Algebraic–Differential Cryptanalysis of DES

Jean–Charles Faugère Ludovic Perret
Pierre–Jean Spaenlehauer

UPMC – LIP6
CNRS
INRIA Paris - Rocquencourt
SALSA team

Séminaire Crypto – Caen
2009/10/29

Plan

Plan

Algebraic Cryptanalysis

Claude Shannon :

" Breaking a good cipher should require as much work as solving a system of simultaneous equations in a large number of unknowns "

Algebraic Cryptanalysis

- Algebraic representation of a cryptographic primitive.
- Tools for efficient polynomial system solving.
  1. Gröbner Bases algorithms (Buchberger, Faugère F4 and F5).
  2. SAT Solvers.

Remark

There is a very strong link between the modeling and the tools used for the resolution.

Challenge

Can algebraic cryptanalysis be efficient against block ciphers ?

Our work

- SAT Solvers attacks against DES using different modelings of the DES S-boxes.
- Including elements from differential cryptanalysis
  - new attacks against 6,7 and 8 rounds of DES using dedicated characteristics.
- Tradeoff between time and data complexity.

The CNF-SAT problem

Given a set of clauses, decide if there exists an assignment of the variables such that all clauses are satisfied.

Cook-Levin Theorem

The CNF-SAT problem is NP-complete.

$$\bar{a} \vee \bar{b} \vee \bar{c}$$
$$a \vee \bar{d} \vee \bar{c}$$
$$d \vee \bar{c}$$
$$c \vee e$$
$$\bar{c} \vee \bar{e}$$
$$c \vee \bar{e}$$

The CNF-SAT problem

Given a set of clauses, decide if there exists an assignment of the variables such that all clauses are satisfied.

Cook-Levin Theorem

The CNF-SAT problem is NP-complete.

$$\begin{aligned} \bar{a} \vee \bar{b} \vee \bar{c} \\ a \vee \bar{d} \vee \bar{c} \\ d \vee \bar{c} \\ c \vee e \\ \bar{c} \vee \bar{e} \\ c \vee \bar{e} \end{aligned}$$

$$\begin{aligned} a &= true \\ b &= false \\ c &= true \\ d &= true \\ e &= false \end{aligned}$$

Polynomial System Solving (PoSSo)

Why SAT Solvers ?

- NP-hard problem.
- Very efficient and flexible dedicated softwares.
- SAT-competition. Active research field.
- Easy to use. Low memory comsumption.

Which SAT Solver ?

- DPLL: backtracking.
- Stochastic local search.

Courtois N.T. , Bard G.V. and Jefferson C.
Efficient Methods for Conversion and Solution of Sparse Systems of Low-Degree Multivariate Polynomials over GF (2) via SAT-Solvers.
http://eprint.iacr.org/2007/024.pdf

MiniSat2

Een, N. and Sorensson, N.
MiniSat: A SAT solver with conflict-clause minimization

Converting a Polynomial system into a SAT instance

Algebraic Normal Form (ANF) to Conjonctive Normal Form (CNF)

An example

$$x_1x_2 + x_3 + x_3x_1 + x_4 + x_5 = 0$$

- Replace each monomial by a new variable.

$$x_1x_2 + x_3 + x_1x_3 + x_4 + x_5 = 0.$$

Converting a Polynomial system into a SAT instance

Algebraic Normal Form (ANF) to Conjonctive Normal Form (CNF)

An example

$$x_1x_2 + x_3 + x_3x_1 + x_4 + x_5 = 0$$

- Replace each monomial by a new variable.

$$x_1x_2 + x_3 + x_1x_3 + x_4 + x_5 = 0.$$

Converting a Polynomial system into a SAT instance

Algebraic Normal Form (ANF) to Conjonctive Normal Form (CNF)

An example

$$x_1x_2 + x_3 + x_3x_1 + x_4 + x_5 = 0$$

- Replace each monomial by a new variable.

$$\begin{aligned} y_1 + x_3 + y_2 + x_4 + x_5 &= 0 \\ y_1 + x_1x_2 &= 0 \\ y_2 + x_1x_3 &= 0 \end{aligned}$$

Converting a Polynomial system into a SAT instance

Algebraic Normal Form (ANF) to Conjonctive Normal Form (CNF)

An example

$$x_1x_2 + x_3 + x_3x_1 + x_4 + x_5 = 0$$

- "Cut" linear equations.

$$\begin{aligned} y_1 + x_3 + y_2 + x_4 + x_5 &= 0 \\ y_1 + x_1x_2 &= 0 \\ y_2 + x_1x_3 &= 0 \end{aligned}$$

Converting a Polynomial system into a SAT instance

Algebraic Normal Form (ANF) to Conjonctive Normal Form (CNF)

An example

$$x_1x_2 + x_3 + x_3x_1 + x_4 + x_5 = 0$$

- "Cut" linear equations.

