

ALMOST BOOLEAN FUNCTIONS: THE DESIGN OF BOOLEAN FUNCTIONS BY SPECTRAL INVERSION

JOHN A. CLARK AND JEREMY L. JACOB

Department of Computer Science, University of York, York, United Kingdom

SUBHAMOY MAITRA

Indian Statistical Institute, Calcutta, India

PANTELIMON STĂNICĂ

Mathematics Department, Auburn University Montgomery, Montgomery, Alabama, USA

The design of Boolean functions with properties of cryptographic significance is a hard task. In this paper, we adopt an unorthodox approach to the design of such functions. Our search space is the set of functions that possess the required properties. It is “Boolean-ness” that is evolved.

Key words: boolean functions, rotational symmetry, nonlinearity, autocorrelation, correlation immunity, simulated annealing.

1. INTRODUCTION

Boolean functions form crucial components in cryptographic systems. The design of suitable functions has received significant attention from cryptographers for decades. Recently, metaheuristic search (particularly hill-climbing, genetic algorithms, and simulated annealing) has emerged as a potentially very powerful tool for the design of such functions (Millan, Clark, and Dawson 1997, 1998, 1999; Clark and Jacob 2000). Most recently, metaheuristic search in combination with well-established theory has found functions with properties unattained by any other means (Clark et al. 2002).

All the optimization work so far has searched the space of Boolean functions for those with particular properties. In this paper, we invert this notion and search the space of artifacts with the required properties and seek one that is actually a Boolean function. Combining this general approach with very recent theory allows functions with hitherto unobtained properties to be found.

In this paper, our primary aim is to report new research results that compete with or improve on the best results of available theoretical constructions. A full motivation or explanation of why what we wish to do is of interest to the cryptographic community is beyond the scope of this paper. However, we have included sufficient background material to make the paper accessible to the non-cryptographer. We hope also that this material may encourage researchers whose primary expertise lies in metaheuristic search to work in the area.

2. PRELIMINARIES

2.1. Motivation

Boolean functions play a critical role in cryptography. A standard block cipher such as the data encryption standard (DES) or its more recent replacement, the advanced encryption

Pantelimon Stănică is also associated to the Institute of Mathematics “Simion Stoilow” of the Romanian Academy, Bucharest, Romania.

Address correspondence to John A. Clark, The Department of Computer Science at The University of York, York, YO10 5DD, UK; e-mail: jac@cs.york.ac.uk.

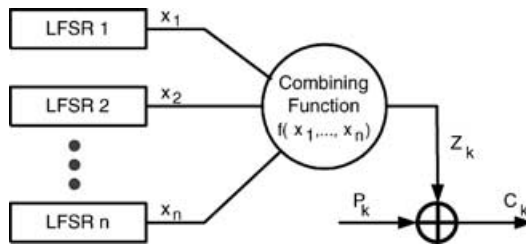


FIGURE 1. Combining bit streams.

standard (AES), can be thought of as functions that take vectors of Boolean inputs (e.g., 64- or 80-bit plaintext data blocks) and output similar vectors of Booleans (i.e., the ciphertext blocks). The exact functional transformation is generally defined by a secret key K . Block ciphers typically comprise a number of iterations of a much simple cipher, usually referred to as a “round”, the output from one round instance forming the input to the next. DES, for example, comprises 16 rounds. Within the round cipher, smaller functions may be deployed. The “substitution boxes” of DES take 6-bit input vectors and give 4-bit output vectors. Much of the security of the algorithm rests in the design of these boxes. Vector-output functions can be thought of as combinations of single-output functions, but single-output functions have serious applications in their own right. In this paper, we restrict our attention to single-output functions. Hereafter, the term “Boolean functions” will refer to such functions.

In some stream ciphers Boolean functions may be used to combine the outputs of several bit streams. Figure 1 illustrates a classic stream cipher model. The plaintext stream $\{P_k\}$ of bits is XOR-ed with a pseudo-random bit stream $\{Z_k\}$ to give a ciphertext stream $\{C_k\}$. The plaintext is recovered by the receiver by XOR-ing the cipherstream with the same pseudo-random stream. The pseudo-random stream is formed from several bit streams generated by linear feedback shift registers (LFSRs) using a suitable combining function f . At each step the rightmost bit of register LFSR i becomes the i th input x_i to the combining function f , the register is shifted right by 1 and the new value for the leftmost bit is calculated by XOR-ing a number of the other bits of the register (before shifting). The initial state of the registers forms the *secret key*. The function f must be sufficiently complex that cryptanalysts will not be able to determine the initial state of the registers, even when they know what the plaintext is (and thus can recover the pseudo-random stream $\{Z_k\}$).

