Algebraic Techniques in Differential Cryptanalysis

Martin Albrecht and Carlos Cid

Information Security Group,
Royal Holloway, University of London

FSE 2009, Leuven, 24.02.2009

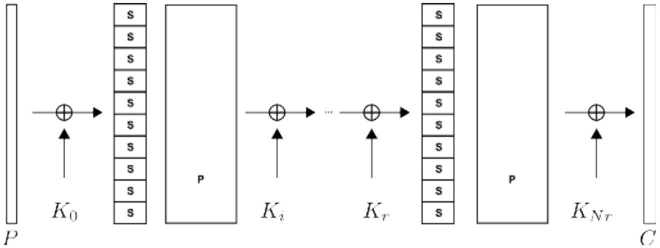
Outline

Outline

The Blockcipher PRESENT

Information Security Group

PRESENT [2] was proposed by Bogdanov et al. at CHES 2007.



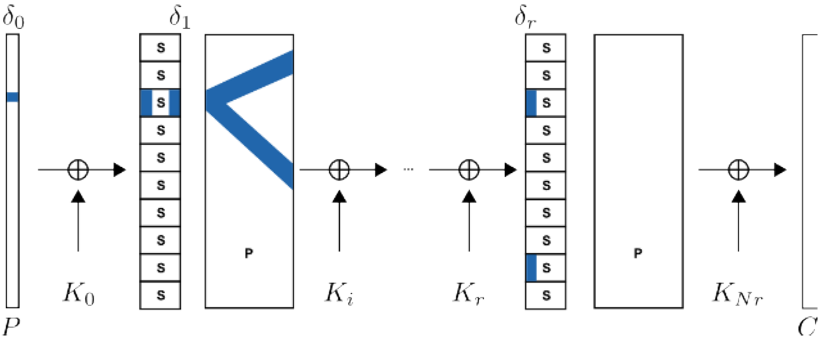
Where S is the 4-bit S-Box and P a permutation of bit positions.

We define reduced round variants and denote them by PRESENT-Ks-Nr.

Prior DC on Reduced Round Versions

Differential characteristics and two round filter function available in [3].

Differential Cryptanalysis I



$$Pr(\delta_i) = p_i \longrightarrow Pr(\Delta) = \prod p_i$$

Differential Cryptanalysis II

Key Recovery:

- **backward key guessing** to recover subkey bits of last rounds not covered by characteristic
- **right pairs** suggest correct and wrong key bits
- **wrong pairs** suggest random key bits
- **filter functions** used to remove wrong pairs
- **candidate key arrays** to count suggestions and observe peak

Differential Cryptanalysis of 16-round DES [1]

- distinguishes right pairs,
- uses outer round active S-Boxes to recover key bits and
- does not rely on candidate key arrays.

Algebraic Cryptanalysis

$$\longrightarrow$$

$$\begin{array}{l}
y_2x_3 + y_3x_3 + x_1x_3 + x_2x_3 + x_3, \\
y_0x_3 + y_3x_3 + x_1x_3 + x_2x_3 + \dots, \\
x_1x_2 + y_3 + x_0 + x_1 + x_3, \\
x_0x_2 + y_3x_3 + x_1x_3 + x_2x_3 + \dots \\
y_3x_2+y_3x_3+x_1x_3+y_0+y_1+y_3 \dots \\
y_0x_2 + y_1x_2 + y_1x_3 + y_3x_3 + \dots \\
x_0x_1 + y_3x_3 + x_1x_3 + x_2x_3 + \dots \\
y_3x_1 + y_3x_3 + x_2x_3 + \dots, \dots
\end{array}$$

We call $X_{i,j}$ and $Y_{i,j}$ the input resp. output variable for the j-th bit of the i-th S-Box application (i.e. round).

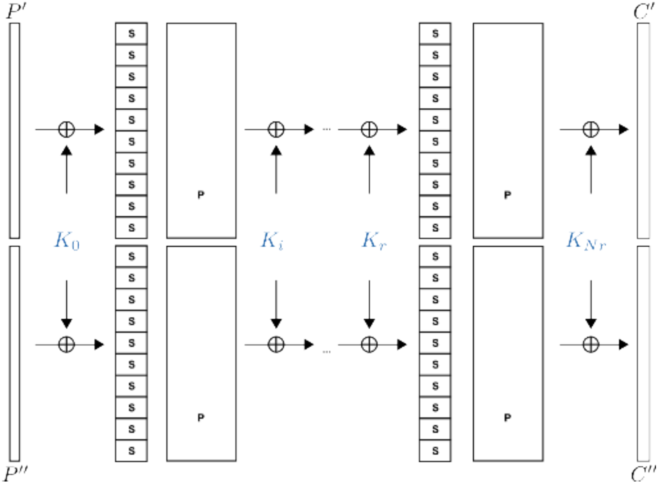
For example, for PRESENT-80-31 we have a system of 4172 variables in 13642 equations.