$$\begin{aligned} y_1 + x_3 + z_1 &= 0 \\ z_1 + y_2 + z_2 &= 0 \\ z_2 + x_4 + x_5 &= 0 \\ y_1 + x_1x_2 &= 0 \\ y_2 + x_1x_3 &= 0 \end{aligned}$$

Converting a Polynomial system into a SAT instance

Algebraic Normal Form (ANF) to Conjonctive Normal Form (CNF)

An example

$$x_1x_2 + x_3 + x_3x_1 + x_4 + x_5 = 0$$

- Convert into clauses

$$\begin{aligned} y_1 + x_3 + z_1 &= 0 \\ z_1 + y_2 + z_2 &= 0 \\ z_2 + x_4 + x_5 &= 0 \\ y_1 + x_1x_2 &= 0 \\ y_2 + x_1x_3 &= 0 \end{aligned} \Rightarrow \begin{aligned} &\bar{y_1} \vee \bar{x_3} \vee \bar{z_1}, \quad \bar{y_1} \vee x_3 \vee z_1 \\ &y_1 \vee \bar{x_3} \vee z_1, \quad y_1 \vee x_3 \vee \bar{z_1} \\ &\bar{z_1} \vee \bar{y_2} \vee \bar{z_2}, \quad \bar{z_1} \vee y_2 \vee z_2 \\ &z_1 \vee \bar{y_2} \vee z_2, \quad z_1 \vee y_2 \vee \bar{z_2} \\ &\bar{z_2} \vee \bar{x_4} \vee \bar{x_5}, \quad \bar{z_2} \vee x_4 \vee x_5 \\ &z_2 \vee \bar{x_4} \vee x_5, \quad z_2 \vee x_4 \vee \bar{x_5} \\ &\bar{y_1} \vee x_1, \quad \bar{y_1} \vee x_2 \\ &\bar{y_2} \vee x_1, \quad \bar{y_2} \vee x_3 \\ &y_1 \vee \bar{x_1} \vee \bar{x_2}, \quad y_2 \vee \bar{x_1} \vee \bar{x_3} \end{aligned}$$

Solving HFE algebraic systems ?

HFE: a public key algorithm

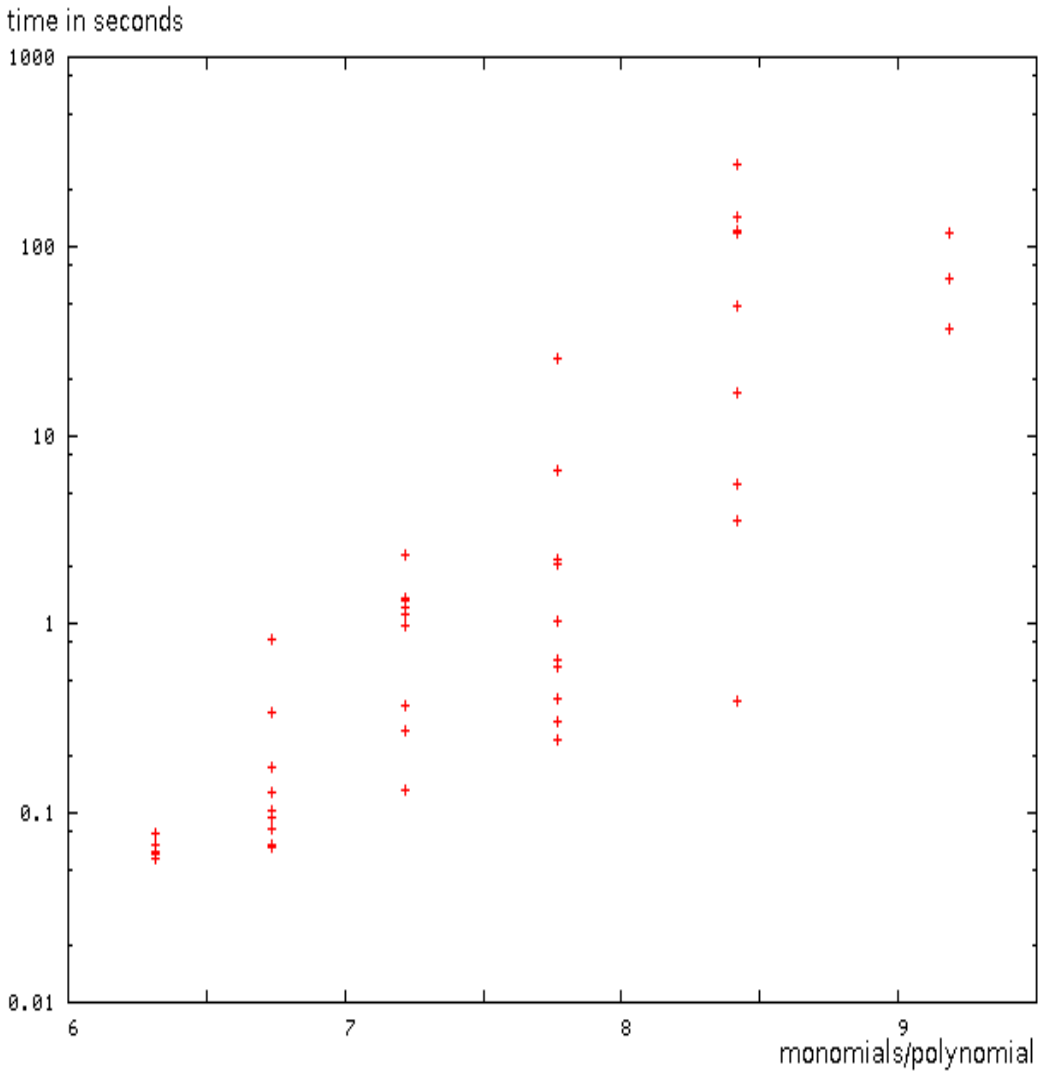
Patarin J.
Hidden fields equations (HFE) and isomorphisms of polynomials (IP): Two new families of asymmetric algorithms
EUROCRYPT'96

