COVERING RADIUS OF THE $(N-3)$-RD ORDER REED-MULLER CODE IN THE SET OF RESILIENT FUNCTIONS

Yuri Borissov
Institute of Mathematics and Informatics,
Bulgarian Academy of Sciences,
8 G.Bonchev, 1113 Sofia, Bulgaria
yborisov@moi.math.bas.bg
An Braeken, Svetla Nikova
Department Electrical Engineering, ESAT/COSIC,
Katholieke Universiteit Leuven, Kasteelpark Arenberg 10,
B-3001 Heverlee-Leuven, Belgium
an.braeken,svetla.nikova@kuleuven.ac.be

INTRODUCTION

In an important class of stream ciphers, called combination generators, the key stream is produced by combining the outputs of several independent Linear Feedback Shift Register (LFSR) sequences with a nonlinear Boolean function. Siegenthaler [12] was the first to point out that the combining function should possess certain properties in order to resist divide-and-conquer attacks.A Boolean function to be used in the combination generator (or more general also in stream ciphers) should satisfy several properties. *Balancedness* – the Boolean function has to output zeros and ones with equal probabilities. *High nonlinearity* - the Boolean function has to be at sufficiently high distance from any affine function. *Correlation-immunity* (of order t) - the output of the function should be statistically independent of the combination of any t of its inputs. A balanced correlation-immune function is called *resilient*.

Besides the divide-and-conquer attacks, another important class of attacks on combination generators are the algebraic attacks [4, 5]. The central idea in the algebraic attacks is to use a lower degree approximation of the combining Boolean function and then to solve an over-defined system of nonlinear multivariate equations of low degree by efficient methods such as XL or simple linearization [3]. In order to resist these attacks, the Boolean function should have not only a a high algebraic degree but also a high distance to lower order degree functions. The trade-off between resiliency and algebraic degree is well-known. To achieve the desired trade-off designers typically fix one or two parameters and try to optimize the others.

In this paper, we investigate the generalization of the trade-off between resiliency and algebraic degree. In particular, we study the relation between resiliency and distance to lower order degree functions. In order to define a theoretic model for combining these properties, Kurosawa *et al.* [6] have introduced a new covering radius $\hat{\varrho}(t, r, n)$, which measures the maximum distance between t-resilient functions and r-th degree functions or the r-th order Reed-Muller code $RM(r, n)$. That is $\hat{\varrho}(t, r, n) = \max d(f(\overline{x}), RM(r, n))$, where the maximum is taken over the set $\mathcal{R}_{t,n}$ of t-resilient Boolean functions of n variables. Note that as the covering radius of Reed-Muller codes is defined by $\varrho(r, n) = \max d(f, RM(r, n))$ where the maximum is taken over all Boolean functions f, it holds that $0 \leq \hat{\varrho}(t, r, n) \leq \varrho(r, n)$. Kurosawa *et al.* also provide a table with certain lower and upper bounds for $\hat{\varrho}(t, r, n)$. In [1] some exact values and new bounds for the covering radius of the second order Reed-Muller codes in the set of resilient functions were found.

In this paper we find the exact value of the covering radius of $RM(n-3, n)$ in the set of 1-resilient Boolean functions of n variables, when $\lfloor n/2 \rfloor = 1 \bmod 2$. We also improve the lower bounds for covering radius of the Reed-Muller codes $RM(r, n)$ in the set of t-resilient functions, where $\lceil r/2 \rceil = 0 \bmod 2$, $t \leq n - r - 2$ and $n \geq r + 3$. We start with some background on Boolean functions.

