A Library for Analysis of Substitutions

Oleksandr Kazymyrov

Selmer Center, Department of Informatics,
University of Bergen, Norway
Oleksandr.Kazymyrov@ii.uib.no

PhD seminar
Autumn 2013

Agenda

1. Introduction
2. Preliminaries
3. Criteria of Substitutions
4. Sage and Libraries

Substitutions

Definition

Substitution box (S-box) is an arbitrary mapping of one alphabet to another.

Substitutions for cryptography

S-boxes used in cryptography often map elements from vector space $\mathbb{F}_2^n$ to $\mathbb{F}_2^m$.

Substitutions

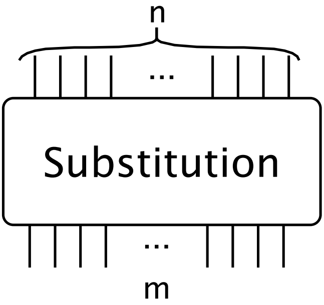

Figure : A Substitution Box

Substitutions

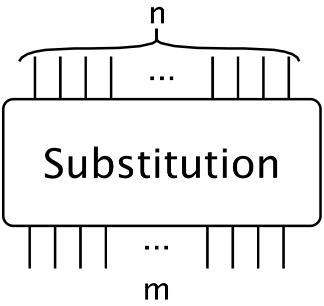

Figure : A Substitution Box

Possible variants

- $n > m$
- $n < m$
- $n = m$

Substitutions

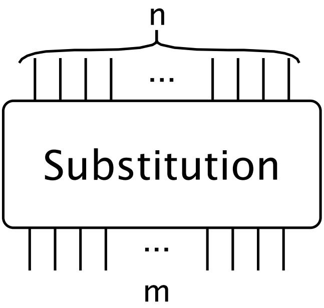

Figure : A Substitution Box

Possible variants

- $n > m$
- $n < m$
- $n = m$
  - $\#img(\text{S-box}) = 2^n$

Substitutions

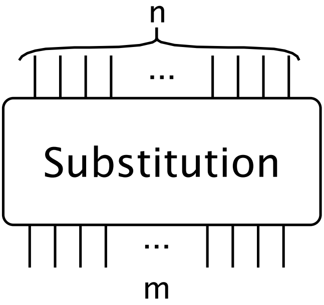

Figure : A Substitution Box

Possible variants

- $n > m$
- $n < m$
- $n = m$
  - $\#img(\text{S-box}) = 2^n$

Representations

- lookup tables
- vectorial Boolean functions
  - Boolean functions
- system of equations

Examples of substitutions

Table : Examples of substitutions for different n and m

n	m	S-box
3	3	$\{7, 1, 0, 4, 2, 3, 5, 6\}$
3	3	$\{3, 0, 0, 1, 1, 7, 7, 5\}$
3	1	$\{1, 1, 0, 0, 1, 1, 1, 0\}$
3	2	$\{1, 1, 0, 0, 1, 1, 1, 0\}$
3	3	$\{1, 1, 0, 0, 1, 1, 1, 0\}$
3	4	$\{1, 1, 0, 0, 1, 1, 1, 0\}$
3	8	$\{255, 5, 83, 11, 3, 7, 5, 1\}$

Application of S-boxes

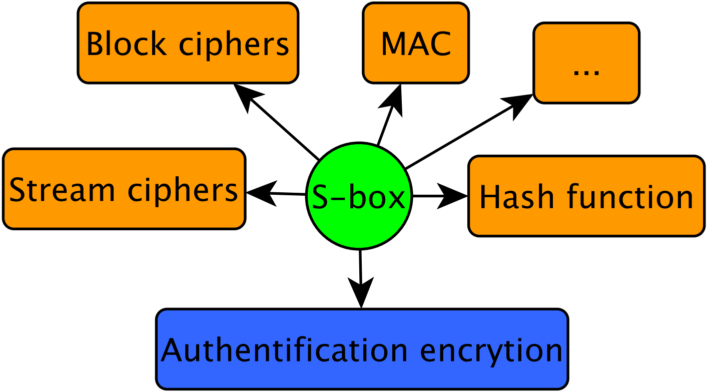

Figure : Application of S-boxes

List of properties

Definition

An S-box is a mapping of an n-bit input message to an m-bit output message.