The possession (or otherwise) of various structural properties by f may facilitate cryptanalysis. Suppose our stream cipher had two 64-bit registers and a combining function f . An enumerative attack on the cipher would involve trying 2^{128} possible keys. However, suppose that $f(x_1, x_2) = x_1 \oplus x_2$. If we know the output from f we know that there are only two possible combinations of x_1, x_2 that would give rise to that output. And thus if we observe the first 64 Z_k outputs we have effectively reduced the size of the key space to be searched to 2^{64} . This illustrates the dangers of *linearity* (formally defined below). A dreadful choice of f would be $f(x_1, x_2) = x_1$; observing the first 64 output bits would reveal all the relevant secret key bits (i.e., for LFSR 1). Once the initial state of any register is known, its subsequent operation is completely determined. In practice, more sophisticated attacks exploiting linearity of f are possible (e.g., the Best Affine Approximation, see Ding, Xiao and Shan (1991), and techniques derived from Matsui’s linear cryptanalysis (Matsui 1993).

It would seem incompetent to use a truly linear function, and in practice the best a cryptanalyst can hope for is a small statistical bias, for example, $f(x_1, x_2) = x_1 \oplus x_2$ might be true (or false) for marginally greater than 50% of input combinations.

Exploiting linear approximations involving only one or two bits is generally easier than exploiting linear approximations of several bits. In a larger stream cipher, with more input

register streams, we might require that there is no effective linear bias involving m or fewer input registers. This leads to the notion of *correlation immunity of order m* (see below).

Autocorrelation is another structural property of f that might cause problems. Essentially, autocorrelation is a measure of a form of periodicity in a function. If autocorrelation is too high, then new forms of attack become possible, such as differential cryptanalysis. The details need not concern us here.

Highly nonlinear Boolean functions on small (up to 11) number of variables can be used efficiently in symmetric cipher systems such as the crossing over system (COS), cipher family) invented by Filliol, Fontaine and Vianne (2001). The acronym COS refers to how each cipher internally works. COS ciphers are block ciphers, built only from stream cipher primitives (nonlinear feedback shift registers and Boolean functions). COS ciphers use Boolean functions of 11 variables for key setup and of 9 variables for the encryption procedure itself. These functions must achieve best possible tradeoffs between important cryptographic properties (balancedness, correlation immunity, high nonlinearity, etc).

The above text is for motivation only. There is no such thing as a “secure Boolean function”; there are only functions with properties that prove useful in particular system contexts. We wish only to convey the idea that Boolean functions are an important research topic with practical significance to modern day cryptography. We are not addressing “toy” problems.

2.2. Single-Output Boolean Functions

We denote the binary truth table of a Boolean function by $f : Z_2^n \rightarrow Z_2$ mapping each combination of n binary values to some binary value. If the number of combinations mapping to 0 is the same as the number mapping to 1 then the function is said to be *balanced*.

The *polarity truth table* is a particularly useful representation for our purposes. It is defined by $\hat{f}(x) = (-1)^{f(x)}$. Two functions f and g are said to be uncorrelated when $\sum_{x \in Z_2^n} \hat{f}(x)\hat{g}(x) = 0$. If so, if you try to approximate f by using g , you will be right half the time and wrong half the time.

For convenience, we often refer to the elements of Z_2^n by their natural decimal interpretation. Thus 11111111 is interpreted as 255, 00000011 as 3, etc. We now give formal definitions of terms used in this paper. Summations will generally be over $0, \dots, (2^n - 1)$ unless otherwise stated.

Linear Boolean Function. A linear Boolean function, selected by $\omega \in Z_2^n$, is denoted by

$$L_\omega(x) = \omega_1 x_1 \oplus \omega_2 x_2, \dots, \oplus \omega_n x_n, \quad (1)$$

where $w_i x_i$ denotes the bitwise AND of the i th bits of ω and x , and \oplus denotes bitwise XOR.

Affine Function. The set of affine functions is the set of linear functions and their complements

$$A_{\omega,c}(x) = L_\omega(x) \oplus c, \quad (2)$$

where $c \in Z_2$.