BACKGROUND

Any Boolean function $f(\overline{x})$ on $\mathbb{F}_2^n$ can be uniquely expressed in the algebraic normal form (ANF):

$$f(\overline{x}) = \sum_{(a_1, \ldots, a_n) \in \mathbb{F}_2^n} h_f(a_1, \ldots, a_n) x_1^{a_1} \cdots x_n^{a_n},$$

with h_f a function on $\mathbb{F}_2^n$, defined by $h_f(\overline{a}) = \sum_{\overline{x} \leq \overline{a}} f(\overline{x})$ for any $\overline{a} \in \mathbb{F}_2^n$, where $\overline{x} \leq \overline{a}$ means that $x_i \leq a_i$ for all $i \in \{1, \ldots, n\}$. The algebraic degree of f, denoted by $\deg(f)$ or shortly d, is defined as the number of variables in the highest term $x_1^{a_1} \cdots x_n^{a_n}$ in the ANF of f for which $h_f(a_1, \ldots, a_n) \neq 0$. The suport of f, denoted by $sup(f)$, is the set of all vectors x for which $f(x) \neq 0$. The Walsh transform of

$f(\overline{x})$ is a real-valued function over $\mathbb{F}_2^n$ that is defined as

$$W_f(\overline{\omega}) = \sum_{\overline{x} \in \mathbb{F}_2^n} (-1)^{f(\overline{x}) + \overline{x} \cdot \overline{\omega}},$$

where $\overline{x} \cdot \overline{w}$ denotes the dot product of the vectors $\overline{x}$ and $\overline{w}$, i.e., $\overline{x} \cdot \overline{w} = x_1 w_1 + \cdots + x_n w_n$.

Definition 1 *A function $f(\overline{x})$ is called t-th order correlation-immune if its Walsh transform satisfies $W_f(\overline{\omega}) = 0$, for $1 \leq wt(\overline{\omega}) \leq t$, where $wt(\overline{x})$ denotes the Hamming weight of $\overline{x}$. Balanced t-th order correlation-immune functions are called t-resilient functions, i.e. $W_f(\overline{\omega}) = 0$, for $0 \leq wt(\overline{\omega}) \leq t$.*

By the well-known *Siegenthaler's inequality* [11] the maximal possible algebraic degree of t-resilient function f of n variables is equal to $n - t - 1$ when $t < n - 1$. The problem for constructing resilient functions (in particular such of maximal possible degree) attracted the attention of many authors in the past. Among other works we mention [11], [2] and [10]. The next theorem shows how we can easily construct $(t+1)$-resilient function on $\mathbb{F}_2^{n+1}$ from t-resilient function on $\mathbb{F}_2^n$.

Lemma 2 *[2] Let x_{n+1} be a linear variable, i.e., $f(x_1, \ldots, x_n, x_{n+1}) = g(x_1, \ldots, x_n) + x_{n+1}$, where $g(x_1, \ldots, x_n)$ is t-resilient. Then $f(x_1, \ldots, x_n, x_{n+1})$ is $(t+1)$-resilient.*

We also make use of the following theorem:

Theorem 3 *[7] The covering radius of $RM(n-3, n)$ is equal to $n+2$ if n is even. If n is odd, the covering radius is equal to $n+1$.*

To prove the theorem, McLoughlin constructed a coset for which the minimal weight is equal to $n+2$ when n is even, and $n+1$ when n is odd. This coset contains σ_{n-2}, the symmetric polynomial consisting of all terms of degree $n-2$.

THE COVERING RADIUS OF $(N-3)$-RD REED-MULLER CODES IN THE SET OF 1-RESILIENT BOOLEAN FUNCTIONS

In order to prove the main theorem of this paper we will need the following lemmas.

