1차 상관면역 함수의 개수

지성택°† 이상진† 박춘식† 성수학‡ ‡

†한국전자통신연구소 ‡배재대학교

Enumerating 1st order Correlation Immune Functions

Seongtaek Chee† Sangjin Lee† Choonsik Park† Soohak Sung‡ ‡

†Electronics and Telecommunications Research Institute ‡Paichai Univ.

요 약

본 논문에서는 새로운 방식으로 1차 상관면역 함수를 설계하는 방법을 제시하고 이 방법을 이용하여 1차 상관면역 함수 개수의 하한값을 구한다. 이 값은 1차 상관면역 함수의 개수에 대한 기존의 제시된 결과를 크게 개선한 값이다.

1 서 론

암호학적으로 우수한 성질을 가지는 부울함수는 스트림 암호와 블럭 암호의 핵심 논리로 사용될 수 있다. 어떤 성질이 좋은 부울함수를 판단하는 기준이라면 그러한 성질을 만족하는 부울함수는 충분히 많아야 한다. 따라서 암호학적으로 중요한 성질을 만족하는 부울함수가 얼마나 되는지를 아는 것은 중요하며 본 논문에서는 이러한 연구를 하고자 한다. 중요한 대표적인 성질로는 균형성(Balance), 비선형성(Nonlinearity), Nondegeneracy, 상관면역(Correlation Immune), 대칭성(Symmetry) 등이 있다. 언급된 성질 중 상관면역을 제외하고는 그러한 성질을 만족하는 부울함수의 개수는 이미 잘 알려져 있다[2],[7]. 상관면역 성질을 만족하는 부울함수의 개수는 정확히 모르나 하한값과 상한값은 알려져 있다[2],[7]. 상관면역 함수는 Siegenthaler[4]에 의해서 소개된 이후 많은 사람들이 연구하였다[1],[3],[5],[6],[8]. 이렇듯 상관면역 함수가 활발히 연구된 이유 중의 하나는 지금까지 많은 암호 시스템이 상관공격에 의해서 해독되었기 때문이다.

상관면역 함수를 설계하는 방법은 크게 두 가지로 나눌 수 있다. Siegenthaler의 방법처럼 상관면역 함수를 이용하여 새로운 상관면역 함수를 설계하는 귀납적(recusive)인 방법과, Camion 등[1]의 방법처럼 직접 상관면역 함수를 설계하는 방법이 있다. Mitchell과 Yang-Guo는 상관면역 함수를 직접 설계하는 방법으로 상관면역 함수 개수의 하한값과 상한값을 구하였다. Yang-Guo는 Mitchell이 구한 하한값을 개선하였으나, 두 하한값의 수렴 속도가 같으므로 많은 개선은 아니다.

본 논문에서는 상관면역 함수를 설계하는 새로운 귀납적인 방법을 이용하여 n(부울함수의 정의역의 차원)이 6이하일 때는 상관면역 함수를 모두 찾으며, n이 7이상 일때는 상관면역함수의 개수에 대한 Yang-Guo가 얻은 결과를 크게 개선하고자 한다.

2 기본적인 정의

n 차원의 벡터공간 $\{0,1\}^n$을 $\mathbb{Z}_2^n$으로 쓰기로 하며, $\mathbb{Z}_2^n$ 상의 벡터를 $x=(x_1,\cdots,x_n)$로 쓰며 n 개의 변수를 갖는 부울함수를 $f(x)=f(x_1,\cdots,x_n)$ 또는 $f:\mathbb{Z}_2^n\to\mathbb{Z}_2$로 쓰며 f를 $\mathbb{Z}_2^n$ 상의 부울함수라고 부르기로 한다. 또 두 벡터 $x=(x_1,\cdots,x_n), y=(y_1,\cdots,y_n)$의 내적을 $x\cdot y$로 표시하며 $x\cdot y=x_1y_1\oplus\cdots\oplus x_ny_n$으로 정의한다. $x_i=1$이고 $x_j=0(j\neq i)$인 $\mathbb{Z}_2^n$ 상의 벡터를 e_i^n으로 나타내기로 한다. 즉

$$e_i^n=(0,\cdots,0,1,0,\cdots,0),\quad 1\le i\le n$$

$\mathbb{Z}_2^n$ 상의 벡터 x의 Hamming 가중치를 $wt(x)$로 쓴다. 부울함수 f의 n개의 변수 중 $k(1\le k\le n)$개의 변수 $x_{i_1},\cdots,x_{i_k}(1\le i_1<\cdots<i_k\le n)$와 함수값이 독립일 때 f를 k차 상관면역이라고 한다. 본 논문에서는 상관면역의 정의를 Hadamard-Walsh 변환으로 된 것을 사용하고자 한다. 이러한 정의는 이미 잘 알려져 있으나 본 논문에서 많이 사용되므로 다시 언급하고자 한다.

