the difference in the plaintext* ΔP *and the difference* $(\Delta L_2, \Delta R_2)$ *in state at the beginning of the third round.*

Proof. From the Feistel construction it leads that $\Delta R_0 = E^{-1}(\Delta f_1)$ and $\Delta L_0 = \Delta\tilde{f}_1 \oplus E^{-1}(\Delta f_2)$. Then the difference ΔP in the plaintext is $\Delta P =$

[2] For example, the attacker can use the probability of the already known characteristic on n rounds as an estimate.

[3] We will see later, that this requirement can be omitted.

$IP^{-1}(\Delta R_0||\Delta L_0)$. Similarly, for the difference at the beginning of the third round we get $\Delta R_2 = \Delta R_0 \oplus \Delta\tilde{f}_2$ and $\Delta L_2 = \Delta L_0 \oplus E^{-1}(\Delta f_2)\Box$.

Therefore, instead of fixing all possible differences ΔP in the plaintext one can fix only the input and the output differences to the S-boxes in rounds 1,2. But, since the active S-boxes of the first round have to hold with a probability of at least $P_n^*/\overline{P_{n-1}}$, and in the first and the second round with at least $P_n^*/\overline{P_{n-2}}$, the number of 2-round characteristics is significantly reduced. For each such characteristic, one can proceed with Matsui's technique, and try to extend it to n-rounds (since the difference at the beginning of round 3 is fixed).

3.2 Applying Matsui's Approach for Related-key Characteristics

One can easily reconstruct Matsui's algorithm to search for related-key characteristics. Note that for a fixed difference in the key, the algorithm still works and it finds the best characteristic with this specific difference. However, since the key has 56 bits, this search has to be repeated 2^{56} times and hence this naive approach is not feasible. We can still run a so-called *limited search* for related-key characteristics, by allowing low Hamming difference in the key. For example, to find the best characteristic that has at most 2-bit difference in the key, we have to rerun Matsui's algorithm $1 + C_{56}^1 + C_{56}^2 = 1597$ times.

Indeed, finding the best related-key characteristic using Matsui's approach can be done efficiently. We only have to find a way to efficiently limit the number of possible differences in the key and in the plaintext. We want to reduce the search space, yet to perform a full search of all possible related-key differential characteristics. The following observation can be used for that purpose.

Observation 2 *Given the input and the output differences* $(\Delta f_1, \Delta\tilde{f}_1)$, $(\Delta f_2, \Delta\tilde{f}_2)$, $(\Delta f_3, \Delta\tilde{f}_3)$ *of the S-boxes layers in the first three rounds, one can find the difference in the plaintext* ΔP*, the difference* $(\Delta L_3, \Delta R_3)$ *in state at the beginning of the fourth round, and all* 2^8 *values for the difference* ΔK *in the master key.*

Proof. Again we use the property of the Feistel construction and the linearity of the key schedule. From the definition of DES we get:

$$\Delta f_1 = E(\Delta R_0) \oplus \Delta K_1 \tag{1}$$

$$\Delta f_3 = E(\Delta R_0 \oplus \Delta\tilde{f}_2) \oplus \Delta K_3 \tag{2}$$

Since E is linear, we get:

$$\Delta K_1 \oplus \Delta K_3 = \Delta f_1 \oplus \Delta f_3 \oplus E(\Delta\tilde{f}_2)$$

The key schedule is linear, and both K_1 and K_3 are obtained from the master K with some linear transformation. Therefore $\Delta K_1 \oplus \Delta K_3$ can be expressed as $\mathcal{L}(\Delta K)$, where $\mathcal{L}$ is a linear transformation. On the other hand, the input-output differences of the S-boxes are given, and therefore, the value $V = \Delta f_1 \oplus$

$\Delta f_3 \oplus E(\Delta \tilde{f}_2)$ is known. Hence, the master key difference ΔK can be found as $\Delta K = \mathcal{L}^{-1}(V)$. However, the key is 56 bits, while V only 48 bits. Therefore we get an underdefined system of linear equations with 2^8 solutions. If we fix a particular solution for the system, and thereby the difference in the key K, we can easily find $\Delta K_1, \Delta K_3$ (and ΔK_2). Then $\Delta R_0 = E^{-1}(\Delta f_1 \oplus \Delta K_1)$ and $\Delta L_0 = E^{-1}(\Delta f_2 \oplus K_2) \oplus \Delta \tilde{f}_1$. Similarly can be found the differences $\Delta L_3, \Delta R_3$ $\square$.

The above observation clearly indicates how to reduce the search space. Instead of trying all possible differences in the key K and running Matsui's algorithm for each of them, one should only fix the input and the output differences to the S-box layers in the first three rounds. Due to restrictions on the probability, all the active S-boxes in first, in the first and second, and in the first, second and third round, should have a combined probability of at least $P_n^*/\overline{P_{n-1}}, P_n^*/\overline{P_{n-2}}, P_n^*/\overline{P_{n-3}}$, respectively. Once the active S-boxes for the first three rounds are fixed, one can easily find all 2^8 candidates for the difference in the master key and the difference in the state after the third round and hence produce 3-round differential characteristic with a fixed difference in the master key. Further, Matsui's approach can be used, and this characteristic can be extended to any number of rounds. The pseudo-code of the whole algorithm is given at Alg. 1.1.