- Minimum degree
- Balancedness
- Non-linearity
- Correlation immunity
- δ-uniformity
- Cyclic structure
- Algebraic immunity
- Absolute indicator
- Absence of fixed points
- Propagation criterion
- Sum-of-squares indicator
- ...

Agenda

1. Introduction
2. Preliminaries
3. Criteria of Substitutions
4. Sage and Libraries

Properties of substitutions (1/5)

Definition

Let n and m be two positive integers. Any function $F : \mathbb{F}_2^n \mapsto \mathbb{F}_2^m$ is called an (n, m)-function or vectorial Boolean function.

δ-uniform

Arbitrary F is differentially δ-uniform if equation

$$b = F(x) + F(x + a),\ \forall a \in \mathbb{F}_2^n, \forall b \in \mathbb{F}_2^m, a \neq 0$$

has at most δ solutions.

Properties of substitutions (2/5)

Walsh transform

The Walsh transform of an (n, m)-function F at $(u, v) \in \mathbb{F}_2^n \times \mathbb{F}_2^m \backslash \{0\}$

$$\lambda(u, v) = \sum_{x \in \mathbb{F}_2^n} (-1)^{v \cdot F(x) \oplus u \cdot x}, \qquad (1)$$

where "$\cdot$" denotes inner products in $\mathbb{F}_2^n$ and $\mathbb{F}_2^m$ respectively.

Non-linearity

$$NL(F) = 2^{n-1} - \frac{1}{2} \max_{v \in \mathbb{F}_2^{m*};\ u \in \mathbb{F}_2^n} |\lambda(u, v)|$$

Properties of substitutions (3/5)

Balancedness

An (n, m)-function F is called balanced if it takes every value of F_2^m the same number of times (2^{n-m}).

Absence of Fixed Points

A substitution must not have fixed point, i.e.

$$F(a) \neq a, \quad \forall a \in \mathbb{F}_2^n.$$

Properties of substitutions (4/5)

The algebraic normal form (ANF) of any (n, m)-function F always exists and is unique:

$$F(x) = \sum_{I \subseteq \{1,\ldots,n\}} a_I \left(\prod_{i \in I} x_i \right) = \sum_{I \subseteq \{1,\ldots,n\}} a_I x^I, \; a_I \in \mathbb{F}_2^m$$

The algebraic degree of F

$$deg(F) = \max\{|I| \mid a_I \neq 0\}$$

Minimum degree

The minimum algebraic degree of all the component functions of F is called the minimum degree.

Properties of substitutions (5/5)

Arbitrary substitution can be represented as the system of equations

$$\begin{cases} g_1(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) = 0; \\ g_2(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) = 0; \\ \ldots \\ g_r(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) = 0. \end{cases} \quad (2)$$

Properties of substitutions (5/5)

Arbitrary substitution can be represented as the system of equations

$$\begin{cases} g_1(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) = 0; \\ g_2(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) = 0; \\ \ldots \\ g_r(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m) = 0. \end{cases} \quad (2)$$

Algebraic immunity

The algebraic immunity $AI(F)$ of any (n, m)-function F is the minimum algebraic degree of all functions in (2).

Agenda

1 Introduction

2 Preliminaries

3 Criteria of Substitutions

4 Sage and Libraries

List of properties

Definition

An S-box is a mapping of an n-bit input message to an m-bit output message.

- Minimum degree
- Balancedness
- Non-linearity
- Correlation immunity
- δ-uniformity
- Cyclic structure
- Algebraic immunity
- Absolute indicator
- Absence of fixed points
- Propagation criterion
- Sum-of-squares indicator
- ...

Necessary properties for stream ciphers (FG)

Definition

An S-box is a mapping of an n-bit input message to an m-bit output message.

- Minimum degree
- Balancedness
- Non-linearity
- Correlation immunity
- δ-uniformity
- Cyclic structure
- Algebraic immunity
- Absolute indicator
- Absence of fixed points
- Propagation criterion
- Sum-of-squares indicator
- ...

Necessary properties for block ciphers

Definition

An S-box is a mapping of an n-bit input message to an m-bit output message.

