T-79.503 Fundamentals of Cryptology
Additional material (Handout 3)
for lectures October 15-22, 2003

1 Euler Phi-Function

In section 1.1.3 of the text-book, Definition 1.3, the Euler phi-function is defined as follows.

Definition 1.3 (Stinson) Suppose $a \geq 1$ and $m \geq 2$ are integers. If $\gcd(a, m) = 1$ then we say that a and m are relatively prime. The number of integers in $\mathbb{Z}_m$ that are relatively prime to m is denoted by $\phi(m)$.

We set $\phi(1) = 1$. The function

$$m \mapsto \phi(m), \ m \geq 1$$

is called the Euler phi-function, or Euler totient function. Clearly, for m prime, we have $\phi(m) = m - 1$. Further, we state the following fact without proof, and leave the proof as an easy exercise.

Fact. If m is a prime power, say, $m = p^e$, where p is prime and $p > 1$, then $\phi(m) = m(1 - \frac{1}{p}) = p^e - p^{e-1}$.

The main purpose of this section is to prove the multiplicative property of the Euler phi-function.

Proposition. Suppose that $m \geq 1$ and $n \geq 1$ are integers such that $\gcd(m, n) = 1$. Then $\phi(m \times n) = \phi(m) \times \phi(n)$.

Proof. If $m = 1$ or $n = 1$, then the claim holds. Suppose now that $m > 1$ and $n > 1$, and denote:

$$\begin{aligned} A &= \{a \,|\, 1 \leq a < m, \ \gcd(a, m) = 1\} \\ B &= \{b \,|\, 1 \leq b < n, \ \gcd(b, n) = 1\} \\ C &= \{c \,|\, 1 \leq c < m \times n, \ \gcd(c, m \times n) = 1\}. \end{aligned}$$

Then we have that $|A| = \phi(m)$, $|B| = \phi(n)$, and $|C| = \phi(m \times n)$. We show that C has equally many elements as the set $A \times B = \{(a, b) \,|\, a \in A, b \in B\}$, from which the claim follows.

Since $\gcd(m, n) = 1$, we can use the Chinese Remainder Theorem, by which the mapping

$$\pi : \mathbb{Z}_{mn} \to \mathbb{Z}_m \times \mathbb{Z}_n, \ \pi(x) = (x \bmod m, x \bmod n)$$

is bijective. Now we observe that $A \subset \mathbb{Z}_m$, $B \subset \mathbb{Z}_n$, and $C \subset \mathbb{Z}_{m \times n}$. Moreover, it holds that $x \in C$ if and only if $\pi(x) \in A \times B$, which we see by writing the following chain of equivalences:

$$\begin{aligned} \gcd(x, m \times n) = 1 &\Leftrightarrow \gcd(x, m) = 1 \text{ and } \gcd(x, n) = 1 \\ &\Leftrightarrow \gcd(x \bmod m, m) = 1 \text{ and } \gcd(x \bmod n, n) = 1. \end{aligned}$$

$\square$

As a corollary, we get Theorem 1.2 of the textbook.

Theorem 1.2 Suppose

$$m = \prod_{i=1}^{k} p_i^{e_i},$$

where the integers p_i are distinct primes and $e_i > 0$, $1 \le i \le k$. Then

$$\phi(m) = \prod_{i=1}^{k} (p_i^{e_i} - p_i^{e_i - 1}).$$

2 Algebraic Normal Form of a Boolean function

Let us now consider a function $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$. Such a function is called a Boolean function of n variables. A Boolean function of n variables $x_1, \ldots, x_n$ has a unique representation in its algebraic normal form

$$\begin{aligned} g(x_1, \ldots, x_n) &= a_0 \oplus a_1 x_1 \oplus \cdots \oplus a_n x_n \oplus a_{12} x_1 x_2 \oplus \cdots \\ &\quad \cdots \oplus a_{(n-1)n} x_{n-1} x_n \oplus a_{123} x_1 x_2 x_3 \oplus \cdots \oplus a_{12\ldots n} x_1 x_2 \cdots x_n. \end{aligned}$$

with coefficients $a_{i_1,\ldots,i_k} \in \mathbb{Z}_2$.

Given the values of the function f, its algebraic normal form ANF(f) can be derived using the following algorithm:

ANF Algorithm.