On the complexity and optimization of the search. Calculating the exact time complexity of the whole search is complex and probably impossible. However, some estimate can be given, under a certain assumption. Our experiments indicate that once the difference in the state (after the third round) and in the key is fixed, extending the characteristic to n rounds becomes fairly easy and computationally cheap task. The main complexity lies in generating all 3-round related-key characteristics that have a certain probability. More precisely, from observation 2 it follows that one should generate all active S-boxes in the first round that hold with a combined probability P_1 of not less than $P_n^*/\overline{P_{n-1}}$, then all active S-boxes in the second round with a combined probability not less than $P_n^*/(\overline{P_{n-2}} \cdot P_1)$ and all active in the third round with probability of not less than $P_n^*/(\overline{P_{n-3}} \cdot P_2)$ (where P_2 is the probability of the active S-boxes in the first two rounds). Therefore, the number of all 3-round related-key characteristics depends only on the values $P_n^*/\overline{P_{n-1}}, P_n^*/\overline{P_{n-2}}$ and $P_n^*/\overline{P_{n-3}}$ – higher the values, less characteristics exist, and the search is faster.

The complexity of creating all these 3-round characteristics is not the same (or proportional) as the number of such characteristics. This comes from the fact that the linear transform E is not a surjective, since it has 32-bit input and 48-bit output. For example, after ΔK_1 is found (see the proof of the observation 2), the value $\Delta R_0 = E^{-1}(\Delta f_1 \oplus \Delta K_1)$ exists only with a probability 2^{-16}. Similar holds for ΔL_0. Hence, the optimal strategy for creating the 3-round characteristics would be to:

1. Fix the probabilities of the first four active S-boxes in the first and the third round and all the active S-boxes of the second round (that have the above

Algorithm 1.1. Search for RK differential characteristic

FullSearch()
{
// The first three rounds
for all $\Delta f_1 \rightarrow \Delta\tilde{f}_1 | P(\Delta f_1 \rightarrow \Delta\tilde{f}_1)\overline{P_{n-1}} > P_n^*$ **do**
  for all $\Delta f_2 \rightarrow \Delta\tilde{f}_2 | P(\Delta f_1 \rightarrow \Delta\tilde{f}_1)P(\Delta f_2 \rightarrow \Delta\tilde{f}_2)\overline{P_{n-2}} > P_n^*$ **do**
    for all $\Delta f_3 \rightarrow \Delta\tilde{f}_3 | P(\Delta f_1 \rightarrow \Delta\tilde{f}_1)P(\Delta f_2 \rightarrow \Delta\tilde{f}_2)P(\Delta f_3 \rightarrow \Delta\tilde{f}_3)\overline{P_{n-3}} > P_n^*$ **do**
      $V = \Delta f_1 \oplus \Delta f_3 \oplus E(\Delta\tilde{f}_2)$
      for all $\Delta K | \mathcal{L}(\Delta K) = V$ **do**
        $\Delta K_1 = PC2(rot(\Delta K, 1))$
        $\Delta K_2 = PC2(rot(\Delta K, 2))$
        if $E^{-1}(\Delta K_1 \oplus \Delta f_1)$ and $E^{-1}(\Delta K_2 \oplus \Delta f_2)$ **then**
          $\Delta R_0 = E^{-1}(\Delta K_1 \oplus \Delta f_1)$
          $\Delta L_0 = E^{-1}(\Delta K_2 \oplus \Delta f_2) \oplus \Delta\tilde{f}_1$
          $\Delta R_3 = \Delta L_0 \oplus \Delta\tilde{f}_1\Delta\tilde{f}_3$
          $\Delta L_3 = \Delta R_0 \oplus \Delta\tilde{f}_2$
          Call NextRound(ΔL_3, ΔR_3, ΔK, $P(\Delta f_1 \rightarrow \Delta\tilde{f}_1)P(\Delta f_2 \rightarrow \Delta\tilde{f}_2)P(\Delta f_3 \rightarrow \Delta\tilde{f}_3)$, 4)
        end if
      end for
    end for
  end for
end for
}

NextRound($\Delta L, \Delta R, \Delta K, p, round$)
{
$\Delta K_r = PC2(rot(\Delta K, round))$
$\Delta f = \Delta K_r \oplus E(\Delta R)$
for all $\Delta f \rightarrow \Delta\tilde{f} | P(\Delta f \rightarrow \Delta\tilde{f}) \cdot p \cdot \overline{P_{n-round}} > P_n^*$ **do**
  $\Delta L_{new} = \Delta R$
  $\Delta R_{new} = \Delta L \oplus \Delta\tilde{f}$
  if $round == n$ **then**
    if $P(\Delta f \rightarrow \Delta\tilde{f}) \cdot p > P_n^*$ **then**
      $P_n^* = P(\Delta f \rightarrow \Delta\tilde{f}) \cdot p$
    end if
  else
    Call NextRound($\Delta L_{new}, \Delta R_{new}, \Delta K, P(\Delta f \rightarrow \Delta\tilde{f}) \cdot p, round + 1$)
  end if
end for
}

limitations), without fixing the exact input-output differences. This can be done by fixing only the possible values from the difference distribution tables of the S-boxes.