Algebraic Normal Form and Degree. A Boolean function can be expressed as a minimal (XOR) sum of (AND) products:

$$f(x_1, \dots, x_n) = a_0 \oplus a_1 x_1 \dots \oplus a_n x_n \oplus a_{1,2} x_1 x_2 \dots \oplus a_{n-1, n} x_{n-1} x_n \dots \oplus a_{1,2,\dots,n} x_1 x_2 \dots x_n. \quad (3)$$

This is the *algebraic normal form* of the function. The *algebraic degree* of a function f is the largest number of inputs appearing in any product in its algebraic normal form. Thus $x_1 \oplus x_2$ has degree 1 (i.e., is linear), $x_1 \oplus x_1x_2x_3$ has degree 3, etc.

Walsh–Hadamard Transform. For a Boolean function f the Walsh–Hadamard transform W_f is defined by

$$W_f(\omega) = \sum_x \hat{f}(x) \hat{L}_\omega(x). \quad (4)$$

Thus, each Walsh–Hadamard transform $W_f(\omega)$ is the vector dot product of the polar forms of f and the linear function L_ω . We denote the maximum absolute value taken by the transform by $WH_{\max}(f) = \max_{\omega \in Z_2^n} |W_f(\omega)|$. It is related to the nonlinearity of f . We refer to the vector $(W_f(0), \dots, W_f(2^n - 1))$ as the *Walsh Spectrum* of the function f .

Nonlinearity. The nonlinearity N_f of a Boolean function f is its minimum distance to any affine function. It is given by

$$N_f = \frac{1}{2}(2^n - WH_{\max}(f)). \quad (5)$$

Correlation Immunity. A function f is correlation immune of order m if and only if

$$|W_f(\omega)| = 0, \quad 1 \leq |\omega| \leq m, \quad (6)$$

where $|\omega|$ denotes the Hamming weight of ω .

Resilience. A function f that is correlation immune of order m is *resilient* if and only if it is also balanced:

$$|W_f(\omega)| = 0, \quad 0 \leq |\omega| \leq m. \quad (7)$$

Parseval’s Theorem. This states that

$$\sum_{\omega} (W_f(\omega))^2 = 2^{2n}. \quad (8)$$

In a sense this can be thought of as a form of generalized Pythagoras’ theorem (see below).

Bent Functions. A function on an even number n of inputs is a bent function if and only if

$$|W_f(\omega)| = 2^{\frac{n}{2}}, \quad \forall \omega \in 0, \dots, (2^n - 1) \quad (9)$$

Bent functions achieve the maximum possible nonlinearity but are unbalanced (because $W_f(0) \neq 0$). They are often used as components in theoretical constructions of larger Boolean functions. The maximum algebraic degree for a bent function on n inputs is $\frac{n}{2}$.

Autocorrelation Transform. The autocorrelation transform of a Boolean function f is given by:

$$\hat{r}_f(s) = \sum_x \hat{f}(x) \hat{f}(x \oplus s). \quad (10)$$

We refer to the vector $(\hat{r}_f(0), \dots, \hat{r}_f(2^n - 1))$ as the *autocorrelation spectrum* of the function f . $\hat{r}_f(0)$ always has value 2^n . We denote the maximum absolute value

in the rest of the autocorrelation spectrum of a function f by Δ_f , i.e., $\Delta_f = \max_{s \neq 0} |\sum_x \hat{f}(x)\hat{f}(x \oplus s)|$. Here x and s range over Z_2^n . The Δ_f was one of the “global avalanche characteristics” suggested by Zhang and Zheng (1995) and is often referred to simply as the *autocorrelation* of function f .

Boolean functions in polar form are represented here as vectors in \mathbf{R}^{2^n} where each component is either +1 or -1. Thus, Boolean functions are particular points on the surface of a hypersphere of radius $2^{\frac{n}{2}}$. Not all points on this surface are Boolean functions; there are uncountably more points with some elements that are not +1 or -1.