- Based on the difficulty of solving quadratic polynomial systems over $GF(2)$.

21 equations, 21 unknowns, $d = 129$.

- F_4 (Magma): 19s.
- MiniSAT (3SAT): 189s.
- MiniSAT (5SAT): 561s.

Sparse random quadratic systems



Plan

Data Encryption Standard

- Iterative Block Cipher.
- Bloc size : 64 bits.
- Effective size of the key : 56 bits.
- Encryption standard between 1976 and 2002.

Why did we choose to study the DES ?

Main attacks against DES

Wiener, M.J.
Efficient DES key search. Technical Report

Meet-in-the-middle attacks

Dunkelman, O. and Sekar, G. and Preneel, B.
Improved Meet-in-the-Middle Attacks on Reduced-Round DES. INDOCRYPT'2007

Main attacks against DES (II)

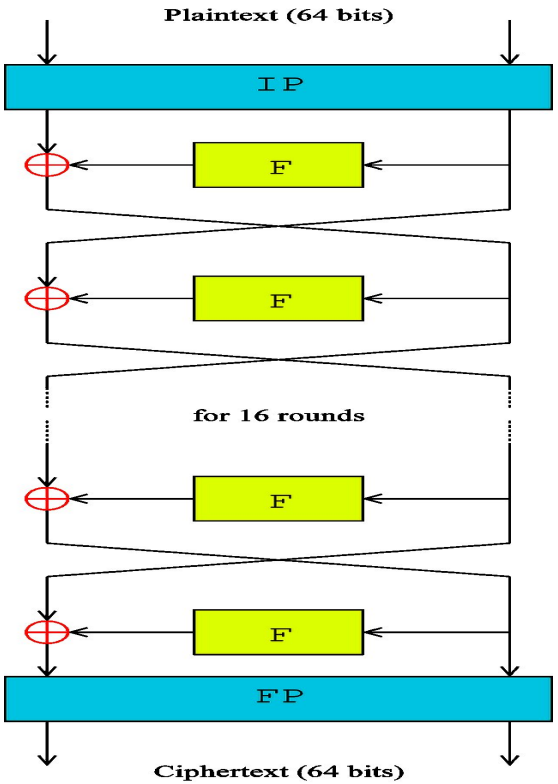

Biham, E. and Shamir, A.
Differential cryptanalysis of the full 16-round DES. Crypto'1992


Knudsen, L.R.
Partial and higher order differentials and applications to the DES. BRICS report