2. Fix the input differences to the four S-boxes of round 1,3, and the output differences of the S-boxes of round 2 (that correspond to the previously fixed distribution values).
3. Find 28 bits of ΔK, then find 28 bits of ΔK_1 and check if there exist preimage of 24 bits of $\Delta f_1 \oplus \Delta K_1$ for E. This can be done, since the left and the right 28-bit halves of the key are independent.
4. If exists, fix the probabilities of the last four active S-boxes in the first and the third round.
5. Fix the input differences to these 8 S-boxes.
6. Find the rest 28 bits of ΔK, then of ΔK_1 and check if there exist preimage of last 24 bits of $\Delta f_1 \oplus \Delta K_1$ for E.
7. If exists, find ΔK_2, fix the input difference to the S-boxes in the second round and check if there exist a preimage of $\Delta f_2 \oplus K_2$ for E.
8. If exists, fix the output differences for the S-boxes of round 3 (it is not necessary to fix the outputs of S-boxes of round 1).

Although we cannot give a precise estimate for the complexity of creating all 3-round characteristics, we can give such estimates for some particular fixed values of P_n^*,$\overline{P_{n-1}}$,$\overline{P_{n-2}}$, and $\overline{P_{n-2}}$. For example, when $P_n^*/\overline{P_{n-1}} = 2^{-3}, P_n^*/\overline{P_{n-2}} = 2^{-6}, P_n^*/\overline{P_{n-3}} = 2^{-9}$, then steps 1-8 are repeated $2^{16.7}$, $2^{28.9}$, $2^{32.9}$, $2^{27.3}$, $2^{30.8}$, $2^{34.9}$, $2^{27.6}$, $2^{20.3}$ times, respectively, leading to a total complexity of around 2^{35}. On the other hand, when $P_n^*/\overline{P_{n-1}} = 2^{-3}, P_n^*/\overline{P_{n-2}} = 2^{-7}, P_n^*/\overline{P_{n-3}} = 2^{-10}$, then steps 1-8 are repeated $2^{18.7}$, $2^{32.4}$, $2^{36.4}$, $2^{30.8}$, $2^{34.3}$, $2^{38.4}$, $2^{30.9}$, $2^{22.6}$ times, respectively, and hence the complexity is around 2^{39}, while there exist around $2^{22.6}$ (step 8) good 3-round related-key characteristics.

3.3 The Split Approach

To build the best n-round characteristic Matsui's approach requires first to build the best characteristics on $1, 2, \ldots, n-1$ rounds because it uses the probabilities of these characteristics. One may be able to skip building the characteristics on some rounds and to assume that they have the same probability as the characteristic on lower number of rounds. Under this assumption, the algorithm still works and finds the best characteristic on n rounds, however the time complexity usually suffers significantly.

Avoiding building all round-reduced characteristics can be done with a different approach. Let us assume we search for characteristic on n rounds that has a probability of at least P_n^*. This n-round characteristic can be seen as a concatenation of two $n/2$-round characteristics, with a combined probability of at least P_n^*. Therefore, one of these two characteristics has a probability of at least $\sqrt{P_n^*}$. Indeed we can split the n-round characteristic on any (reasonable) number of k characteristics, each on n/k rounds, and claim that at least one of them has a probability of $\sqrt[k]{P_n^*}$.

Now, let us assume that $n = 3k$, and the n-round characteristic has been split into k three-round characteristics. One of these characteristics (we do not know exactly which), has to have a probability of at least $\sqrt[k]{P_n^*}$. Since it is

on three rounds, and it has a bound on its probability, we can use our previous method (observation 2), to build all such characteristics. However, unlike in Matsui's approach, where each of the three rounds has some bound on probability, now we build 3-round characteristics that only have the bound on the combined probability (of all three rounds). Once we have built all such the 3-round characteristics we try to extend them to n rounds (recall that if the difference in the state and in the key is fixed, then it is easy to extend it to more rounds – the difficulty lies in creating all such 3-round characteristics). Interestingly, when extending the three round characteristics, we can use the bounds from Matsui's approach.