Hadamard-Walsh 변환은 부울함수와 같은 정의역 상에서 정의되나 그 값은 실수값을 갖는 함수로 아래와 같이 정의한다.

정의 2.1 부울함수 f의 *Hadamard-Walsh* 변환을 $\widehat{(-1)^f}$로 표시하며 다음과 같이 정의한다.

$$\widehat{(-1)^f}(w)=\sum_x(-1)^{f(x)}(-1)^{w\cdot x}$$

Hadamard-Walsh 변환을 이용하여 상관면역에 대한 동치 정의를 얻을 수 있다.

정의 2.2 *Hamming* 가중치가 *1*과 k 사이인 임의 벡터 α, 즉 $1\le wt(\alpha)\le k$에 대해 $\widehat{(-1)^f}(\alpha)=0$인 함수 f를 k차 상관면역이라고 한다. 특히 $k=1$일 때는 간단히 상관면역이라고 부르기로 한다.

3 새로운 부울함수의 설계

이 절에서는 귀납적인 방법으로 부울함수를 설계하는 새로운 방법을 제시한다. 즉 두 개의 부울함수를 이용하여 새로운 부울함수를 설계한다.

f와 g가 $\mathbb{Z}_2^n$ 상의 부울함수일 때 $\mathbb{Z}_2^{n+1}$ 상의 새로운 부울함수 h를 다음과 같이 정의한다.

$$h(x_1,\cdots,x_n,x_{n+1})=\begin{cases}f(x_1,x_3,\cdots,x_{n+1}), & x_1=0,x_2=0,\\ g(x_1,x_3,\cdots,x_{n+1}), & x_1=0,x_2=1,\\ g(x_1,x_3,\cdots,x_{n+1}), & x_1=1,x_2=0,\\ f(x_1,x_3,\cdots,x_{n+1}), & x_1=1,x_2=1\end{cases}$$

즉,

$$\begin{aligned} h(x_1,\cdots,x_n,x_{n+1}) &= (1\oplus x_1)(1\oplus x_2)f(x_1,x_3,\cdots,x_{n+1}) \\ &\oplus(1\oplus x_1)x_2 g(x_1,x_3,\cdots,x_{n+1}) \\ &\oplus x_1(1\oplus x_2)g(x_1,x_3,\cdots,x_{n+1}) \\ &\oplus x_1x_2 f(x_1,x_3,\cdots,x_{n+1}) \end{aligned} \quad (1)$$

이다. 이 때 h를 $< f,g >$로 쓰기로 한다.

[주] $< f,g >$는 $\mathbb{Z}_2^{n+1}$ 상의 모든 부울함수를 생성한다. 즉 $\mathbb{Z}_2^n$ 상의 모든 부울함수의 집합을 Ω_n이라고 표시하면 다음과 같다.

$$\Omega_{n+1} = \{< f,g > | f,g \in \Omega_n\}$$

$\mathbb{Z}_2^n$ 상의 부울함수를 이용하여 $\mathbb{Z}_2^{n+1}$ 상의 모든 부울함수를 생성할 수 있기 때문에 $n+1$이 작을 때$(n+1 \le 6)$ $\mathbb{Z}_2^{n+1}$상의 모든 상관면역 함수를 쉽게 찾을 수 있다.