Matsui, M.
Linear cryptanalysis method for DES cipher. EUROCRYPT'1993

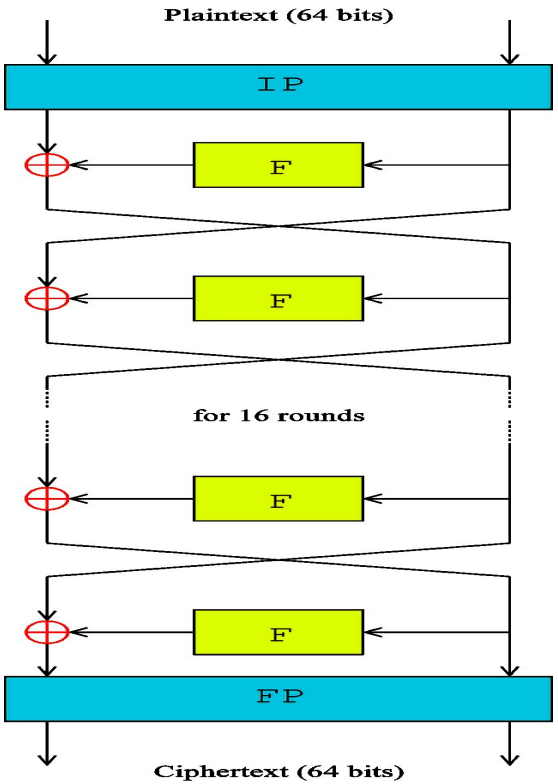
DES Structure

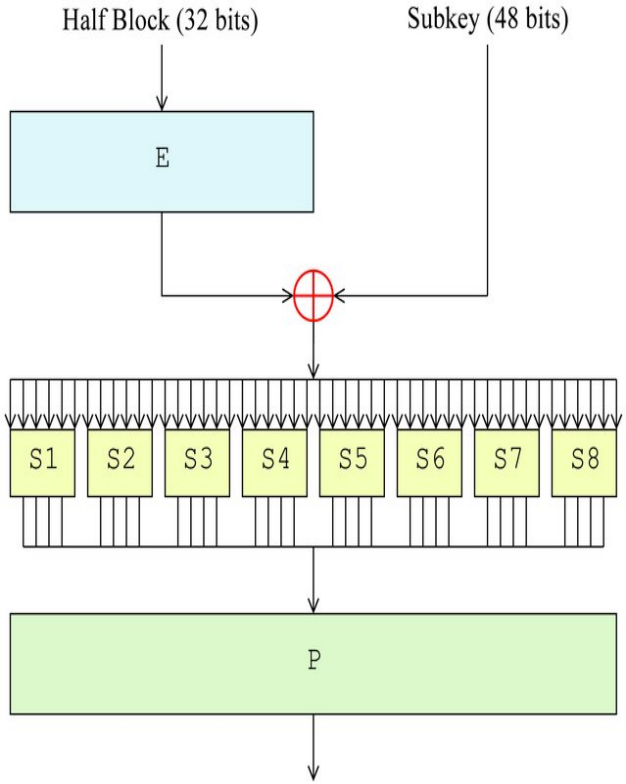
Feistel network



DES Structure

Feistel network





S-boxes : non-linear part of the system

Algebraic cryptanalysis of DES using Minisat

Starting point

N.T. Courtois, G.V. Bard
Algebraic Cryptanalysis of the Data Encryption Standard
IMA Int. Conf. 2007

General principle

- 1 known plaintext.
- Model the cryptosystem by a set of clauses.
- Use Minisat to extract the key.

Remark

We can combine this approach with an exhaustive search over some bits of the key.

S-box modeling (I)

- We have considered several modelings of the DES S-boxes.
- The choice of the modeling is very important.

Our modeling

We search (exhaustively) for the set of polynomials which verify :

$$P(x_1, \ldots, x_6, y_1, \ldots, y_4) = \prod(x_i + \alpha_i) \prod(y_i + \beta_i), \alpha_i, \beta_i \in \{0, 1\}$$

such that

$$S(x_1, \ldots, x_6) = (y_1, \ldots, y_4) \Rightarrow P(x_1, \ldots, x_6, y_1, \ldots, y_4) = 0$$

