Boolean Functions in Cryptology

Dr. Sugata Gangopadhyay

Department of Mathematics
Indian Institute of Technology Roorkee
Roorkee 247667 INDIA,
Tutorial Workshop on Many Facets of Cryptology
October 14–15, 2011,
`gsugata@gmail.com`

Outline of the talk

- A basic model of stream cipher.
- Use of Boolean functions in stream ciphers.
- Some cryptographically significant properties of Boolean functions.

Stream Ciphers

- Alice sends message to Bob over a public channel. Assume that the message is a binary sequence.
- Oscar has access to the channel.
- While sending the binary sequence (message) Alice xor-s it bitwise with a random binary sequence (keystream).
- In case Bob has access to the same random binary sequence (keystream) he can retrieve the original message sent by Alice by xor-ing the random binary sequence to the received message.
- The reason is that if $x \in \{0, 1\}$ then $x \oplus x = 0$.

A Statistical model

- Each message bit is considered as an instance of a random variable X.
- Each keystream bit is considered as an instance of a random variable Z.
- Then each ciphertext bit is an instance of the random variable $Y = X \oplus Z$.
- $Pr[X = 0] = p, Pr[X = 1] = 1 - p$ and $Pr[Z = 0] = Pr[Z = 1] = 1/2$.

Killing the probability distribution of X

- $$\begin{aligned} Pr[X \oplus Z = 0] &= Pr[X = 1]Pr[Z = 1] + Pr[X = 0]Pr[Z = 0] \\ &= (1 - p)(1/2) + p(1/2) = 1/2. \end{aligned} \tag{1}$$

  $$\begin{aligned} Pr[X \oplus Z = 1] &= Pr[X = 0]Pr[Z = 1] + Pr[X = 1]Pr[Z = 0] \\ &= p(1/2) + (1 - p)(1/2) = 1/2. \end{aligned} \tag{2}$$
- Is it possible for X to be such that $p = 1/2$?

Use of Pseudorandom binary sequence generator

- Instead of using a true random binary sequence Alice uses a finite state machine (FSM) which generates a pseudorandom binary sequence which we refer to as the keystream.
- Alice initializes the FSM with a short key (say k bits long) which she shares with Bob.
- Then Bob is also able to generate the same pseudorandom binary sequence and hence able to retrieve the original message sent by Alice.

Oscar's activity

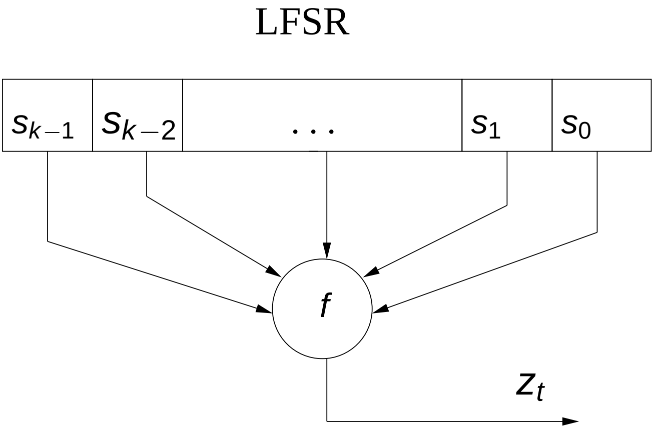
- We assume that Oscar can access any M bits of the keystream.
- If the cipher is optimally secure then Oscar must not be able to compute the key with complexity less than 2^k - that is less than exhaustive search over the entire keyspace.

A model of a stream cipher

- $\mathbf{x} = (x_1, \ldots, x_N) \in GF(2)^N$ be a state of an FSM.
- $L : GF(2)^N \rightarrow GF(2)^N$ be the state update function.
- $f : GF(2)^n \rightarrow GF(2)$, where $n \leq N$ is a Boolean function.
- The generation of keystream is done as follows:

$$f(L^i(\mathbf{x})) = z_i, \text{ for } i = 0, 1, 2, 3, \ldots, M, \ldots . \quad (3)$$

Stream cipher using LFSR



Nonlinear filtering generator

Linear case

- Suppose f and L used in (3) are both linear, i.e., $f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ and $L(\mathbf{x}+\mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y})$.
- The (3) is a system of linear equations.
- This means that if we have N such linearly independent equations then we can solve them for $\mathbf{x}$.
- If we obtain $2N$ such equations it is seen by experiment that we usually obtain N linearly independent equations in the process and hence can solve them and obtain the initial state $\mathbf{x}$.