1. Set $g(x_1, \ldots, x_n) = f(0, 0, \ldots.0)$
2. For $k = 1$ to $2^n - 1$, do
3. use the binary representation of the integer k,
$$k = b_1 + b_2 2 + b_3 2^2 + \cdots + b_n 2^{n-1}$$
4. if $g(b_1, b_2, \ldots, b_n) \ne f(b_1, b_2, \ldots, b_n)$ then
set $g(x_1, \ldots, x_n) = g(x_1, \ldots, x_n) \oplus \prod_{i=1}^{n} (x_i)^{b_i}$
5. ANF(f) $= g(x_1, \ldots, x_n)$

Example 4.

k	b_3	b_2	b_1	$f(b_1, b_2, b_3)$	$g(b_1, b_2, b_3)$
	0	0	0	0	0
1	0	0	1	0	0
2	0	1	0	1	x_2
3	0	1	1	0	$x_2 \oplus x_1 x_2$
4	1	0	0	1	$x_2 \oplus x_1 x_2 \oplus x_3$
5	1	0	1	1	$x_2 \oplus x_1 x_2 \oplus x_3$
6	1	1	0	0	$x_2 \oplus x_1 x_2 \oplus x_3$
7	1	1	1	1	$x_2 \oplus x_1 x_2 \oplus x_3$

Now the values of f given in the third column of the table can also be calculated from the expression $f(x_1, x_2, x_3) = x_2 \oplus x_3 \oplus x_1 x_2$.

3 Non-linearity of Boolean Functions

3.1 Correlations

Let $x = (x_1, \ldots, x_m) \in \mathbb{Z}_2^m$. The *Hamming weight* of x is defined as

$$H_W(x) = |\{i \in \{1, 2, \ldots, m\} \mid x_i = 1\}|.$$

For two vectors $x = (x_1, \ldots, x_m) \in \mathbb{Z}_2^m$ and $y = (y_1, \ldots, y_m) \in \mathbb{Z}_2^m$ the *Hamming distance* is defined as

$$d_H(x, y) = H_W(x \oplus y) = |\{i \in \{1, 2, \ldots, m\} \mid x_i \neq y_i\}|.$$

Given two Boolean functions $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ and $g : \mathbb{Z}_2^n \to \mathbb{Z}_2$ the *Hamming weight* of f is defined as

$$H_W(f) = |\{x \in \mathbb{Z}_2^n \mid f(x) = 1\}|,$$

and the *Hamming distance* between f and g is

$$d_H(f, g) = |\{x \in \mathbb{Z}_2^n \mid f(x) \neq g(x)\}|.$$

A Boolean function $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ is *balanced* if $H_W(f) = 2^{n-1}$, which happens if and only if

$$|\{x \in \mathbb{Z}_2^n \mid f(x) = 1\}| = |\{x \in \mathbb{Z}_2^n \mid f(x) = 0\}|.$$

Example 5. Let $f_{00} : \mathbb{Z}_2^4 \to \mathbb{Z}_2$ be the Boolean function defined as the first outputbit of the s-box S_1 of the DES, when the first and the last (sixth) input bits are set equal to zero. Then f_{00} has the following values

$$f_{00} = (1, 0, 1, 0, 0, 1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 0)$$

arranged in the *lexicographical order* with respect to the input (x_2, x_3, x_4, x_5). Clearly, f_{00} is balanced, that is, $H_W(f_{00}) = 8$. Further we see that

$$d_H(f_{00}, s_5) = 6, \text{ and } d_H(f_{00}, s_2) = 10,$$

where we have denoted by s_i the ith input bit to S_1 as a Boolean function of the four middle input bits. That is, $s_i(x_2, x_3, x_4, x_5) = x_i$, for $i = 2, 3, 4, 5$.

Let $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ and $g : \mathbb{Z}_2^n \to \mathbb{Z}_2$ be two Boolean functions. The *correlation* between f and g is defined as

$$\begin{aligned} c(f, g) &= 2^{-n}(|\{x \in \mathbb{Z}_2^n \mid f(x) = g(x)\}| - |\{x \in \mathbb{Z}_2^n \mid f(x) \neq g(x)\}|) \\ &= 2^{-n}(2^n - 2|\{x \in \mathbb{Z}_2^n \mid f(x) \neq g(x)\}|) = 1 - 2^{1-n} d_H(f, g). \end{aligned}$$

A Boolean function $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ is *linear* if it has an ANF of the form

$$f(x) = a \cdot x = a_1 x_1 \oplus a_2 x_2 \oplus \cdots \oplus a_n x_n$$

for some $a = (a_1, a_2, \ldots, a_n) \in \mathbb{Z}_2^n$. Then f is just a linear combination of its input bits. In such a case we denote $f = L_a$. A Boolean function is *affine* if it has an ANF of the form $f(x) = a \cdot x \oplus 1$.

Nonlinearity of a Boolean function $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ is defined as its minimum distance from the set consisting all affine and linear Boolean functions

$$\mathcal{N}(f) = \min_{L \text{ linear}}\{\min\{d_H(f, L), d_H(f, L \oplus 1)\}\}.$$

Example 5(continued)
From $d_H(f_{00}, s_5) = 6$ and $d_H(f_{00}, s_2) = 10$, it follows that the nonlinearity of f is at most 6. Further we see that

$$\begin{aligned} c(f_{00}, s_5) &= 1 - \frac{1}{8} \cdot 6 = \frac{1}{4}, \text{ and} \\ c(f_{00}, s_2) &= 1 - \frac{10}{8} = -\frac{1}{4}. \end{aligned}$$