Lemma 4 *Let $\sigma_i(\overline{x})$ be the symmetric polynomial of n variables containing all terms of degree i ($\sigma_0(\overline{x}) = 1$) and $S(\overline{x}) = \sum_{i=0}^{n-2} \sigma_i(\overline{x})$. Then*

$$\overline{v} \in sup(S) \textit{ if and only if } wt(\overline{v}) = \begin{cases} 0, n-1, n & \textit{when } n \textit{ is even;} \\ 0, n-1 & \textit{when } n \textit{ is odd.} \end{cases}$$

Proof. Let $\overline{v} \in \mathbb{F}_2^n$ be a vector of weight w. It is easy to see that the number of terms in $\sigma_i(\overline{v})$ equal to 1 is $\binom{w}{i}$ (as usual $\binom{w}{i} = 0$, when $w < i$). Therefore the number of terms in $S(\overline{v})$ that are equal to 1 is $N(w) = \sum_{i=0}^{n-2} \binom{w}{i}$ i.e. $S(\overline{v}) = N(w) \mod 2$. There are four cases to be considered:

1. If $w = 0$, then $S(\overline{0}) = 1$;
2. If $0 < w < n-1$, then $N(w) = 2^w$ and thus $S(\overline{v}) = N(w) \mod 2 = 0$;
3. If $w = n-1$, we have $N(n-1) = \sum_{i=0}^{n-2} \binom{n-1}{i} = 2^{n-1} - 1$ and therefore $S(\overline{v}) = 1$;
4. If $w = n$, we have $N(n) = \sum_{i=0}^{n-2} \binom{n}{i} = 2^n - (n+1)$. Therefore

$$S(\overline{1}) = \begin{cases} 1 & \text{when n is even;} \\ 0 & \text{when n is odd.} \end{cases}$$

This completes the proof. □

Lemma 5 *Let $S(\overline{x})$ be the symmetric Boolean function of n variables, defined in Lemma 4, where n is equal to $4k+2$ or equal to $4k+3$. Let $\overline{v}$ be an arbitrary vector of weight $2k+1$ or of weight $2k+2$. Then the Walsh transform value $W_S(\overline{v}) = 0$.*

Proof. Let us consider the following two linear functions: $L_1(\overline{x}) = \sum_{i=1}^{2k+1} x_i$ and $L_2(\overline{x}) = \sum_{i=1}^{2k+2} x_i$. Arranging the set *sup*(S) in decreasing lexicographic order, it is easy to see that $L_j = 0, j = 1, 2$ for the half of the vectors from *sup*(S). Since the linear functions are balanced the same is true for the complement set of *sup*(S), in which S takes value 0. Therefore L_1 and L_2 differ from S in 2^{n-1} points i.e. $d(L_j, S) = 2^{n-1}, j = 1, 2$. By using the relation $W_f(\overline{\omega}) = 2^n - 2\, d(\langle\overline{\omega}, \overline{x}\rangle, f)$ we get $W_S(\overline{v}) = 0$ where $\overline{v}$ is either the vector having only ones in the first $2k+1$ or in the first $2k+2$ coordinates. Since $S(\overline{x})$ is a symmetric function this holds for any vector of weight $2k+1$ or $2k+2$. □

Let T be a subset of $\mathbb{F}_2^n$. The rank of T, denoted by $rank(T)$, is defined as the maximal number of linearly independent elements from T.

Lemma 6 *Let n be equal to $4k+2$ or equal to $4k+3$ and $Z = \{\overline{v} \in \mathbb{F}_2^n : wt(\overline{v}) = 2k+1$ or $2k+2\}$. Denote by $\overline{v}_1$ the vector $(1,1,1,..1,0,0,0,...0)$ of weight $2k+1$. Then the set $Z + \overline{v}_1$ has rank n.*

Proof. Note that the following vectors of weight 2

$$(1,0,0,...,0,1,0,...0),\ (0,1,0,...,0,1,0...0),\ \ldots,\ (0,0,0,...,1,1,0...0),$$

where the second "1" is in the $(2k+2)$-nd position, belong to $Z + \overline{v}_1$. The same is valid for the vectors having only one "1" in positions $2k+2$ till n. Obviously, these are n linearly independent vectors and the proof is complete. □