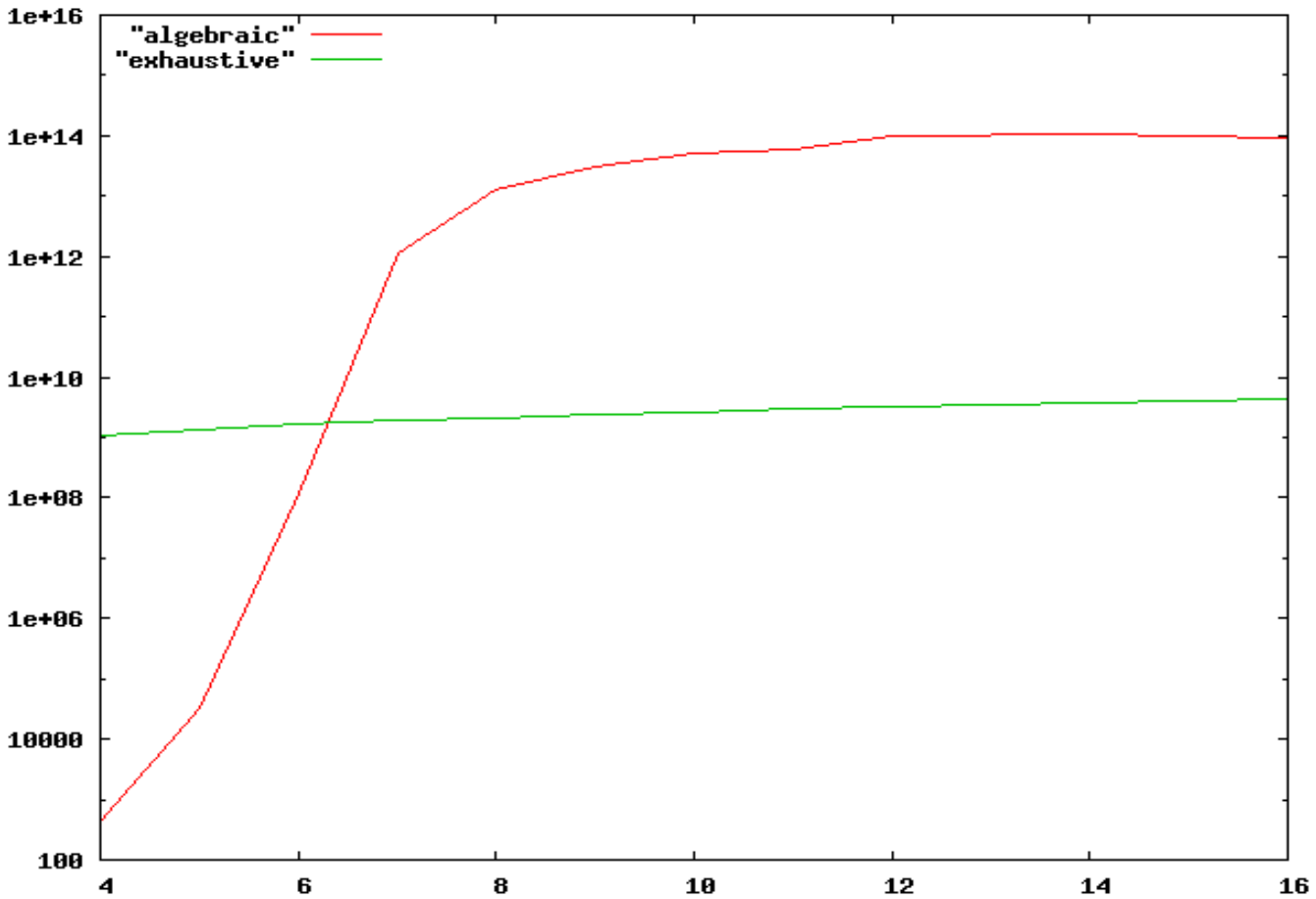
Complexity : 3^{10}

S-box modeling (II)

S-box	Nb of clauses
S1	1624
S2	1844
S3	1767
S4	1881
S5	1812
S6	1705
S7	1673
S8	2047

For 6 rounds : 792 variables and 90086 clauses.
+ partial exhaustive search on 28 bits of the key.

Experimental results



Plan

Our approach

Limit

Algebraic cryptanalysis usually consider only one known plaintext.

We combine algebraic cryptanalysis and statistical techniques to exploit efficiently the knowledge of several plaintexts.

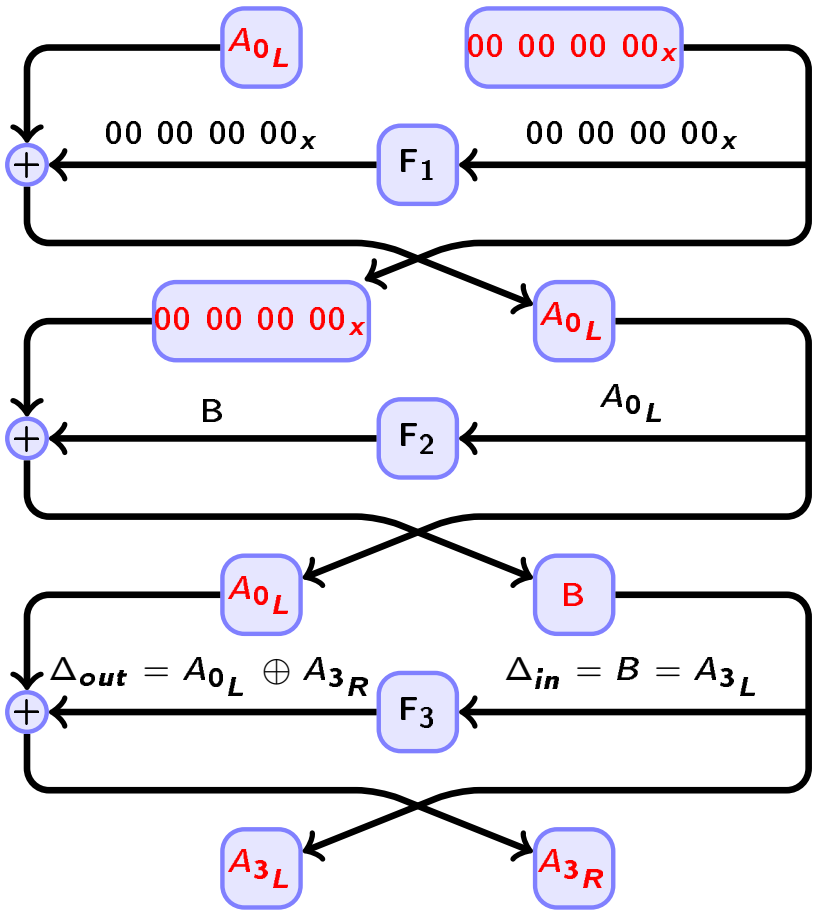
- Tradeoff time/plaintexts.
- In particular, we consider differential cryptanalysis.