Hadamard-Walsh 변환의 정의에 의해서 $h =< f,g >$의 Hadamard-Walsh 변환은 다음과 같이 쓸 수 있다.

$$\begin{aligned} \widehat{(-1)^h} \quad & (w_1,\cdots,w_n,w_{n+1}) \\ = \quad & \sum_{x_1=0,x_2=0} (-1)^{f(x_1,x_3,\cdots,x_{n+1})}(-1)^{w_1x_1\oplus w_3x_3\oplus\cdots\oplus w_{n+1}x_{n+1}} \\ & + \sum_{x_1=0,x_2=1} (-1)^{g(x_1,x_3,\cdots,x_{n+1})}(-1)^{w_1x_1\oplus w_2x_2\oplus\cdots\oplus w_{n+1}x_{n+1}} \\ & + \sum_{x_1=1,x_2=0} (-1)^{g(x_1,x_3,\cdots,x_{n+1})}(-1)^{w_1x_1\oplus w_3x_3\oplus\cdots\oplus w_{n+1}x_{n+1}} \\ & + \sum_{x_1=1,x_2=1} (-1)^{f(x_1,x_3,\cdots,x_{n+1})}(-1)^{w_1x_1\oplus w_2x_2\oplus\cdots\oplus w_{n+1}x_{n+1}} \end{aligned} \quad (2)$$

여기서 $\sum_{x_1=a_1,x_2=a_2}$은 $x_1=a_1, x_2=a_2$인 $\mathbb{Z}_2^{n+1}$상의 모든 벡터 $(x_1,\cdots,x_n,x_{n+1})$에 대해서 더하는 것을 나타낸다. $(-1)^h$의 Hadamard-Walsh 변환의 값은 위와 같이 복잡하게 쓸 수 밖에 없으나 w의 성분 중 w_2가 0일 때는 아주 간단히 쓸 수 있다.

보조정리 3.1 $h =< f,g >$가 *(1)*과 같이 정의되었을 때 $w_2=0$인 w에 대한 $(-1)^h$의 *Hadamard-Walsh* 변환값은 다음과 같다.

$$\widehat{(-1)^h}(w_1,0,w_3,\cdots,w_{n+1}) = \widehat{(-1)^f}(w_1,w_3,\cdots,w_{n+1}) + \widehat{(-1)^g}(w_1,w_3,\cdots,w_{n+1})$$

(증명) $w_2=0$일 때 식 (2)의 첫째식과 넷째식을 합하면

$$\begin{aligned} & \sum_{x_1=0,x_2=0} (-1)^{f(x_1,x_3,\cdots,x_{n+1})}(-1)^{w_1x_1\oplus w_3x_3\oplus\cdots\oplus w_{n+1}x_{n+1}} \\ & + \sum_{x_1=1,x_2=1} (-1)^{f(x_1,x_3,\cdots,x_{n+1})}(-1)^{w_1x_1\oplus w_3x_3\oplus\cdots\oplus w_{n+1}x_{n+1}} \\ = \quad & \widehat{(-1)^f}(w_1,w_3,\cdots,w_{n+1}) \end{aligned}$$

이며, 같은 방법으로 $w_2 = 0$일 때 식 (2)의 둘째식과 셋째식을 합하면

$$\widehat{(-1)}^g(w_1, w_3, \cdots, w_{n+1})$$

이다. 따라서 증명이 완성된다. □

$h =< f, g >$가 상관면역인지 비상관면역인지를 판정하기 위해서는 가중치가 1인 벡터 $w \in \mathbb{Z}_2^{n+1}$에 대해서만 $(-1)^h$의 Hadamard-Walsh 변환값을 구하면 된다. 기호를 간단히 쓰기 위해서 이미 제2절에서 언급하였드시 e_i^n은 i번째 성분은 1이고 나머지 성분은 0인 $\mathbb{Z}_2^n$ 상의 벡터를 나타낸다.

보조정리 3.2 $h =< f, g >$가 *(1)*과 같이 정의되었을 때 *Hamming* 가중치가 *1*인 $\mathbb{Z}_2^{n+1}$ 상의 벡터 $e_i^{n+1}(1 \le i \le n+1)$에 대해 $(-1)^h$의 *Hadamard-Walsh* 변환값은 다음과 같다.

$$\begin{aligned}
\widehat{(-1)}^h(e_1^{n+1}) &= \widehat{(-1)}^f(e_1^n) + \widehat{(-1)}^g(e_1^n) \\
\widehat{(-1)}^h(e_2^{n+1}) &= \widehat{(-1)}^f(e_1^n) - \widehat{(-1)}^g(e_1^n) \\
\widehat{(-1)}^h(e_i^{n+1}) &= \widehat{(-1)}^f(e_{i-1}^n) + \widehat{(-1)}^g(e_{i-1}^n), 3 \le i \le n+1
\end{aligned}$$