- Minimum degree
- Balancedness
- Non-linearity
- Correlation immunity
- δ-uniformity
- Cyclic structure
- Algebraic immunity
- Absolute indicator
- Absence of fixed points
- Propagation criterion
- Sum-of-squares indicator
- ...

Perfect nonlinear substitutions

Definition

An S-box is a mapping of an n-bit input message to an m-bit output message.

- Minimum degree
- Balancedness
- Non-linearity
- Correlation immunity
- δ-uniformity
- Cyclic structure
- Algebraic immunity
- Absolute indicator
- Absence of fixed points
- Propagation criterion
- Sum-of-squares indicator
- ...

Optimal substitutions

Definition

Substitutions satisfying only mandatory criteria essential for a particular cryptographyc algorithm are called optimal.

Optimal substitutions

Definition

Substitutions satisfying only mandatory criteria essential for a particular cryptographyc algorithm are called optimal.

An optimal substitution for a block cipher

- permutation
- maximum value of minimum degree
- without fixed points (cycles of length 1)
- maximum algebraic immunity/minimum number of equations

Optimal substitutions

Definition

Substitutions satisfying only mandatory criteria essential for a particular cryptographyc algorithm are called optimal.

An optimal substitution for a block cipher

- permutation
- maximum value of minimum degree
- without fixed points (cycles of length 1)
- maximum algebraic immunity/minimum number of equations
  - minimum δ-uniformity
  - maximum non-linearity

Optimal substitutions

Definition

Substitutions satisfying only mandatory criteria essential for a particular cryptographyc algorithm are called optimal.

An optimal permutation for a block cipher

- ~~permutation~~
- maximum value of minimum degree
- ~~without fixed points (cycles of length 1)~~
- maximum algebraic immunity/~~minimum number of equations~~
  - minimum δ-uniformity
  - maximum non-linearity

Example of criteria

An optimal permutation without fixed points for $n = m = 8$ must have

- minimum degree 7
- algebraic immunity 3 (441 equations)
- $\delta \leq 8$
- $NL \geq 104$

Problems

- Generation
  - How to generate effectively?
  - How to foresee properties in advance?
  - Find optimal and general structures
- Cryptanalysis
  - Check for all known criteria
  - Algebraic properties (e.g. cyclic properties, algebraic degree)
  - Protection against physical attack (e.g. fault attacks)
  - Create new criteria for generation

Agenda

1 Introduction

2 Preliminaries

3 Criteria of Substitutions

4 Sage and Libraries

System for Algebra and Geometry Experimentation (Sage)

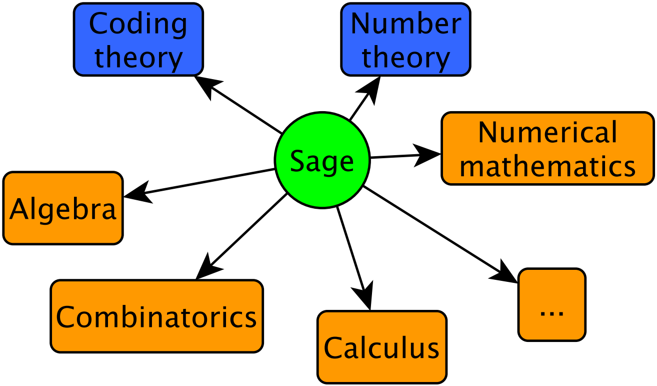

Figure : One can use Sage for ...

System for Algebra and Geometry Experimentation (Sage)

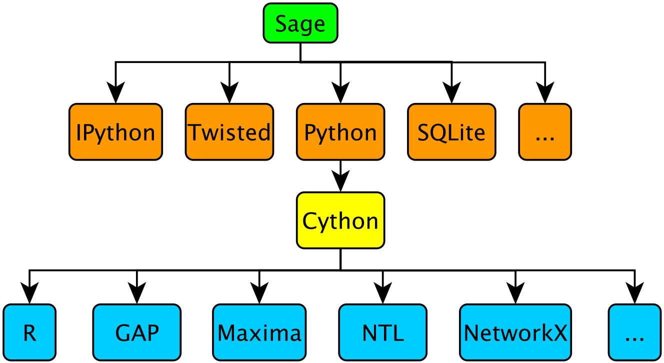

Figure : Sage components

Documentation

- William Stein. Sage for Power Users // Link
- Martin R. Albrecht. Sage for Cryptographers // ECrypt II PhD Summer School
- Martin R. Albrecht. Sage & Algebraic Techniques for the Lazy Symmetric Cryptographer // IceBreak, Reykjavik, Iceland
- Martin R. Albrecht et al. Documentation of SBox class // Link