Introduction of nonlinearity

- In large number of stream ciphers L is fixed to be linear while f is not linear.
- However use of any nonlinear function may not solve the problem.
- Example:

$$f(x_1, \ldots, x_4) = x_1 x_2 x_3 x_4 \oplus x_1 \oplus x_2.$$

Affine approximation

$$f(x_1,\ldots,x_4) = x_1x_2x_3x_4 \oplus x_1 \oplus x_2; \ell(x_1,\ldots,x_4) = x_1 \oplus x_2.$$

x_4	x_3	x_2	x_1	f	ℓ
0	0	0	0	0	0
0	0	0	1	1	1
0	0	1	0	1	1
0	0	1	1	0	0
0	1	0	0	0	0
0	1	0	1	1	1
0	1	1	0	0	1
0	1	1	1	1	0
1	0	0	0	0	0
1	0	0	1	1	1
1	0	1	0	0	1
1	0	1	1	0	0
1	1	0	0	0	0
1	1	0	1	1	1
1	1	1	0	1	1
1	1	1	1	1	0

Affine approximation

- Suppose that the initial state of the FSM, $\mathbf{x}$ is randomly chosen and we compute $f(L^i(\mathbf{x})) = z_i$ for $i = 0, 1, \ldots M$.
- The probability that the above system of equations is same as $\ell(L^i(\mathbf{x})) = z_i$ for $i = 0, 1, \ldots, M$ is $(1 - \frac{1}{16})^M$. (Of course under certain assumptions)
- Since the second system is linear it can be solved for the initial state $\mathbf{x}$.
- Thus the function f must be "far away" from all the affine functions

$$a_1x_1 \oplus a_2x_2 \oplus \ldots \oplus a_nx_n \oplus \epsilon,$$

for all $(a_1, \ldots, a_n) \in GF(2)^n$ and $\epsilon \in GF(2)$. In the present case $n = 4$.
- This can also be written as

$$\ell_{\mathbf{a},\epsilon}(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} \oplus \epsilon,$$

where $\mathbf{a} \cdot \mathbf{x} = \bigoplus_{i=1}^{n} a_ix_i$ is the inner product of $\mathbf{a}$ and $\mathbf{x}$.

Affine approximation

- Suppose that the initial state of the FSM, $\mathbf{x}$ is randomly chosen and we compute $f(L^i(\mathbf{x})) = z_i$ for $i = 0, 1, \ldots M$.
- The probability that the above system of equations is same as $\ell(L^i(\mathbf{x})) = z_i$ for $i = 0, 1, \ldots, M$ is $(1 - \frac{1}{16})^M$. (Of course under certain assumptions)
- Since the second system is linear it can be solved for the initial state $\mathbf{x}$.
- Thus the function f must be “far away” from all the affine functions

$$a_1x_1 \oplus a_2x_2 \oplus \ldots \oplus a_nx_n \oplus \epsilon,$$

  for all $(a_1, \ldots, a_n) \in GF(2)^n$ and $\epsilon \in GF(2)$. In the present case $n = 4$.
- This can also be written as

$$\ell_{\mathbf{a},\epsilon}(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} \oplus \epsilon,$$

  where $\mathbf{a} \cdot \mathbf{x} = \bigoplus_{i=1}^{n} a_ix_i$ is the inner product of $\mathbf{a}$ and $\mathbf{x}$.

Affine approximation

- Suppose that the initial state of the FSM, $\mathbf{x}$ is randomly chosen and we compute $f(L^i(\mathbf{x})) = z_i$ for $i = 0, 1, \ldots M$.
- The probability that the above system of equations is same as $\ell(L^i(\mathbf{x})) = z_i$ for $i = 0, 1, \ldots, M$ is $(1 - \frac{1}{16})^M$. (Of course under certain assumptions)
- Since the second system is linear it can be solved for the initial state $\mathbf{x}$.
- Thus the function f must be "far away" from all the affine functions

$$a_1x_1 \oplus a_2x_2 \oplus \ldots \oplus a_nx_n \oplus \epsilon,$$

for all $(a_1, \ldots, a_n) \in GF(2)^n$ and $\epsilon \in GF(2)$. In the present case $n = 4$.
- This can also be written as

$$\ell_{\mathbf{a},\epsilon}(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} \oplus \epsilon,$$

where $\mathbf{a} \cdot \mathbf{x} = \bigoplus_{i=1}^n a_ix_i$ is the inner product of $\mathbf{a}$ and $\mathbf{x}$.