(증명) $e_i^{n+1}(1 \le i \le n+1)$ 중 e_2^{n+1}를 제외한 나머지 것들은 둘째 성분이 0인 벡터이다. 따라서 첫째식과 셋째식은 보조정리 3.1에 의해서 바로 유도된다. 이젠 둘째 식을 증명해 보자. 식 (2)에 의해서

$$\begin{aligned}
\widehat{(-1)}^h(0,1,0,\cdots,0) = & \sum_{x_1=0,x_2=0} (-1)^{f(x_1,x_3,\cdots,x_{n+1})} \\
& - \sum_{x_1=0,x_2=1} (-1)^{g(x_1,x_3,\cdots,x_{n+1})} \\
& + \sum_{x_1=1,x_2=0} (-1)^{g(x_1,x_3,\cdots,x_{n+1})} \\
& - \sum_{x_1=1,x_2=1} (-1)^{f(x_1,x_3,\cdots,x_{n+1})}
\end{aligned} \tag{3}$$

이다. (3)의 오른쪽의 첫째식과 넷째식을 더하면

$$\sum_{x_1,x_3,\cdots,x_{n+1}} (-1)^{f(x_1,x_3,\cdots,x_{n+1})}(-1)^{x_1} = \widehat{(-1)}^f(e_1^n)$$

이고 둘째식과 셋째식을 더하면 $-\widehat{(-1)}^g(e_1^n)$이므로

$$\widehat{(-1)}^h(0,1,0,\cdots,0) = \widehat{(-1)}^f(e_1^n) - \widehat{(-1)}^g(e_1^n)$$

이다. 따라서 증명이 완성된다. □

4 상관면역 함수 개수의 하한값

상관면역 함수 개수의 하한값을 구하기 위해서 $h =< f, g >$가 상관면역 함수가 될 조건을 찾아보자.

정리 4.1 f와 g가 $\mathbb{Z}_2^n$ 상에서 정의된 상관면역 함수이면 $h =< f, g >$도 $\mathbb{Z}_2^{n+1}$ 상의 상관면역 함수이다.

(증명) h가 상관면역 함수임을 증명하기 위해서 $\mathbb{Z}_2^{n+1}$ 상의 Hamming 가중치가 1인 벡터 $e_i^{n+1}(1 \leq i \leq n+1)$에 대해 $(-1)^h$의 Hadamard-Walsh 변환값이 0임을 증명하면 된다. 만일 f와 g가 $\mathbb{Z}_2^n$ 상에서 정의된 상관면역 함수이면 $\widehat{(-1)^f}(e_i^n) = 0(1 \leq i \leq n)$, $\widehat{(-1)^g}(e_i^n) = 0(1 \leq i \leq n)$이다. 따라서 보조정리 3.2에 의해서 $\widehat{(-1)^h}(e_i^{n+1}) = 0(1 \leq i \leq n+1)$ 이다. 즉 $h =< f, g >$도 $\mathbb{Z}_2^{n+1}$ 상의 상관면역 함수이다. □

[주] 정리 4.1은 귀납적인 방법으로 새로운 상관면역 함수를 설계하는 방법이다.

$\mathbb{Z}_2^n$ 상의 상관면역 함수의 전체 집합을 A_n이라고 두자. 즉

$$A_n = \{f : \mathbb{Z}_2^n \rightarrow \mathbb{Z}_2 | f \text{ 는 상관면역 함수}\}$$

또 집합 S의 원소의 개수를 $|S|$ 또는 $\sharp S$로 표시한다.

정리 4.2 $|A_{n+1}| \geq |A_n|^2$이다.

(증명) 함수 $F : A_n \times A_n \rightarrow A_{n+1}$를 다음과 같이 정의한다.

$$F(f, g) =< f, g >$$

그러면 정리 4.1에 의해서 F는 잘 정의된다. 또한 F는 1-1 함수이다. 왜냐하면, 만일 $F(f, g) = F(f', g')$(즉 $< f, g >=< f', g' >$)이면 $<, >$의 정의에 의해서 $f = f'$이고 $g = g'$이기 때문이다. F가 1-1 함수이므로

$$|A_{n+1}| \geq |A_n \times A_n| = |A_n|^2$$

이다. 따라서 증명이 완성된다. □

따름정리 4.1 $|A_n| \geq |A_{n-k}|^{2^k}$이다.