Theorem 7 *The covering radius of RM(n-3,n) in the set of 1-resilient Boolean functions of n variables is equal to:*

$$\hat{\varrho}(1, n-3, n) = \begin{cases} n+2, & \textit{when } n = \textit{4k+2;} \\ n+1, & \textit{when } n = \textit{4k+3.} \end{cases}$$

Proof. By the result of McLoughlin [7] (see Theorem 3), the Boolean function $S(\overline{x})$ defined in Lemma 4, belongs to the coset of $RM(n-3,n)$ with a maximal possible minimal weight. By Lemma 5 and Lemma 6 and using the procedure for "change the basis" described by Maitra and Pasalic [9] the function $S(\overline{x})$ is affine reducible to 1-resilient function. □

Finally, let us consider the case $n = 4$. It is easy to see that σ_2 is affine equivalent to some function in the coset of $RM(1,4)$ containing the function $f = x_1x_2 + x_3x_4$. However f is a bent function and therefore the coset $\sigma_2 + RM(1,4)$ contains no balanced functions. By Dickson [8] theorem the remaining two types of cosets (which are interesting when consider 1−resilient functions of 4 variables), are RM(1,4) itself and these equivalent to $x_1x_2 + RM(1,4)$. In fact the function $g = x_1x_2 + x_3 + x_4$ is 1-resilient and the minimal weight of its coset is 4. Hence the covering radius of interest is 4 (see also numerical results in [6]).

By induction, using Theorem 3 and Theorem 7, we can also generalize the lower bounds for $RM(r,n)$ in the set of t-resilient functions where $\lceil r/2 \rceil = 0 \bmod 2$, $t \le n-r-2$ and $n \ge r+3$.

Theorem 8 *The covering radius of the Reed-Muller code $RM(r,n)$ in the set $\mathcal{R}_{t,n}$ for $\lceil r/2 \rceil = 0 \bmod 2$, $t \le n-r-2$ and $n \ge r+3$ is bounded from below by 2^{n-3}.*

In particular, for $r = 3$ and $r = 4$, this leads to the following lower bound:

Corollary 9 *The covering radius of the Reed-Muller code $RM(3,n)$ in the set $\mathcal{R}_{t,n}$ for $t \le n-5$ is bounded from below by 2^{n-3}, when $n \ge 6$. The covering radius of the Reed-Muller code $RM(4,n)$ in the set $\mathcal{R}_{t,n}$ for $t \le n-6$ is bounded from below by 2^{n-3}, when $n \ge 7$, i.e.*

$$\begin{aligned} \hat{\varrho}(t,3,n) &\ge 2^{n-3} \quad \textit{for} \;\; t \le n-5,\; n \ge 6 \\ \hat{\varrho}(t,4,n) &\ge 2^{n-3} \quad \textit{for} \;\; t \le n-6,\; n \ge 7. \end{aligned}$$

CONCLUSION

In this paper, we continued the study of the covering radius in the set of resilient functions, which has been defined by Kurosawa *et al.* [6]. This new concept is meaningful to cryptography especially in the context of the new class of algebraic attacks on stream ciphers proposed by Courtois and Meier at Eurocrypt 2003 [4] and Courtois at Crypto 2003 [5]. In order to resist such attacks the combining Boolean function should be at high distance from lower degree functions.

Using a result from coding theory on the covering radius of $(n-3)$-rd Reed-Muller codes, we establish exact values of the the covering radius of $RM(n-3,n)$ in the set of 1-resilient Boolean functions of n variables, when $\lfloor n/2 \rfloor = 1 \bmod 2$. We also improve the lower bounds for covering radius of the Reed-Muller codes $RM(r,n)$ in the set of t-resilient functions, where $\lceil r/2 \rceil = 0 \bmod 2$, $t \le n-r-2$ and $n \ge r+3$.

In the table below we present the improved numerical values of the covering radius for resilient functions. The entry $\alpha - \beta$ means that $\alpha \le \hat{\varrho}(t,r,n) \le \beta$.