For example, let us assume we want to build a characteristic on 9 rounds with a probability at least 2^{-24}. Then we know that one of the three 3-round characteristics has a probability of at least 2^{-8}. First we assume that this is the characteristic on the first three rounds. We build all first 3-round characteristics with probability at least 2^{-8}, i.e. $P_3 \geq 2^{-8}$, and then try to extend them 6 rounds forward, thus obtaining a characteristic on 9 rounds. If we have the probabilities $\overline{P_1}, \ldots, \overline{P_6}$ for the best characteristics on the last 6 rounds, then for rounds 4-9, we can use Matsui's approach, e.g. for 4 rounds we take only those with P_4 such that $P_4 \cdot \overline{P_5} \geq 2^{-24}$, for 5 rounds $P_5 \cdot \overline{P_4} \geq 2^{-24}$, etc. If we do not have the best probabilities than for each round i ($i \geq 4$) we only check if $P_i \geq 2^{-24}$. Then we assume the characteristic on rounds 4-6 has a probability of at least 2^{-8}. Again, we build all 3-round characteristics with at most 2^{-8} and extend them three rounds forward and three backwards (by using Matsui's bounds). Finally, we assume this is the 3-round characteristic on the last three rounds (7-9). We build all such characteristics and extend them 6 rounds backwards (again we can use Matsui's bounds if we have the best probabilities for the first 6 rounds). Among all 9-round characteristics we have produced in these three iterations, we take the one with the highest probability. If such characteristic exist than it is the best characteristic on 9 rounds and it has a probability at least 2^{-24}. If it does not exist then it means all the characteristics on 9 rounds have probability lower than 2^{-24}.

What is the real advantage of this approach compared to related-key Matsui's approach? To find this out, we have to compare the number of possible 3-round related-key characteristic built in the two approaches. In Matsui's algorithm, this number depends on the values $P_n^*/\overline{P_{n-1}}$, $P_n^*/\overline{P_{n-2}}$ and $P_n^*/\overline{P_{n-3}}$, while in the split approach, the number depends only on P_n^*. Hence, when the probabilities $\overline{P_{n-1}}$, $\overline{P_{n-2}}$, $\overline{P_{n-3}}$ are really high, then it is computationally cheaper to build the n-round characteristic with the split approach.

4 The Case of DES

The notion of (single-key) differentials and differential characteristics was introduced in the seminal paper of Biham and Shamir [2] on cryptanalysis of DES, where the authors presented characteristic on 15 rounds of DES with a probability higher than 2^{-56}. Later in [3], the authors used 13-round characteristic

to give the first attack on all 16 rounds of DES. By performing a full search, Matsui [7] has shown that the characteristics found by Biham and Shamir were actually the best round-reduced single-key characteristics for DES. It is well known that S-boxes and the permutation used in the round function of DES are very carefully chosen to avoid single-key differential cryptanalysis and even subtle changes in them can weaken the cipher [3]. Our study of related-key attacks on DES is motivated by the fact that differences in the subkeys could violate some of the design principles and this could lead to new attacks on DES.

We would like to run a full search of the space of all related-key differential characteristics in DES by using the approaches of the previous section. We start with the related-key version of Matsui's algorithm and try to find the best related-key characteristics on as many rounds as possible. Although our search will always find the best characteristics, we should keep in mind that we have a limited computational power. For example, if we try to find the best n-round related-key characteristic that holds with a probability at least P_n^*, then the time complexity of the search mostly depends on the probability $\overline{P_{n-3}}$ of the best characteristics on $(n-3)$ rounds (but also depends on $\overline{P_{n-1}}, \overline{P_{n-2}}$). Our experimental results show that when $P_n^*/\overline{P_{n-3}} < 2^{-12} \sim 2^{-14}$ we do not have the resources to perform the search, hence if for some n this holds, then we will switch to the split approach and continue further with this approach. Note that even in the case of single-key characteristics a similar limitation holds when for some n the ratio $P_n^*/\overline{P_{n-2}}$ is too low.

We start the search by finding the best related-key characteristic on 3 rounds (we assume that $\overline{P_0} = \overline{P_1} = \overline{P_2} = 1$). We fix P_3^* (the probability of the best related-key 3-round characteristic) to 2^{-1} and then gradually decrease by a factor of 2^{-1} if we do not find a characteristic that holds with this probability. There is always a lower bound on this probability – the case of the single-key characteristic (our tool does not make distinction between these two cases, and searches for both). Hence, we can be sure that P_3^* cannot be lower than 2^{-4} (this is the probability of the best single-key characteristic on 3 rounds). Having found the highest P_3^*, we fix $\overline{P_3} = P_3^*$, and then search for $\overline{P_4}$. We fix P_4^* to $\overline{P_3}$, i.e. we assume that the characteristic on 4-rounds has the same probability as the best characteristic on 3 rounds, and then gradually decrease this probability by a factor 2^{-1} each time when we cannot find 4-round characteristic with such probability. Up to $\overline{P_6}$ we could easily perform the search. However, when searching for $\overline{P_7}$ we could not find anything even when P_7^* was set up to 2^{-18}. We knew that $\overline{P_7}$ could not be lower than $2^{-23.6}$ (the probability of the single-key characteristic on 7 rounds), however if we set $P_7^* = 2^{-23.6}$, then $P_7^*/\overline{P_4} = 2^{-19}$ which is lower than our maximal computational limit of $2^{-12} \sim 2^{-14}$. Therefore, we switched to the split approach for finding the best 7-round related-key characteristic. We started with all possible 3.5-round characteristic (with the first 3.5 rounds and the last 3.5 rounds) with probability of at least 2^{-11} and tried to extend it to 7 rounds, thus we allowed a probability of 2^{-22}. The split approach found that the best related-key characteristic on 7 rounds has a probability of $2^{-20.38}$.