Affine approximation

- Suppose that the initial state of the FSM, $\mathbf{x}$ is randomly chosen and we compute $f(L^i(\mathbf{x})) = z_i$ for $i = 0, 1, \ldots M$.
- The probability that the above system of equations is same as $\ell(L^i(\mathbf{x})) = z_i$ for $i = 0, 1, \ldots, M$ is $(1 - \frac{1}{16})^M$. (Of course under certain assumptions)
- Since the second system is linear it can be solved for the initial state $\mathbf{x}$.
- Thus the function f must be "far away" from all the affine functions

$$a_1x_1 \oplus a_2x_2 \oplus \ldots \oplus a_nx_n \oplus \epsilon,$$

for all $(a_1, \ldots, a_n) \in GF(2)^n$ and $\epsilon \in GF(2)$. In the present case $n = 4$.
- This can also be written as

$$\ell_{\mathbf{a},\epsilon}(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} \oplus \epsilon,$$

where $\mathbf{a} \cdot \mathbf{x} = \bigoplus_{i=1}^{n} a_ix_i$ is the inner product of $\mathbf{a}$ and $\mathbf{x}$.

Affine approximation

- Suppose that the initial state of the FSM, $\mathbf{x}$ is randomly chosen and we compute $f(L^i(\mathbf{x})) = z_i$ for $i = 0, 1, \dots M$.
- The probability that the above system of equations is same as $\ell(L^i(\mathbf{x})) = z_i$ for $i = 0, 1, \dots, M$ is $(1 - \frac{1}{16})^M$. (Of course under certain assumptions)
- Since the second system is linear it can be solved for the initial state $\mathbf{x}$.
- Thus the function f must be "far away" from all the affine functions

  $$a_1x_1 \oplus a_2x_2 \oplus \dots \oplus a_nx_n \oplus \epsilon,$$

  for all $(a_1, \dots, a_n) \in GF(2)^n$ and $\epsilon \in GF(2)$. In the present case $n = 4$.
- This can also be written as

  $$\ell_{\mathbf{a},\epsilon}(\mathbf{x}) = \mathbf{a} \cdot \mathbf{x} \oplus \epsilon,$$

  where $\mathbf{a} \cdot \mathbf{x} = \bigoplus_{i=1}^{n} a_ix_i$ is the inner product of $\mathbf{a}$ and $\mathbf{x}$.

The distance between two Boolean functions

- The distance between two Boolean functions f and g is

$$
\begin{aligned}
d_H(f,g) &= \#(f \neq g) \\
&= \frac{1}{2}(\#(f = g) + \#(f \neq g)) - \frac{1}{2}(\#(f = g) - \#(f \neq g)) \\
&= 2^{n-1} - \frac{1}{2} \sum_{\mathbf{x} \in GF(2)^n} (-1)^{f(x) \oplus g(x)}
\end{aligned}
\tag{4}
$$

-

$$
\begin{aligned}
d_H(f, \ell_{\mathbf{a},\epsilon}) &= 2^{n-1} - \frac{1}{2} \sum_{\mathbf{x} \in GF(2)^n} (-1)^{f(x) \oplus \ell_{\mathbf{a},\epsilon}(x)} \\
&= 2^{n-1} - \frac{1}{2} \sum_{\mathbf{x} \in GF(2)^n} (-1)^{f(x) \oplus \mathbf{a} \cdot \mathbf{x} \oplus \epsilon} \\
&= 2^{n-1} - (-1)^{\epsilon} \frac{1}{2} \sum_{\mathbf{x} \in GF(2)^n} (-1)^{f(x) \oplus \mathbf{a} \cdot \mathbf{x}}
\end{aligned}
\tag{5}
$$