Multiple $P - C$ Pairs I

Information Security Group

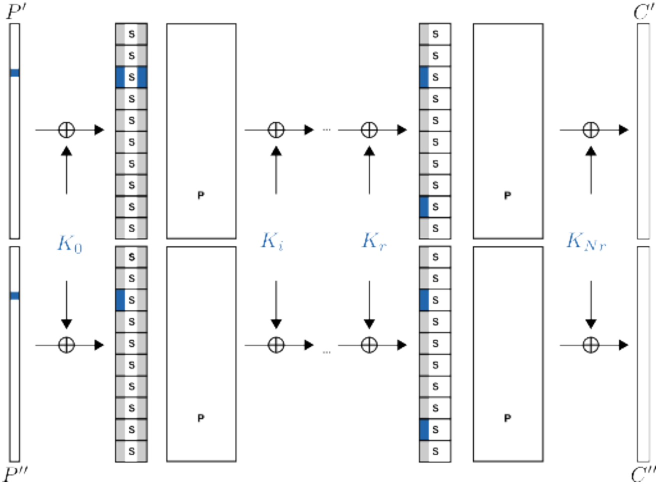
- Given two equation systems F' and $F"$ for two plaintext-ciphertext pairs (P', C') and $(P", C")$ under same encryption key K.
- We can combine these equation systems to form a system $F = F' \cup F"$.
- While F' and $F"$ do not share most of the state variables $X', X", Y', Y"$ but they share the key K and key schedule variables K_i.
- Thus by considering two plaintext–ciphertext pairs the cryptanalyst gathers twice as many equations, involving however many new variables.

Multiple $P - C$ Pairs II



Outline

Attack-A I



Attack-A II

- Each one-round difference gives rise to equations relating the input and output pairs for active S-Boxes.
- We have that the expressions

$$X'_{j,k} + X''_{j,k} = \Delta X_{j,k} \rightarrow \Delta Y_{j,k} = Y'_{j,k} + Y''_{j,k},$$

  where $\Delta X_{j,k}, \Delta Y_{j,k}$are known values predicted by the characteristic, are valid with some non-negligible probability $p_{j,k}$.
- For non-active S-Boxes we have the relations

$$X'_{j,k} + X''_{j,k} = 0 = Y'_{j,k} + Y''_{j,k}$$

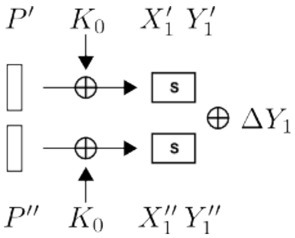
  also valid with a non-negligible probability.

These are $2n$ linear equations per round we can add to our equation system F. The resulting system $\overline{F}$ is expected to be easier to solve **but we need to solve $1/Pr(\Delta)$ such systems**.

Attack-B I

Restrict the first round bits to one active S-Box and assume we have a right pair. The S-Box can be written as a vectorial Boolean function

$$S(X_i) = \begin{array}{l} f_0(X_{i,0}, \ldots, X_{i,n-1}) \\ \ldots \\ f_{n-1}(X_{i,0}, \ldots, X_{i,n-1}) \end{array}.$$



If P', C' and $P"$, $C"$ is a right pair, we have

- $S(P' \oplus K_0) = S(X'_1) = Y'_1$
- $S(P" \oplus K_0) = S(X_1") = Y_1"$
- $Y'_1 \oplus Y_1" = \Delta Y_1$

$\rightarrow S(P'_1 \oplus K_0) \oplus S(P_1" \oplus K_0) = \Delta Y_1$

Attack-B II

We can use this small equation system F_s to recover bits of information about the subkey. Specifically:

Lemma

Given a differential characteristic Δ with a first round active S-Box with a difference that is true with probability 2^{-b}, then by considering F_s we can recover b bits of information about the key from this S-Box.

This is the algebraic equivalent of the well known subkey bit recovery from outer rounds in differential cryptanalysis.

In the case of PRESENT and Wang's differentials we can learn 4-bit of information per characteristic Δ.

Attack-B III

Experimental Observation

For some ciphers **Attack-A** can be used to distinguish **right pairs** and thus enables this attack.

Attack-B proceeds by measuring the time t it maximally takes to find that the system is inconsistent and assume we have a right pair if this time t elapsed without a contradiction.