Differential cryptanalysis (I)

Biham, E. and Shamir, A.
Differential cryptanalysis of DES-like cryptosystems.
Journal of Cryptology. 1991.

- The principle was known by the DES designers.
- Based on a statistical bias of the S-boxes.
- Key recovery attack.
- We try to predict how the difference of a pair of plaintexts will diffuse through the cipher.

Differential cryptanalysis of the 3 round-reduced DES



Differential cryptanalysis (II)

more than 3 rounds

- Statistical method.
- *differential characteristics*.
- A lot of plaintexts needed.

Motivations

Compromise between algebraic cryptanalysis and differential cryptanalysis.

- We can use the strong correlation between the subkeys.
- The notion of difference is easy to represent with clauses.
- We only need one pair following the characteristic to retrieve the key.

General algorithm

Repeat until the key is found :

- Choose a differential characteristic.
- Choose two plaintexts with difference fixed by the characteristic.
- Construct the system of clauses for DES, and add the clauses corresponding to the characteristic.
- Solve with MiniSat. If the result is UNSATISFIABLE, restart (it means that the pair didn't follow the characteristic). If the result is SATISFIABLE, then MiniSat returns the key.

Six rounds

Approach

- Classical differential characteristics.
- Combination of different characteristics to reduce the data complexity.

Six rounds

Approach

- Classical differential characteristics.
- Combination of different characteristics to reduce the data complexity.

3-round characteristics

- Classical characteristics [BihamShamir], probability: 1/16

  (4008000004000000, 0400000000000000, 0000000004000000, 0400000040080000)

  (0020000800000400, 0000040000000000, 0000000000000400, 0000040000200008)

- 4 extra characteristics, probability: 1/20

  (0080820060000000, 6000000000000000, 0000000060000000, 6000000000808200)

  (4000401002000000, 0200000000000000, 0000000002000000, 0200000040004010)

  (0000401006000000, 0600000000000000, 0000000006000000, 0600000000004010)

  (0010000100000060, 0000006000000000, 0000000000000060, 0000006000100001)

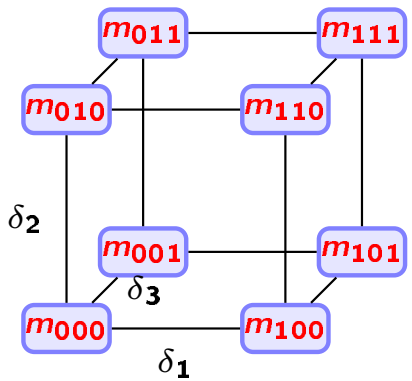
How to combine characteristics ?

A first attempt: "cube combination"

Principle

Combining n characteristics:

- Choose $m_{0,\ldots,0}$ uniformly at random.
- For $v \in \{0,1\}^n$, use $m_{\mathbf{v}} = m_0 \oplus \sum_{i=1}^{n} v_i \delta_i$.
- Number of trials = number of edges of a n-dimensional cube.
- Number of chosen plaintexts = number of vertices of a n-dimensional cube.



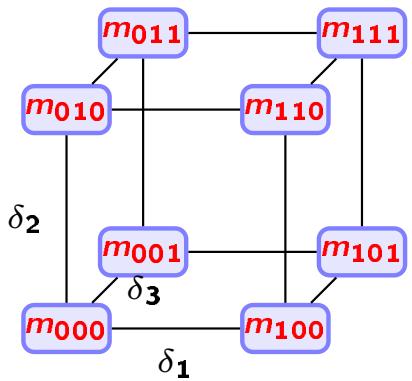
How to combine characteristics ?

A first attempt: "cube combination"

Principle

Combining n characteristics:

- Choose $m_{0,\ldots,0}$ uniformly at random.
- For $v \in \{0,1\}^n$, use $m_{\mathbf{v}} = m_0 \oplus \sum_{i=1}^{n} v_i \delta_i$.
- Number of trials = number of edges of a n-dimensional cube.
- Number of chosen plaintexts = number of vertices of a n-dimensional cube.



but... not independant events

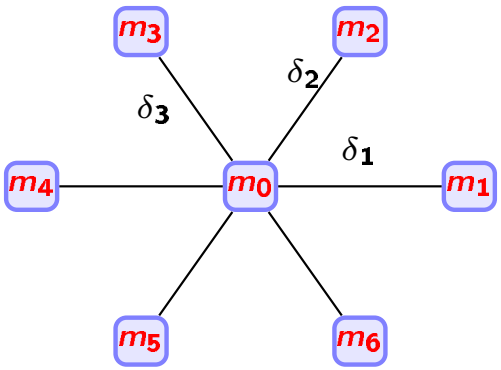
How to combine characteristics ?