Hamming distance and Walsh–Hadamard transform

- $$\begin{aligned} d_H(f,\ell_{\mathbf{a},\epsilon}) &= 2^{n-1} - (-1)^{\epsilon}\frac{1}{2}\sum_{\mathbf{x}\in GF(2)^n}(-1)^{f(x)\oplus\mathbf{a}\cdot\mathbf{x}} \\ &= 2^{n-1} - (-1)^{\epsilon}\frac{1}{2}W_f(\mathbf{a}). \end{aligned} \tag{6}$$
- $\min_{\epsilon\in GF(2)}(d_H(f,\ell_{\mathbf{a},\epsilon})) = 2^{n-1} - \frac{1}{2}|W_f(\mathbf{a})|$.
- The nonlinearity of f is defined as:

$$\min_{\mathbf{a}\in GF(2)^n}\min_{\epsilon\in GF(2)}(d_H(f,\ell_{\mathbf{a},\epsilon})) = 2^{n-1} - \frac{1}{2}\max_{\mathbf{a}\in GF(2)^n}|W_f(\mathbf{a})|.$$

Parseval's equation

- $$\begin{aligned} \sum_{\mathbf{a}\in GF(2)^n} W_f(\mathbf{a})^2 &= \sum_{\mathbf{a}\in GF(2)^n} \sum_{\mathbf{x}\in GF(2)^n} \sum_{\mathbf{y}\in GF(2)^n} (-1)^{f(\mathbf{x})\oplus f(\mathbf{y})\oplus \mathbf{a}\cdot(\mathbf{x}+\mathbf{y})} \\ &= \sum_{\mathbf{x}\in GF(2)^n} \sum_{\mathbf{y}\in GF(2)^n} (-1)^{f(\mathbf{x})\oplus f(\mathbf{y})} \\ &\quad \times \sum_{\mathbf{a}\in GF(2)^n} (-1)^{\mathbf{a}\cdot(\mathbf{x}+\mathbf{y})} \\ &= 2^{2n}.(\text{Why!}) \end{aligned} \tag{7}$$
- Therefore $W_f(\mathbf{a}) \geq 2^{\frac{n}{2}}$, which implies that nonlinearity of f is bounded above by $2^{n-1} - 2^{\frac{n}{2}-1}$.

Bent functions - the functions with maximum nonlinearity

- Consider $f(x, y) = x \cdot y$ and $n = 2t$
-

$$
\begin{aligned}
W_f(\mathbf{a}, \mathbf{b}) &= \sum_{\mathbf{x} \in GF(2)^t} \sum_{\mathbf{y} \in GF(2)^t} (-1)^{\mathbf{x} \cdot \mathbf{y} \oplus \mathbf{a} \cdot \mathbf{x} \oplus \mathbf{b} \cdot \mathbf{y}} \\
&= \sum_{\mathbf{x} \in GF(2)^t} (-1)^{\mathbf{x} \cdot \mathbf{a}} \sum_{\mathbf{y} \in GF(2)^t} (-1)^{\mathbf{y} \cdot (\mathbf{x} \oplus \mathbf{b})} \qquad (8) \\
&= (-1)^{\mathbf{a} \cdot \mathbf{b}} 2^t .
\end{aligned}
$$

- These functions where first constructed by Rothaus in 1966.
- These are called bent functions.

The case when n is odd

- For a long time it was believed that the maximum nonlinearity possible for odd n is $2^{n-1} - 2^{\frac{n-1}{2}}$.
- It was known that such is the case for $n \leq 7$.
- It was shown by Patterson and Wiedemann in 1983 that for $n = 15$ there are functions with larger value of nonlinearity.
- The cases $n = 9, 11, 13$ remained open for long time and finally settled in the following paper:
  Kavut S., Maitra S., Sarkar S., Yücel M. D., Enumeration of 9-variable Rotation Symmetric Boolean Functions having Nonlinearity $>$ 240, INDOCRYPT 2006, pp. 266-279, 2006.