(증명) A_n을 정리 4.2에 적용하면 $|A_n| \geq |A_{n-1}|^2$이고, 다시 A_{n-1}을 정리 4.2에 적용하면

$$|A_n| \geq |A_{n-1}|^2 \geq |A_{n-2}|^{2^2}$$

이다. 정리 4.2를 계속 적용하면(총 k번) $|A_n| \geq |A_{n-k}|^{2^k}$을 얻을 수 있다. □

제 3절에서 제안된 부울함수의 설계 방법을 이용하여 n이 6이하일 때 $\mathbb{Z}_2^n$ 상의 모든 상관면역 함수를 쉽게 찾을 수 있다.

$$|A_1| = 2, |A_2| = 4, |A_3| = 18, |A_4| = 648, |A_5| = 3,140,062, |A_6| = 503,483,766,022,188$$

정리 4.3 $n \geq 7$일 때 $|A_n| \geq (503,483,766,022,188)^{2^{n-6}}$ 이다.

(증명) $|A_6| = 503,483,766,022,188$이므로 따름정리 4.1에서 k대신 $n-6$를 대입하면 바로 증명된다. □

Mitchell(1990)은 $|A_n|$의 하한값이 $2^{2^{n-1}}$, Yang-Guo(1995)는 $2^{2^{n-1}} + 2^n - 2n + 2^{2^{n-4}} - 2^{n-3}$을 얻었다. 일반적으로 $|A_n|$은 2^{2^n}보다 클수 없으므로 $|A_n| = (2^{2^{n-1}})^c$을 만 족하는 $c \in (1,2)$가 존재하는데 이를 기준으로 $|A_n|$의 하한값을 비교하면 표 1과 같다.

표 1: 하한값$((2^{2^{n-1}})^c)$의 비교, $\lim_{n\to\infty} \varepsilon = 0$

	하한값	c
Mitchell	$2^{2^{n-1}}$	1
Yang-Guo	$2^{2^{n-1}} + 2^n - 2n + 2^{2^{n-4}} - 2^{n-3}$	1+ε
우리	$(503,483,766,022,188)^{2^{n-6}}$	1.52

5 결론

본 논문에서는 상관면역 함수의 설계 방법을 제시하였다. 제시된 방법으로 정의역의 차수가 6이하일 때는 상관면역 함수를 모두 구할 수 있다. 또 정의역의 차수가 클 경우에는 상관면역 함수 개수의 하한값을 구하였다. 이들 값은 기존의 제시된 결과보다 훨씬 우수하다.

참고 문헌

[1] P. Camion, C. Carlet, P. Charpin and N.Spndrier, "On correlation-immune functions", Advanced in Cryptology-CRYPTO'91, Springer-Verlag, pp. 87-100, 1992.

[2] C.J. Mitchell, "Enumerating boolean functions of cryptographic significance", J. Cryptology, 2, pp. 155-170, 1990.

[3] J. Seberry, X.M. Zhang, and Y. Zheng, "On constructions and nonlinearity of correlation immune functions", Advanced in Cryptology-EUROCRYPT'93, Springer-Verlag, pp. 181-199, 1994.

[4] T. Siegenthaler, "Correlation immunity of nonlinear combining functions for cryptographic applications", IEEE Trans. on Inf. Th., IT-30, pp. 776-780, 1984.

[5] Y. Xian, "Correlation-immunity of boolean functions", Electronics Letters 23, pp. 1335-1336, 1987.

[6] G. Xiao and J. Massey, "A spectral characterization of correlation-immune combining functions", IEEE Trans. on Inf. Th., IT-34, pp. 569-571, 1988.

[7] Y.X. Yang and B. Guo, "Further enumerating boolean functions of crtptographic significance", J. Cryptology, 8, pp. 115-122, 1995.

[8] 성수학, 지성택, 이상진, 김광조, "상관면역 함수와 비선형치", 한국통신정보보호학회 논문집 투고, 1996.