Alternatively, we may measure other features of a Gröbner basis computation (degree reached, matrix dimensions, . . .).

Attack-B IV

Information Security Group

N_r	K_s	r	$Pr(\Delta)$	SINGULAR	POLYBORI
4	80	3	2^{-12}	106.55-118.15	6.18 – 7.10
4	80	2	2^{-8}	119.24-128.49	5.94 – 13.30
4	80	1	2^{-4}	137.84-144.37	11.83 – **33.47**
16	80	14	2^{-62}	N/A	43.42 – 64.11
16	128	14	2^{-62}	N/A	45.59 – 65.03
16	80	13	2^{-58}	N/A	80.35 – 262.73
16	128	13	2^{-58}	N/A	81.06 – 320.53
16	80	12	2^{-52}	N/A	$>$4 hours
17	80	14	2^{-62}	12,317.49-13,201.99	55.51 - **221.77**
17	128	14	2^{-62}	12,031.97-13,631.52	94.19 - 172.46
17	80	13	2^{-58}	N/A	$>$4 hours

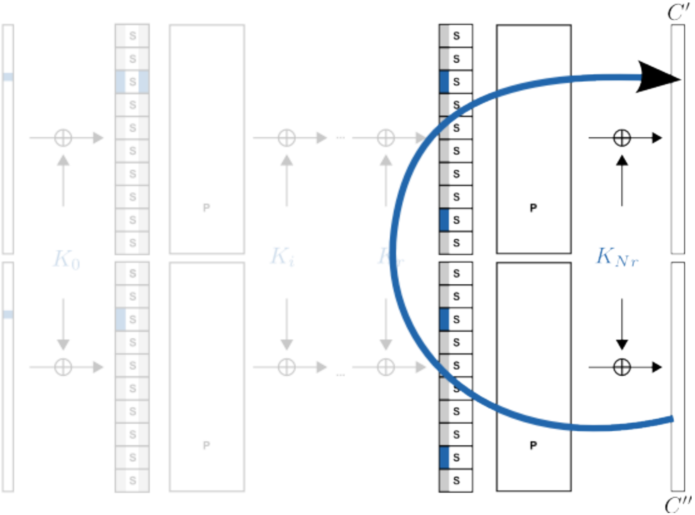
Table: Times in seconds for **Attack-B**

Times obtained on William Stein's `sage.math.washington.edu` computer purchased under NSF Grant No. 0555776.

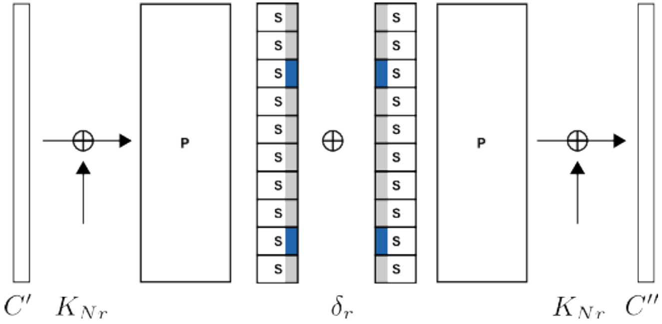
Why?

$$\frac{221.77\ s}{33.47\ s} \approx 6.626$$

Attack-C I



Attack-C II



The algebraic computation is essentially equivalent to solving a related cipher of $2(N_r - r)$ rounds (from C' to C" via the predicted difference δ_r) with a symmetric key schedule, using an algebraic meet-in-the-middle attack.

Attack-C III

Information Security Group

In a Nutshell

Attack-C is an algebraic filter.

Attack-C IV

Information Security Group

N	K_s	r	$Pr(\Delta)$	SINGULAR	POLYBORI	MINISAT2
4	80	4	2^{-16}	0.07 – 0.09	0.05 – 0.06	N/A
4	80	3	2^{-12}	6.69 – 6.79	0.88 – 1.00	0.14 – 0.18
4	80	2	2^{-8}	28.68 – 29.04	2.16 – 5.07	0.32 – 0.82
4	80	1	2^{-4}	70.95 – 76.08	8.10 – 18.30	1.21 – 286.40
16	80	14	2^{-62}	123.82 – 132.47	2.38 – 5.99	N/A
16	128	14	2^{-62}	N/A	2.38 – 5.15	N/A
16	80	13	2^{-58}	301.70 – 319.90	8.69 – 19.36	N/A
16	128	13	2^{-58}	N/A	9.58 – 18.64	N/A
16	80	12	2^{-52}	N/A	$>$ 4 hours	N/A
17	80	14	2^{-62}	318.53 – 341.84	9.03 – 16.93	0.70 – 58.96
17	128	14	2^{-62}	N/A	8.36 – 17.53	0.52 – 8.87
17	80	13	2^{-58}	N/A	$>$ 4 hours	$>$ 4 hours

Table: Times in seconds for **Attack-C**

Outline

PRESENT-80-6 and PRESENT-80-7

- We ran **Attack-C** against PRESENT-80-6 and PRESENT-80-7;
- the algorithm always suggested some key bits after the expected number of runs;
- the algorithm did return false positives (as expected);
- however, a simple majority vote over three experiments, always gave the correct answer.