Second attempt

Principle

Combining n characteristics:

- Choose m_0 uniformly at random.
- For $v \in \{1, \ldots, n\}$, use $m_i = m_0 \oplus \delta_i$.
- Number of trials $= n$.
- Number of chosen plaintexts $= n - 1$.



Experimental results

Cryptanalysis	Plaintexts	Time
Differential (Biham, Shamir)	240	0,3 seconds
Differential (Knudsen)	46	a few seconds
Algebraic with SAT Solver (Courtois, Bard)	1	2^{25} seconds
Algebraic-differential	32	3000 seconds
Algebraic-differential (combination of characteristics)	22	<10 hours

Seven rounds

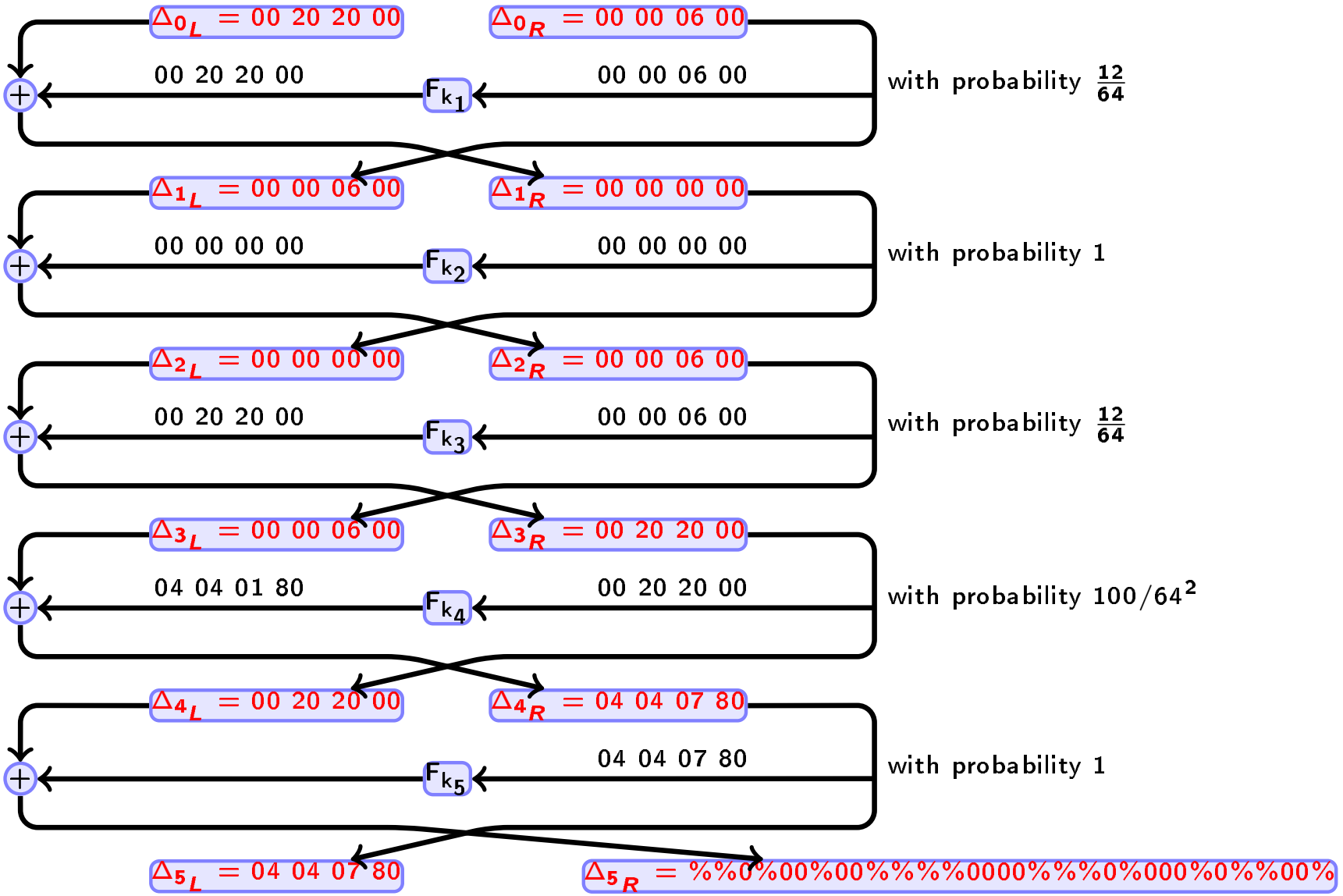
For seven rounds and more, the classical differential characteristics don't seem to be adapted.