RSA

```
sage: p=random_prime(2^512)
sage: q=random_prime(2^512)
sage: n=p*q
sage: phi=(p-1)*(q-1)
sage: d=randint(2,phi-1)
sage: e=xgcd(d,phi)[1]
sage: print "Is '{0}' one?".format((e*d)%phi)
sage: M=randint(0,n-1)
sage: C=power_mod(M,e,n)
sage: print "Is '{0}' True?".format(power_mod(C,d,n)==
```

sage.crypto.mq.sbox.SBox

Listing 1 : Initialization Step

```
sage: S = mq.SBox(1, 3, 0, 2); S
(1, 3, 0, 2)
sage: S(1)
3
```

Listing 2 : Example of functions

```
sage: S.maximal_difference_probability_absolute()
4
sage: S.difference_distribution_matrix()
[4 0 0 0]
[0 0 4 0]
[0 4 0 0]
[0 0 0 4]
```

Functions in SBox

- $2^n - 2NL(F)$

  S.maximal_linear_bias_absolute()

- δ-uniformity

  S.maximal_difference_probability_absolute()

- System of equation for algebraic attack

  S.polynomials()

- Univariate form

  S.interpolation_polynomial()

Fail example

Listing 3 : AES Sbox

```
sage: sbox = [0x63,0x7c,0x77,0x7b,0xf2,0x6b,...
sage: S = mq.SBox(sbox)
sage: S.maximal_difference_probability_absolute()
4
sage: S.maximal_linear_bias_absolute()
16
sage: S.interpolation_polynomial
(a + 1)*x^254 + (a^6 + a^5 + a^2)*x^253 + ...
sage: S.polynomials(degree=2)
[]
```

Pros and Cons of SBox

Advantages	Disadvantages
Integrated to Sage	Slow
Important functions are in	A few functions
	Known bugs

Pros and Cons of SBox

Advantages	Disadvantages
Integrated to Sage	Slow
Important functions are in	A few functions
	Known bugs

Citation of Martin R. Albrecht

"How do I do ... in Sage?" ...It's easy: implement it and send us a patch.

SBox vs Sbox

Pluses

- Oriented on arbitrary n and m
- Optimized for performance
- Implemented lots of cryptographic criteria

Minuses

- Quite hard to compile
    - Works only in console

SBox vs Sbox

Pluses

- Oriented on arbitrary n and m
- Optimized for performance
- Implemented lots of cryptographic criteria

Minuses

- Quite hard to compile
    - Works only in console

Example

One look is worth a thousand words.

List of supported characteristics

- Minimum degree
- Balancedness
- Non-linearity
- Correlation immunity
- δ-uniformity
- Cyclic structure
- Algebraic immunity
- Absolute indicator
- Absence of fixed points
- Propagation criterion
- Sum-of-squares indicator
- ...

Additional functionality

- Extra functions
  - Resilience (balancedness and correlation immunity)
  - Maximum of linear approximation table
  - Check function on APN (optimized)
- Convert linear functions to matrices and vice versa
- Apply EA- and CCZ-equivalence
- Generation of substitutions
  - Based on user polynomial (trace supported)
  - Predefined functions (APN for $n = 6$, Welch, Kasami, Dickson, Dobbertin ...)
  - Random substitution/permutation

Comparison of known substitutions

Properties	AES	GOST R 34.11-2012	STB 34.101.31-2011	Kalyna S0	Proposed S-box
δ-uniformity	4	8	8	8	8
Non-linearity	112	100	102	96	104
Absolute Indicator	32	96	80	88	80
SSI	133120	258688	232960	244480	194944
Minimum Degree	7	7	6	7	7
Algebraic Immunity	2(39)	3(441)	3(441)	3(441)	3(441)

Table : Substitutions comparison