Other important properties

- A Boolean function f is said to be α-*resilient* (i.e., f has *resiliency order* α) if it is balanced (i.e., $d_H(f, \mathbf{0}) = 2^{n-1}$, where $\mathbf{0}$ is the constant function with output 0) and remains balanced as an $(n-t)$-variable function if any set of $t \leq \alpha$ variables are fixed. This definition of resiliency is due to Siegenthaler:
  T. Siegenthaler, "Correlation-immunity of nonlinear combining functions for cryptographic applications", IEEE Trans. Inform. Theory, *vol. 30, no. 5, pp. 776-780, 1984.*

Resiliency and Walsh–Hadamard transform

- A Boolean function $f \in \mathcal{B}_n$ is α-resilient if and only if

$$W_f(\mathbf{u}) = 0 \tag{9}$$

  for all $\mathbf{u} \in GF(2)^n$ such that $wt(\mathbf{u}) \leq \alpha$.
- This is proved by
  Xiao Guo-Zhen and J. L. Massey, "A Spectral Characterization of Correlation-Immune Combining Functions", IEEE Trans. Inform. Theory, *vol. 34, no. 3, pp. 569-571, 1988.*

Algebraic degree

- The algebraic normal form (ANF) of a Boolean function f is as follows:

$$f(x_1, \ldots, x_n) = \bigoplus_{\mathbf{a}=(a_1,\ldots,a_n)\in GF(2)^n} \mu_{\mathbf{a}} \prod_{i=1}^{n} x_i^{a_i}.$$

- The algebraic degree of f,

$$\deg(f) := \max_{\mathbf{a}\in GF(2)^n} \{wt(\mathbf{a}) : \mu_{\mathbf{a}} \neq 0\}.$$

Optimization of nonlinearity, resiliency and algebraic degree

- Then Siegenthaler proved that deg $\leq n - \alpha - 1$. This is known as the *Siegenthaler's bound*.

- Nonlinearity of f is bounded above by $2^{n-1} - 2^{\alpha+1}$, which is known as the *Sarkar and Maitra's bound*.

  S. Maitra and P. Sarkar, "Highly nonlinear resilient functions optimizing Siegnethaler's Inequality", CRYPTO'99, Lecture Notes in Comput. Sci.*, vol. 1666, pp. 198-215, 1999.*

- Sarkar and Maitra obtained the optimal functions with respect to nonlinearity, resiliency and algebraic degree.

Algebraic immunity

- Another class of attack was introduced by Courtois, Meier and others which is known as algebraic attack.

- The property of Boolean function which has to be maximized in order to resist algebraic attack is said to be *algebraic immunity*.

- The Boolean functions with maximum algebraic immunity were constructed for the first time by Dalai, Gupta and Maitra.
  Dalai D. K., Gupta K. C., Maitra S., Results on Algebraic Immunity for Cryptographically Significant Boolean Functions, INDOCRYPT 2004, Lecture Notes in Computer Science, Vol. 3348, pp. 92-106, 2004.

Is it the end of the story of cryptographically significant Boolean functions?

- A Boolean function $f \in \mathcal{B}_n$ is said to be k-normal (weakly k-normal) if its restriction on a k-dimensional flat is constant (affine). The positive integer k is said to be the normality order of the Boolean function f.
- Mihaljević, Gangopadhyay, Paul and Imai demonstrated that in some cases even the optimal functions constructed above are vulnerable to dedicated time-memory-date-tradeoff attacks due to high order of normality.

Normality-order as a cryptographically significant property

- *Mihaljevic M. J., Gangopadhyay S., Paul G. and Imai H., An Algorithm for the Internal State Recovery of Grain-v1, presented in the 11th Central European Conference on Cryptology, Debrecen, Hugary, 30 June to 2 July, 2011.*
- *Mihaljevic M. J., Gangopadhyay S., Paul G. and Imai H., A Generic Weakness of the k-normal Boolean Functions Exposed to Dedicated Algebraic Attack, 2010 Int. Symp. on Inform. Theory and its Appl. - ISITA 2010, Taichung, Taiwan, Oct. 17-20, 2010, IEEE Proceedings, pp. 911-916. (IEEE Catalog Number: CFP 10767-USB, ISBN: 078-1-4244-6014-4, ISSN: 1943-7439)*

Normality-order as a cryptographically significant property

- It is observed that all the standard constructions of optimal functions have high normality-order.
- Thus it seems that the story is not complete yet.

Thank you
Questions Please!