We have used a dedicated differential characteristic.

- Truncated characteristic with probability 1/1000.

S-Box	δ_{in}	δ_{out}	Proba
S1	4_x	6_x	10/64
	8_x	3_x	12/64
S2	4_x	10_x	10/64
	4_x	12_x	10/64
	8_x	9_x	10/64
	8_x	10_x	16/64
	12_x	5_x	14/64
S3	4_x	9_x	12/64
	8_x	3_x	10/64
	12_x	5_x	12/64
	12_x	6_x	12/64
S4	4_x	6_x	12/64
	4_x	9_x	12/64

Boîte-S	δ_{in}	δ_{out}	Proba
S5	4_x	6_x	10/64
	8_x	6_x	10/64
	8_x	10_x	10/64
	12_x	3_x	10/64
	12_x	6_x	10/64
	12_x	10_x	10/64
S6	8_x	6_x	16/64
	12_x	3_x	12/64
	12_x	5_x	10/64
S7	8_x	10_x	12/64
	12_x	12_x	14/64
S8	4_x	12_x	10/64
	12_x	5_x	10/64
	12_x	6_x	10/64


Δ_{0_L} = 00 20 20 00
Δ_{0_R} = 00 00 06 00
00 20 20 00
F_{k_1}
00 00 06 00
with probability $\frac{12}{64}$
Δ_{1_L} = 00 00 06 00
Δ_{1_R} = 00 00 00 00
00 00 00 00
F_{k_2}
00 00 00 00
with probability 1
Δ_{2_L} = 00 00 00 00
Δ_{2_R} = 00 00 06 00
00 20 20 00
F_{k_3}
00 00 06 00
with probability $\frac{12}{64}$
Δ_{3_L} = 00 00 06 00
Δ_{3_R} = 00 20 20 00
04 04 01 80
F_{k_4}
00 20 20 00
with probability $100/64^2$
Δ_{4_L} = 00 20 20 00
Δ_{4_R} = 04 04 07 80
04 04 07 80
F_{k_5}
with probability 1
Δ_{5_L} = 04 04 07 80
Δ_{5_R} = %%0%00%00%%%%0000%%%0%000%0%%00%

Experimental results

7 rounds cryptanalysis

- 2000 chosen plaintexts
- 3 hours

Not so much results on 7 rounds in the literature.

Eight rounds

We have found a 5-round truncated differential characteristic with probability 1/5800.
+ partial exhaustive search over 8 bits of the key.

8 rounds cryptanalysis

- 11600 chosen plaintexts and 2^{25} seconds.

Summary

Rounds	Cryptanalysis	Nb of plaintexts	Time
6	diff (Biham,Shamir)	240 (chosen)	0,3 s
	diff (Knudsen)	46 (chosen)	$<$10 s
	alg (Courtois,Bard)	1 (known)	2^{25} s
	diff + alg	**32 (chosen)**	**3000 s**
	diff + alg	**22 (chosen)**	**$<$10 h**
7	**diff + alg**	**2000 (chosen)**	**10000 s**
8	diff (Biham,Shamir)	50000 (chosen)	100 s
	lin (Matsui)	2^{20} (known)	40 s
	diff + alg	**11500 (chosen)**	2^{25} **s**
	diff+lin (Hellman,Langford)	512 (chosen)	few seconds

Conclusion

- Use of statistical methods in algebraic cryptanalysis.

  $\longrightarrow$ New attacks on 6, 7 and 8 rounds of DES using dedicated characteristics.
- Tradeoff plaintexts/time.

Conclusion

- Use of statistical methods in algebraic cryptanalysis.

  $\longrightarrow$ New attacks on 6, 7 and 8 rounds of DES using dedicated characteristics.
- Tradeoff plaintexts/time.

Related work (Cryptanalysis of Present)

M. Albrecht and C. Cid
Algebraic Techniques in Differential Cryptanalysis
FSE2009

Conclusion

- Use of statistical methods in algebraic cryptanalysis.

  $\longrightarrow$ New attacks on 6, 7 and 8 rounds of DES using dedicated characteristics.
- Tradeoff plaintexts/time.

Related work (Cryptanalysis of Present)

M. Albrecht and C. Cid
Algebraic Techniques in Differential Cryptanalysis
FSE2009

Future work

- Extension of this attack for more rounds ?
- Algebraic-differential cryptanalysis of DES with Gröbner Bases ?
- Other cryptosystems ?
- Other statistical tools (differential-linear cryptanalysis, ...) ?