The results of the split search on 7 rounds can be used to find if 8-round characteristic with 2^{-22} exist, which in our case was negative. If we try to apply the related-key Matsui's approach for 8 rounds and allow $P_8^* = 2^{-22}$, then $P_8^*/\overline{P_5} = 2^{-22}/2^{-7.6} = 2^{-14.4}$, which is low. Hence, for 8 rounds we could not use neither Matsui's nor the split approach. However, we noted that the best characteristics of the first 7 rounds have a difference only in a few bits of the key. Hence, we ran a limited search for 8-round characteristic by allowing only a few bit difference in the key. The limited search gave us a characteristic with a probability $2^{-29.75}$ – better than the best single-key characteristic with $2^{30.8}$.

For higher rounds, the related-key Matsui's approach could not work because of the low probabilities ($P_n^*/\overline{P_{n-3}} < 2^{-12} \sim 2^{-14}$). However, if we assume that the 8-round characteristic found by the limited Matsui's approach is the best, then we can still run related-key Matsui's algorithm for the characteristic on 11 rounds. We found that if this holds, then the best related-key characteristics on 11 rounds is the best single-key characteristics.

For finding the best related-key characteristics on 9, 12, and 13 rounds we used our split approach. For 9 rounds, we allowed the 3-round characteristics to have at least $2^{-10.55}$ (because $(2^{-10.55})^3 = 2^{-31.65}$ and the best single-key on 9 rounds has $2^{-31.48}$). The search found that the best 9-round related-key characteristic is the best single-key characteristic. For 12 and 13 rounds, we allowed the starting 3-round characteristics with probability at least $2^{-11.85}$ (because $(2^{-11.85})^4 = 2^{-47.4}$ and the best single-key on 13 rounds has $2^{-47.22}$). Again, we obtained similar results – the best related-key characteristics on 12 and 13 rounds have no difference in the key, i.e. they are the single-key characteristics.

The result for the 13-round[4] related-key characteristic is especially interesting since Biham-Shamir analysis uses it for the attack on the whole DES. This means that *if the attacker uses related-key characteristics, he cannot improve the complexity of Biham-Shamir attack.*

The summary of our findings is presented in Tbl. 1. The related-key characteristics for 7 and 8 rounds are given in the Appendix (Fig. 1, 2).

5 The Case of DESL

DESL [6] uses a single S-box instead of eight different S-boxes as in DES. This S-box has a special design criteria to discard high probability (single-key) differential characteristics. Indeed, our initial analysis for single-key differential characteristics in DESL confirmed this result. Moreover, we could not find the best single-key differential characteristics (using the original Matsui's tool) for DESL for higher rounds (the absence of the probabilities for the best round-reduced single-key differential characteristics in the submission paper of DESL [6] seems to confirm our findings). Therefore, even the original Matsui's tool cannot be used (it is infeasible) for finding single-key characteristics, when they hold with low probabilities.

[4] We rerun the search for characteristics that cover rounds 1 to 12.

Table 1. Comparison of the probabilities of the best round-reduced differential single-key and related-key characteristics for DES.

rounds	Single-key	Related-key	Method used
3	$2^{-4.0}$	2^{0}	RK Matsui's
4	$2^{-9.6}$	$2^{-4.61}$	RK Matsui's
5	$2^{-13.21}$	$2^{-7.83}$	RK Matsui's
6	$2^{-19.94}$	$2^{-12.92}$	RK Matsui's
7	$2^{-23.60}$	$2^{-20.38}$	Split
8	$2^{-30.48}$	$2^{-29.75} \leq \overline{P_8} < 2^{-22}$	Limited Matsui's
9	$2^{-31.48}$	$2^{-31.48}$	Split + Matsui's
10	$2^{-38.35}$	$\leq \overline{P_9}$	
11	$2^{-39.35}$	$2^{-39.35}$ if $\overline{P_8} = 2^{-29.75}$	RK Matsui's
12	$2^{-46.22}$	$2^{-46.22}$	Split + Matsui's
13	$2^{-47.22}$	$2^{-47.22}$	Split + Matsui's
14	$2^{-54.09}$	$\leq \overline{P_{13}}$	
15	$2^{-55.09}$	$2^{-55.09}$	RK Matsui's
16	$2^{-61.97}$	$\leq \overline{P_{15}}$	