PRESENT-80-16 I

4 bits:

- **Filter:** $\approx 2^{62}$ ciphertext checks
- **Algebraic Filter:** $\approx 2^{11.93} \cdot 6 \cdot 1.8 \cdot 10^9 \approx 2^{46}$ CPU cycles

Full Key Recovery:

- **Characteristics:** 6 characteristics from [4]
- **Filter:** $\approx 6 \cdot 2^{62}$ ciphertext checks
- **Algebraic Filter:** $\approx 6 \cdot 2^{46}$ CPU cycles
- **Guess:** $80 - 18 = 62$ bits

PRESENT-128-19

Consider the input difference for round 15 and iterate over all possible output differences. For the example difference we have 36 possible output differences for round 15 and $2^{13.93}$ possible output difference for round 16.

4 bits $\approx 2^{13.97} \cdot 1.8 \cdot 10^9 \cdot (18 \cdot 2^{62}) \approx 2^{111}$ CPU cycles.

full key $\approx 2^{13.97} \cdot 1.8 \cdot 10^9 \cdot (18 \cdot 2^{62} + 2 \cdot 6 \cdot 2^{64}) \approx 2^{116}$ CPU cycles.

Complexity Estimates

Information Security Group

Attack	N_r	K_s	r	#pairs	time	#bits
Wang	16	80	14	2^{63}	2^{65} MA	57
Attack-C	16	80	14	2^{62}	2^{62} MA	4
Attack-C	16	80	14	$6 \cdot 2^{62}$	2^{62} encr.	18
Attack-C	19	128	14	2^{62}	2^{111} cycles	4
Attack-C	19	128	14	$6 \cdot 2^{62}$	2^{116} cycles	128

Outline

Discussion

Properties:

- One right pair is sufficient to learn some information about the key.
- No requirement for candidate key counter.
- Silimar to DC attack on full DES [1] but **in theory** applicable to any block cipher.

Some open problems:

- Is this idea applicable to other ciphers?
- How long would it take to solve the small cipher system in Attack-C after a right pair has been identified?
- How about other techniques: linear cryptanalysis, saturation attacks, higher order differentials, ...
- Can we do PRESENT-128-20 with $r = 14$: "a situation without precedent" [2]?

Conclusion

Information Security Group

- We presented a new promising research direction: combining statistical and algebraic cryptanalysis instead of holding on to the "low data complexity dream" normally attached to algebraic cryptanalysis.
- In particular, we presented a new approach which uses algebraic techniques in differential cryptanalysis and showed how to invest more time in the last rounds not covered by a differential using algebraic techniques.
- To illustrate the viability of the attack we applied it against round reduced variants of PRESENT. Of course, this attack has no implication for the security of PRESENT!

Thank you!

Literature I

- Eli Biham and Adi Shamir.
  Differential Cryptanalysis of the Full 16-round DES.
  In *Advances in Cryptology — CRYPTO 1992*, volume 740 of *Lecture Notes in Computer Science*, pages 487–496, Berlin Heidelberg New York, 1991. Springer Verlag.

- A. Bogdanov, L.R. Knudsen, G. Leander, C. Paar, A. Poschmann, Matthew Robshaw, Y. Seurin, and C. Vikkelsoe.
  PRESENT: An ultra-lightweight block cipher.
  In *CHES 2007*, volume 7427 of *Lecture Notes in Computer Science*, pages 450–466. Springer Verlag, 2007.
  Available at http://www.crypto.rub.de/imperia/md/content/texte/publications/conferences/present_ches2007.pdf.

Literature II

Meiqin Wang.
Differential Cryptanalysis of reduced-round PRESENT.
In Serge Vaudenay, editor, *Africacrypt 2008*, volume 5023 of *Lecture Notes in Computer Science*, pages 40–49. Springer Verlag, 2008.

Meiqin Wang.
Private communication: 24 differential characteristics for 14-round present we have found, 2008.