The \hat{L}_ω , where ω ranges over $0, \dots, (2^n - 1)$, form a set of orthogonal Boolean functions that span \mathbf{R}^{2^n} (i.e., they form a basis set). A function vector \hat{f} can be expressed as:

$$\hat{f} = \sum_{\omega} \frac{1}{2^n} (\hat{f} \cdot \hat{L}_\omega) \hat{L}_\omega = \sum_{\omega} \frac{W_f(\omega)}{2^{\frac{n}{2}}} \frac{\hat{L}_\omega}{2^{\frac{n}{2}}}. \quad (11)$$

The $\frac{W_f(\omega)}{2^{\frac{n}{2}}}$ values are essentially the coordinates of \hat{f} with respect to the *orthonormal* basis provided by the various $\frac{\hat{L}_\omega}{2^{\frac{n}{2}}}$. The sum of the squares of these coordinates gives the square of the magnitude of \hat{f} (i.e., 2^n). This leads to Parseval’s theorem of equation (8). With the function viewed as the hypotenuse, we can now understand why this theorem can be thought of as an instance of Pythagoras’ Theorem in 2^n dimensions. The magnitude of $\frac{1}{2^n} W_f(\omega)$ indicates the extent to which \hat{f} is approximated by the linear function \hat{L}_ω . The vector of Walsh–Hadamard values (the Walsh–Hadamard *spectrum*) provides an alternative representation for \hat{f} .

A function is *balanced* if $W_f(0) = 0$ indicating that \hat{f} has equal numbers of +1 and -1 elements. A function f is correlation immune of order m (*CI(m)* for short) if all nonempty subsets of inputs of size m are statistically independent of the output of that function. It is simpler, however, to work with Xiao and Massey’s characterization in terms of the Walsh–Hadamard values (Xiao and Massey 1988). This is the form given in equation (6). It simply states that no linear function of m or fewer variables gives any degree of approximation to f . In general, low-order linear approximations are easier to exploit than higher order ones. Similarly, low algebraic degree may significantly facilitate cryptanalysis.

2.3. Conflicts, Tradeoffs and Profiles

We might typically wish to obtain balanced functions with high nonlinearity, high algebraic degree, high order of correlation immunity, and low autocorrelation. However, it is apparent that some of these criteria are in conflict. Consider the bent functions defined in equation (9). These achieve the maximum possible nonlinearity (they minimize the maximum magnitude of Walsh values) but are unbalanced. Now if we require a function to be balanced ($W_f(0) = 0$), then some other $W_f(\omega)$ must have magnitude greater than $2^{\frac{n}{2}}$ to compensate (by Parseval’s theorem given in equation (8)). Similarly, increasing order of immunity can never lead to an increase in achievable nonlinearity.

This conflict means that tradeoffs have to be made. Considerable research has been carried out to derive bounds on achievable profiles (combinations) of properties and to demonstrate functions achieving those bounds (Sarkar and Maitra 2000a,b; Tarannikov 2000; Zheng and Zhang 2000; Pasalic et al. 2001; Carlet and Sarkar 2002; Maitra and Pasalic 2002; Maitra and Sarkar 2002; Maity and Johansson 2002). We denote the profile of a function by a quadruplet: (n, m, d, nl) is the profile of a function on n inputs that is balanced, correlation immune of order m , and which has algebraic degree d and nonlinearity nl . Where we do not impose the requirement for balancedness, we will use profiles with square brackets $[n, m, d, nl]$.

It has been known for a long time that $m + d \leq n$ and that $m + d \leq (n - 1)$ for balanced functions. Much of the above research has concerned with deriving bounds for the nonlinearity components. A summary of current profile bounds can be found in Clark et al. (2002) (which also records the most extreme profiles demonstrated by theoretical or optimization techniques so far).

2.4. Observations on Spectral Properties

We can see that various important cryptographic criteria (balance, correlation immunity, resilience, nonlinearity, etc.) are defined in terms of the Walsh–Hadamard values of that function. Given a spectrum of Walsh–Hadamard values we can see easily whether the various criteria are met. This motivates the new approach described below. A Walsh spectrum satisfying the criteria forms the starting point for the search and we move amongst spectra that preserve the desired properties, in search of a spectrum that is also that of a Boolean function.