Our related-key Matsui's search algorithm, however, did find the best related-key characteristics for up to 7 rounds. Interestingly, the probabilities of these related-key characteristics are higher in DESL, than in DES (see Tbl. 2). For more rounds, we used the split approach as well. Nonetheless, for these characteristics, we were able to find only the upper bounds on their probabilities. For example, for 9-round related-key characteristic we used the split approach with 3-round probability of 2^{-10}. After running the search for the first, middle, and third three rounds, the algorithm did not return any characteristic. This means, there are no related-key characteristics on 9 rounds with probability at least 2^{-30}. Similarly, we used the split approach for finding the upper bound on the probability of the best characteristics for 12-rounds, and the related-key Matsui's approach for the bounds on 10,13, and 15 rounds. Our findings are presented in Tbl. 2.

The related-key characteristics that we have found can be used to launch boomerang attacks on the round-reduced cipher. For example, we can launch a related-key boomerang attack on 12 rounds (from round 4 to round 15), with two characteristics on 6 rounds – the first on rounds 4-9, the second on 10-15. The probability of the first characteristic is $2^{-14.68}$ (it is lower because we consider rounds 4-9), while the probability of the second is $2^{-12.09}$. Therefore, the probability of the whole boomerang is $2^{-2 \cdot 14.68 - 2 \cdot 12.09} = 2^{-53.54}$.

Table 2. Probabilities of the best round-reduced related-key differential characteristics for DESL.

Round	Probability
3	2^{0}
4	$2^{-4.67}$
5	$2^{-7.24}$
6	$2^{-12.09}$
7	$2^{-19.95}$
8	$\leq \overline{P_7}$
9	$< 2^{-30}$
10	$< 2^{-31}$
11	$\leq \overline{P_{10}}$
12	$< 2^{-40}$
13	$< 2^{-41}$
14	$\leq \overline{P_{13}}$
15	$< 2^{-50}$
16	$< 2^{-51}$

6 The Case of s^2DES

Another variant of DES called s^2DES was proposed in [5]. The search for the best single-key differential characteristics in s^2DES was performed in [10]. For this purpose the authors used Matsui's tool. This analysis showed that the best round-reduced differential characteristics in s^2DES have higher probabilities than in DES.

We ran our search for related-key characteristics using only our related-key approached based on Matsui's algorithm. We noted that for each single-key characteristic on n-rounds, the value $\overline{P_n}/\overline{P_{n-3}}$ is at least $2^{-12.75}$ (for $n = 8$, see Tbl. 3), hence building all 3-round related-key characteristic might be feasible. However, the values $\overline{P_{n-3}}$ for different n could be updated, because they were the probabilities in the single-key scenario (the probability in the related-key scenario is not less than in the single-key). Indeed, the probabilities of the round-reduced related-key characteristics for the first 6 rounds, were higher than the probabilities of the single-key characteristics. This made $\overline{P_5}$ to be 2^{-8} instead of $2^{-9.22}$ as in the single-key case. Hence, for the related-key characteristic on 8 rounds, we had to allow $\overline{P_8}/\overline{P_5} = 2^{-22}/2^{-8} = 2^{-14}$ for the active S-boxes in the three rounds, instead of the previous $2^{-12.75}$. However, we were able to perform the search for this 7-round characteristic but with a significant computational cost – the search took around 3 weeks on 64 CPU cores.

After the sixth round, we found that all the best related-key characteristics have the same probability as the single-key (indeed they are single-key). The

probabilities of the best single and related-key round-reduced characteristics are given in Tbl. 3.

Table 3. Comparison of the probabilities of the best round-reduce differential single-key and related-key characteristics for s^2DES.

rounds	Single-key	Related-key
3	$2^{-4.39}$	2^{0}
4	$2^{-6.8}$	$2^{-5.19}$
5	$2^{-9.22}$	$2^{-8.0}$
6	$2^{-14.35}$	$2^{-12.61}$
7	$2^{-17.03}$	$2^{-17.03}$
8	$2^{-21.96}$	$2^{-21.96}$
9	$2^{-22.71}$	$2^{-22.71}$
10	$2^{-27.35}$	$2^{-27.35}$
11	$2^{-28.39}$	$2^{-28.39}$
12	$2^{-34.07}$	$2^{-34.07}$
13	$2^{-34.07}$	$2^{-34.07}$
14	$2^{-39.75}$	$2^{-39.75}$
15	$2^{-39.75}$	$2^{-39.75}$
16	$2^{-45.42}$	$2^{-45.42}$

7 Conclusions

We have presented the first algorithms for automatic search of the best round-reduced related-key differential characteristics in DES-like family of ciphers, DES, DESL, and s^2DES. We have shown that there is no significant difference between the probabilities of the best related-key and the best single-key characteristics on higher number of rounds of DES, and thus, the key schedule of DES has no notable weakness regarding differential attacks.