Table 1: Numerical data of the bounds on $\hat{\varrho}(t, r, n)$

	n	1	2	3	4	5	6	7
	$r=1$		0	2	4	12	26	56
	$r=2$			0	2	6	16	$40-44$
$t=0$	$r=3$				0	2	8	$20-22$
	$r=4$					0	2	8
	$r=5$						0	2
	$r=6$							0
	n	1	2	3	4	5	6	7
	$r=1$			0	4	12	24	56
	$r=2$				0	6	16	$36-44$
$t=1$	$r=3$					0	8	$20-22$
	$r=4$						0	8
	$r=5$							0
	n	1	2	3	4	5	6	7
	$r=1$				0	8	24	56
	$r=2$					0	16	$32-44$
$t=2$	$r=3$						0	$16-22$
	$r=4$							0
	n	1	2	3	4	5	6	7
	$r=1$					0	24	48
	$r=2$						0	32
$t=3$	$r=3$							0

REFERENCES

[1] Y. Borissov, A. Braeken, S. Nikova, B. Preneel, On the Covering Radius of Second Order Binary Reed-Muller Code in the Set of Resilient Boolean Functions, IMA International Conference on Cryptography and Coding, Springer-Verlag LNCS 2898, 2003, pp. 82-92.

[2] P. Camion, C. Carlet, P. Charpin, N. Sendrier, On Correlation Immune Functions,*CRYPTO'91*, LNCS 576, Springer-Verlag 1991, pp. 87-100.

[3] N. Courtois, A. Klimov, J. Patarin, A. Shamir, Efficient Algorithms for Solving Overdefined Systems of Multivariate Polynomial Equations, *Eurocrypt'00*, LNCS 1807, Springer-Verlag, 2000, pp. 392-407.

[4] N. Courtois, W. Meier, Algebraic Attacks on Stream Ciphers with Linear Feedback, *Eurocrypt'03*, LNCS 2656, Springer-Verlag 2003, pp. 345-359.

[5] N. Courtois, Fast Algebraic Attacks on Stream Ciphers with Linear Feedback *Crypto'03*, LNCS 2729, Springer-Verlag 2003, pp. 176-194.

[6] K. Kurosawa, T. Iwata, T. Yoshiwara, New Covering Radius of Reed-Muller Codes for t-Resilient Functions, *SAC'01*, LNCS 2259, Springer-Verlag 2001, pp. 75-86.

[7] A. McLoughlin, The Covering Radius of the $(m-3)-$rd Order Reed-Muller Codes and a Lower Bound on the $(m-4)-$th Order Reed-Muller Codes, *SIAM J. Appl. Mathematics*, vol. 37, No. 2, October 1979, pp. 419-422.

[8] F. J. MacWilliams, N. J. A. Sloane, The Theory of Error-Correcting Codes, North-Holland Publishing Company 1977.

[9] S. Maitra, E. Pasalic, Further Constructions of Resilient Boolean Functions with Very High Nonlinearity, *IEEE Transactions on Information Theory*, vol. 48, No.7, July 2002, pp. 1825-1834.

[10] J. Seberry, J. Zhang, Y. Zheng, On Constructions and Nonlinearity of Correlation Immune Functions, *Eurocrypt'93*, LNCS 765, Springer-Verlag 1994, pp. 181-199.

[11] T. Siegenthaler, Correlation-Immunity of Non-linear Combining Functions for Cryptographic Applications, *IEEE IT*, vol. 30, No. 5, 1984, pp. 776-780.

[12] T. Siegenthaler, Decrypting a Class of Stream Ciphers Using Ciphertext Only, *IEEE Trans. Comp.*, vol 34, No. 1, 1985, pp. 81-85.

[13] Y. Tarannikov, On Resilient Functions with Maximun Possible Nonlinearity, *Indocrypt 2000*, LNCS 1977, pp. 19-30.