3. THE GUIDED SEARCH BASED ON SIMULATED ANNEALING

3.1. Casting the Problem as One of Guided Search

Let us consider that the (currently unknown) Walsh spectrum W_f of a balanced Boolean function f on n input variables with desired nonlinearity nl and order of correlation immunity m , is given by $\{W_f(0), \dots, W_f(2^n - 1)\}$. Let $P = \{P(0), \dots, P(2^n - 1)\}$ be a permutation of the values of W_f . Now consider the set of all such permutations P with $P(\omega) = 0$ for all $0 \leq |\omega| \leq m$. This is the set of permutations that maintain the properties required for balancedness, nonlinearity and correlation immunity in the permuted Walsh spectra. However, it is not guaranteed that a permuted spectrum will be the Walsh spectrum for some Boolean function. We can obtain a function $\hat{p}(x)$ by applying the inverse Walsh transform to the spectrum P as shown below

$$\hat{p}(x) = 2^{-n} \sum_{\omega} P(\omega)(-1)^{\omega \cdot x}. \quad (12)$$

While a few permutations P , after inverse Walsh transform, will correspond to polar forms of Boolean functions, most will not. Thus, the values of the various $\hat{p}(x)$ obtained by inversion may not actually be $+1$ or -1 . (All such $\hat{p}(x)$ are still points on the surface of the 2^n -dimensional hypersphere of radius $2^{\frac{n}{2}}$.) However, the values of $\hat{p}(x)$ can be associated heuristically with a Boolean function $\hat{b}(x)$ (in polar form), by choosing,

$$\begin{aligned} \hat{b}(x) &= +1 && \text{if } \hat{p}(x) > 0, \\ \hat{b}(x) &= -1 && \text{if } \hat{p}(x) < 0, \\ \hat{b}(x) &= +1 \text{ or } -1 && \text{(chosen randomly) if } \hat{p}(x) = 0. \end{aligned}$$

Thus, any inverted spectrum can be “collapsed” to a “nearest” Boolean function. This makes heuristic search over the space of Walsh permutations easier. We need only guide the search to a permutation that *collapses* to a desired function. Many permuted spectra may collapse to the same desired function. With each permutation P we can associate a cost that indicates

how far P is from the spectrum of a valid Boolean function. Our initial cost function is given below

$$\text{cost}(P) = \sum_x (\hat{p}(x) - \hat{b}(x))^2, \quad (13)$$

where \hat{p} is the function arising from P under the inverse Walsh transform. If a permutation P gives rise to zero cost, then it is the spectrum of a Boolean function satisfying the same desired criteria of f . In practice, if the spectrum P is sufficiently “close” to that of a valid Boolean function, it will “collapse” to that function. Thus, the sought spectra are those corresponding to zero or very low values of the cost function in equation (13). The problem is now one of cost minimization over the space of permutations P indicated above.

3.2. Getting Started and Traversing the Space

A standard form of simulated annealing has been used as the search heuristic. A description is given in the Appendix A. For present purposes the details are irrelevant. We note only that it is a local search technique with an excellent ability to escape from local optima.

Our local search moves by swapping two elements $P(\alpha)$ and $P(\beta)$ of the current spectrum with $|\alpha| > m$, $|\beta| > m$ (i.e., the indices α and β must have Hamming weight greater than the required order of correlation immunity m) and $P(\alpha) \neq P(\beta)$. If the initial spectrum has the required properties, this move strategy preserves them. We now need to find an appropriate initial spectrum for the search.

The initial spectrum will be a permutation of the spectrum of a desired Boolean function. In general, it may be rather hard to generate such spectra, but in some cases it is relatively easier. Consider the spectrum of a function with profile (7, 2, 4, 56). The maximum absolute value of elements in the Walsh spectrum is 16. Theory has shown that if $n \geq 3$ and $m \leq n - 3$ then the Walsh values of an m th order resiliant function f on n variables must satisfy $|W_f(\omega)| \equiv 0 \pmod{2^{m+2}}$ (Sarkar and Maitra 2000b). Thus, the Walsh values for (7, 2, 4, 56) must be 0, 16, or -16 (an absolute Walsh value of 32 or above would give rise to a nonlinearity of 48 or less). The formula

$$2^n \times \hat{f}(0) = \sum_{\omega} W_f(\omega) \quad (14)$$

defines the value of $\hat{f}(0)$. Arbitrarily fixing this to be 1 and using Parseval’s equation allows us to determine that the Walsh spectrum must contain 36 ‘+16’s, 28 ‘-16’s, and 64 ‘0’s. Zeroes are placed in the 29 positions in the starting spectrum corresponding to $0 \leq |\omega| \leq 2$ (these elements remain fixed throughout the search) and the remaining Walsh values (i.e., 36 ‘+16’s, 28 ‘-16’s, and 35 ‘0’s) are arbitrarily allocated to the remaining positions. Using the cost function given in equation (13), 500 runs were carried out resulting in five successes. Although this is not efficient, it did generate some example desired functions.¹ A function with this profile was first demonstrated in 2001 (Pasalic et al. 2001) and then more recently in Stănică and Maitra (2002).