We believe our algorithms can be applied to similar 64-bit state and 64-bit key bit-oriented ciphers with linear key schedule. Moreover, our approaches can be used to search for high probability (up to 2^{-20}) related-key differential characteristics in any bit oriented ciphers with linear key schedule.

References

1. K. Aoki, K. Kobayashi, and S. Moriai. Best differential characteristic search of FEAL. In E. Biham, editor, *FSE*, volume 1267 of *Lecture Notes in Computer Science*, pages 41–53. Springer, 1997.
2. E. Biham and A. Shamir. Differential cryptanalysis of DES-like cryptosystems. *J. Cryptology*, 4(1):3–72, 1991.

3. E. Biham and A. Shamir. Differential cryptanalysis of the full 16-round DES. In E. F. Brickell, editor, *CRYPTO*, volume 740 of *Lecture Notes in Computer Science*, pages 487–496. Springer, 1992.
4. A. Biryukov and I. Nikolić. Automatic Search for Related-Key Differential Characteristics in Byte-Oriented Block Ciphers: Application to AES, Camellia, Khazad and Others. In H. Gilbert, editor, *EUROCRYPT*, volume 6110 of *Lecture Notes in Computer Science*, pages 322–344. Springer, 2010.
5. K. Kim. Construction of DES-like S-boxes based on boolean functions satisfyieng the SAC. In H. Imai, R. L. Rivest, and T. Matsumoto, editors, *ASIACRYPT*, volume 739 of *Lecture Notes in Computer Science*, pages 59–72. Springer, 1991.
6. G. Leander, C. Paar, A. Poschmann, and K. Schramm. New lightweight DES variants. In A. Biryukov, editor, *FSE*, volume 4593 of *Lecture Notes in Computer Science*, pages 196–210. Springer, 2007.
7. M. Matsui. On correlation between the order of S-boxes and the strength of DES. In A. D. Santis, editor, *EUROCRYPT*, volume 950 of *Lecture Notes in Computer Science*, pages 366–375. Springer, 1994.
8. National Bureau of Standards. Data Encryption Standard. U.S. Department of Commerce, FIPS pub. 46, January 1977.
9. B. Schneier, J. Kelsey, D. Whiting, D. Wagner, C. Hall, and N. Ferguson. *The Twofish encryption algorithm: a 128-bit block cipher*. John Wiley & Sons, Inc., New York, NY, USA, 1999.
10. T. Tokita, T. Sorimachi, and M. Matsui. Linear cryptanalysis of LOKI and s^2DES. In J. Pieprzyk and R. Safavi-Naini, editors, *ASIACRYPT*, volume 917 of *Lecture Notes in Computer Science*, pages 293–303. Springer, 1994.

A Related-key Characteristics for DES

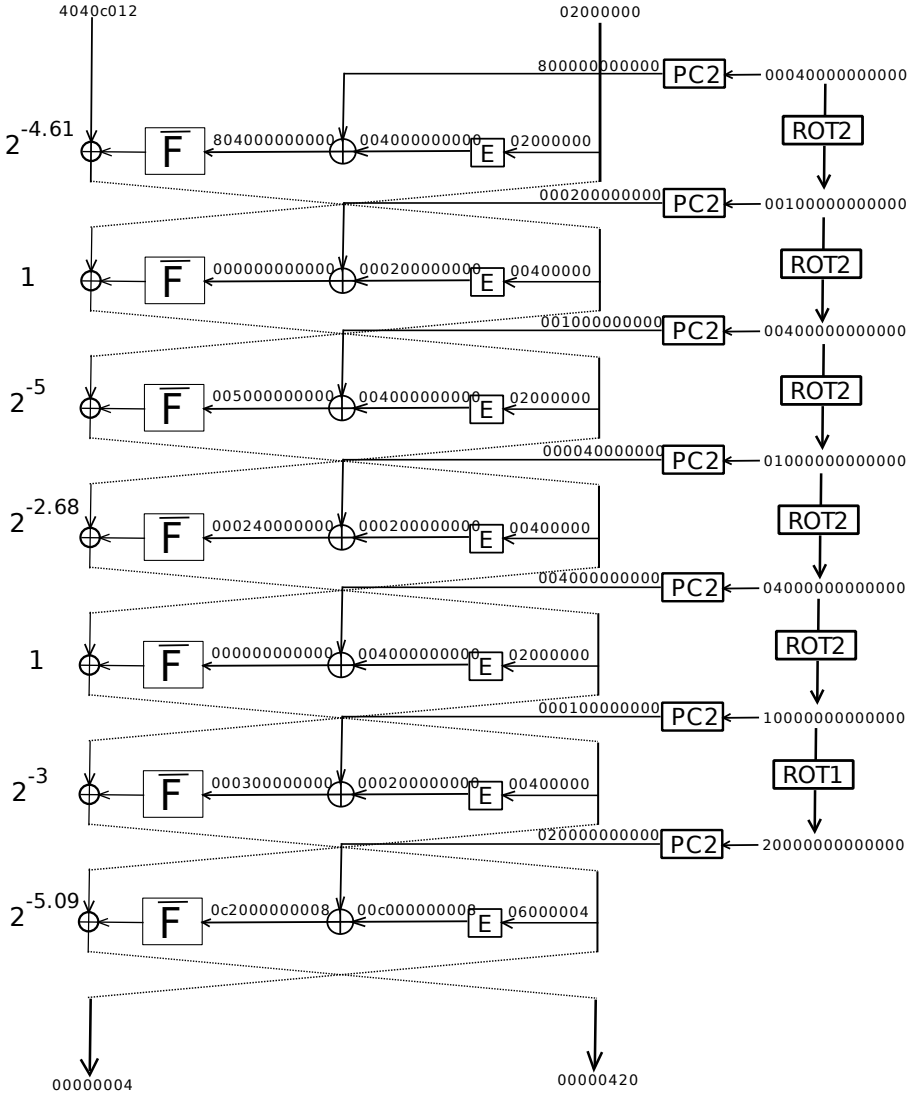

Fig. 1. The best related-key differential characteristic (with probability $2^{-20.38}$) on the last 7 rounds of DES.

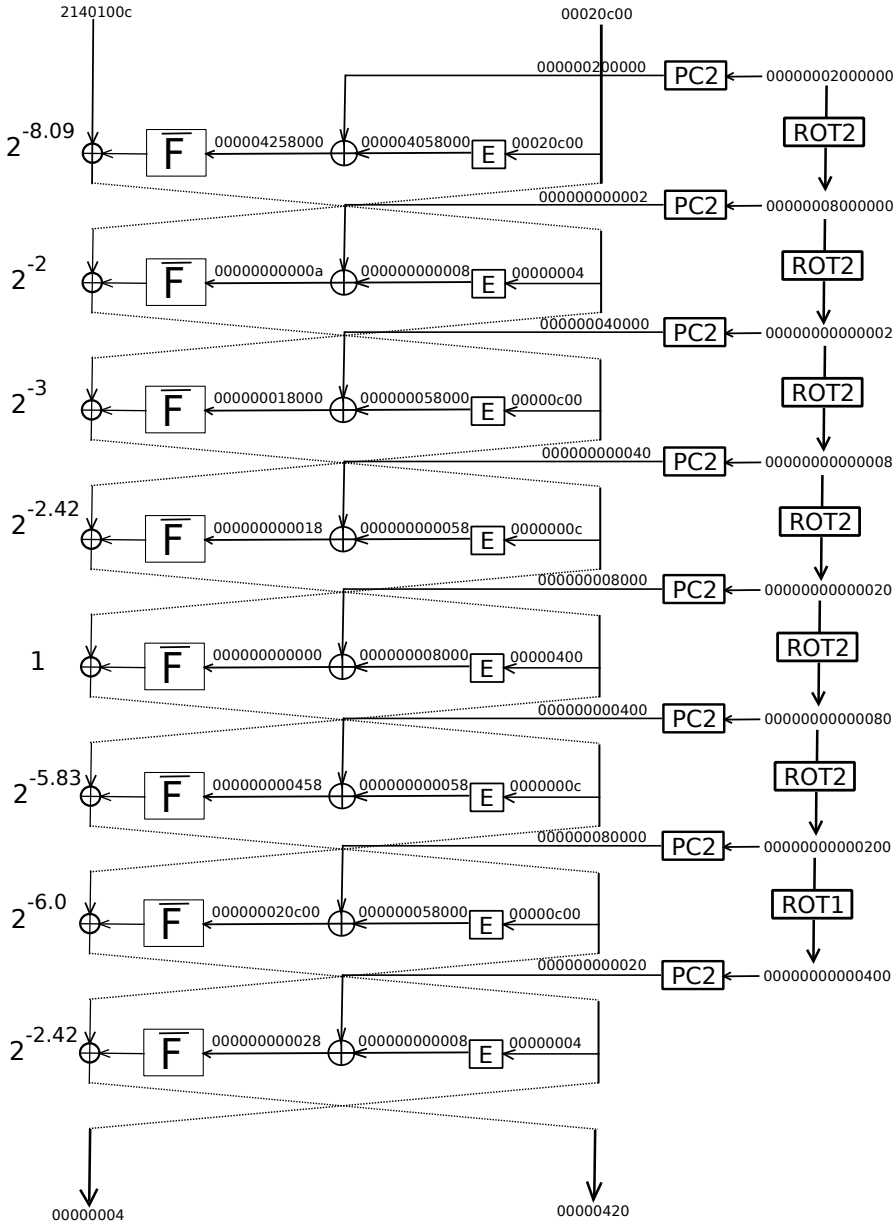

Fig. 2. Related-key differential characteristic (with probability $2^{-29.75}$) on the last 8 rounds of DES.