¹Also, we were able to apply a change of basis to transform some of the “failed” solutions into successful examples. However, finding a suitable change of basis is a highly nonlinear problem, and one which we addressed via annealing in Clark et al. (2002).

3.3. Refining the Approach

Attempts to generate bent functions on eight variables (with 136 ‘16’s and 120 ‘-16’s) proved very inefficient and thus a second cost function was developed using Titsworth’s theorem (Ding et al. 1991). This states that $W_f(\omega)$ is the Walsh spectrum of a binary Boolean function if and only if

$$\sum_{\omega} W_f(\omega) W_f(s \oplus \omega) = 2^{2^n} \delta(s), \quad (15)$$

where $\delta(s) = 1$ if $s = 0$ and $\delta(s) = 0$ otherwise. This immediately suggests a cost function that punishes deviation from this:

$$\text{cost}(W_f) = \sum_s \left(\left| \sum_{\omega} W_f(\omega) W_f(s \oplus \omega) \right| \right)^R. \quad (16)$$

When $s = 0$ the inner sum should be nonzero (and is constant for all permutations of the spectrum W_f). For $s > 0$ the inner sum should be zero. This cost function punishes deviation from zero for these sums. Two sets of experiments have been performed with $R = 2$ aimed at evolving bent functions on six and eight variables. Obtaining spectra for bent functions is easy because it must be the case that $|W_f(\omega)| = 2^{\frac{n}{2}}$ for all ω . Equation (14) was used again to determine the numbers of positive and negative Walsh values.

Fifty runs were carried out in each case. For six input variables the technique generated a bent function with (maximal) nonlinearity 28 and (maximal) algebraic degree 3 every time. The average time per run was 20.6 s. For eight input variables the average time per run was 4 m 58 s. Out of 50 experiments, we got bent functions (nonlinearity 120) 14 times, and all of them were of degree 4, which is the maximum possible. Note that for all the experiments we used a 1GHz Pentium processor with 512 MB RAM. Subsequent attempts to evolve bent functions on 10 inputs failed, despite significantly increasing computational resources for the search. Attempts to derive (9, 3, 5, 240) functions failed similarly. The experiments were performed with $R = 2$, a cooling parameter $\alpha = 0.95$, the number of moves within an inner loop MIL = 400, the maximum number of inner loops MaxIL = 800, and the number of unproductive loops being 50. (See Appendix A for details of the annealing algorithm.)

Subsequent attempts to evolve bent functions on 10 inputs failed, despite significantly increasing computational resources for the search: $\alpha = 0.99$, MIL = 1,000, MaxIL = 3,000. Attempts to derive (9, 3, 5, 240) functions failed similarly.

3.4. ROTS Enabled Guided Search

With the technique described above we could achieve some important resilient functions on seven variables and bent functions on eight variables. However, due to the superexponential search space of Boolean functions, the method showed its limitations from nine variables onward. Restricting the search space seems the order of the day. One class of function that has attracted attention recently is that of rotation symmetric boolean functions (RSBFs) (Pieprzyk and Qu 1999; Cusick and Stănică 2002; Stănică and Maitra 2002). Here, all indices which are rotationally equivalent have the same value for f . Thus, for a five variable Boolean function f , we would have $f(10001) = f(11000) = f(01100) = f(00110) = f(00011)$. We used the same strategy, but only in the space of RSBFs introduced earlier, not the total space of Boolean functions and the result of such an effort produced results which had not previously been demonstrated by any technique.

Rotational equivalence partitions the elements $0, \dots, (2^n - 1)$. For RSBFs rotationally equivalent elements have the same Walsh value. Rather than swapping dissimilar pairs of

Walsh values in the full spectrum, we now swap dissimilar Walsh values of two equivalence classes. We used the cost function of equation (16) in exactly the same way as before. This enabled us to evolve bent RSBFs on 10 variables (for truth table see Appendix B).

For (9, 3, 5, 240) attempts we can use the same divisibility theory (Sarkar and Maitra 2000b) as for the (7, 2, 4, 56) attempts to deduce that the Walsh values are 0, 32 or -32 . Furthermore, assuming $f(\vec{0}) = 0$, there are 256 '0's, 136 '32's and 120 ' -32 's. A little elementary mathematics allows us to further fix the values of specific classes (see Appendix B for details). The runs produced no (9, 3, 5, 240) ROTS functions but (unbalanced) third order correlation immune RSBFs of algebraic degree 5 (i.e., [9, 3, 5, 240]) could be produced reliably (see Appendix B for details). Similarly, additionally allowing classes to take on values of 16 and -16 allowed us to evolve (9, 2, 6, 240) functions. Both [9, 3, 5, 250] and (9, 2, 6, 240) have never been demonstrated previously. For truth tables of these functions, refer to Appendix B. (10, 3, 5, 480) functions in the RSBF class were also found. This gives confidence that with more computational resource and time, it may be possible to get cryptographically significant functions on higher number of variables.²

4. CONCLUSIONS AND FUTURE WORK

The design of Boolean functions with cryptographic significance by search or by theoretical construction has generally considered only the space of Boolean functions. Here, we have observed that many properties of interest are defined in terms of the Walsh spectrum. Using very recent cryptographic theory, we have been able to work directly in the space of Walsh spectra, starting with the properties we actually require and evolving to a solution that is a Boolean function. Embracing the concept of "almost Boolean functions", the cost function measuring just "how Boolean" a functions is, has allowed us to attain functions with optimal profiles of properties. In some cases, such profiles have never been demonstrated by theoretical construction.

In Clark et al. (2002) it was shown how metaheuristic search could equal the best achievements of all theoretical constructions for eight or fewer inputs (and in many cases demonstrate more extreme profiles). In this paper, we have shown how a novel technique can extend the achievements (particularly to higher numbers of inputs). We have used a "vanilla" variant of simulated annealing. More sophisticated search techniques should bring even better results. (Permutation genetic algorithms are currently under investigation.)

The functions evolved in this work are small (modern day cryptographic applications of Boolean functions tends to favor higher numbers of input variables). For functions of nine or fewer variables, annealing-based approaches have shown themselves more than equal to theoretical approaches. For higher numbers of inputs they have yet to make a real impact. This shows very clearly a direction for research—to significantly extend the size of functions that can be evolved.

Cryptanalysis aims to exploit structure in cryptographic system components. Restricting the search space to a structured subset (functions with rotation symmetry) may provide functions that facilitate cryptanalysis (including forms of analysis not currently known). Thus, although we have demonstrated functions with hitherto unattained profiles of properties, cryptodesigners may be wary of using such functions in particular contexts. However, once we have a rotation symmetric function with the required properties, it may be possible to

²It is fairly trivial to use the technique for seven and eight input functions. For example, 50 attempts to derive bent functions on eight variables met rapidly with 100% success (all with maximal degree 4).

transform it to one without rotation symmetric structure. Indeed, current work suggests that if the magnitude of each Walsh spectrum value is known, annealing can be used to evolve the signs of the magnitude (using the same cost functions as used in this paper) to give the Walsh spectrum of a legitimate Boolean function. The absolute values of the spectrum can be obtained from the evolved rotation symmetric boolean functions. This seems a very exciting avenue to pursue.

Our results are made possible only by embracing leading edge theory. We have used results on Walsh divisibility developed as recently as 2000 (Sarkar and Maitra 2000a) and improved matters by restricting ourselves to a class of functions investigated as recently as 2002 (Pieprzyk and Qu 1999; Cusick and Stănică 2002; Stănică and Maitra 2002; Stănică, Maitra, and Clark, 2004). The message is clear: A fusing of evolutionary approaches and leading-edge cryptographic theory has the potential to be a very significant design tool for functions of cryptographic significance. The two communities should talk more often! We recommend the area to specialists in evolutionary computing.³

REFERENCES

- CARLET, C., and P. SARKAR. 2002. Spectral domain analysis of correlation immune and resilient Boolean functions. *Finite Fields and Its Applications*, **8**(1):120–130.
- CLARK, J. A., and J. L. JACOB. 2000. Two-stage optimization in the design of Boolean functions. *In ACISP 2000, Lecture Notes in Computer Science*, Vol. 1841, Springer-Verlag, Berlin, pp. 242–254.
- CLARK, J. A., J. L. JACOB, S. STEPNEY, S. MAITRA, and W. MILLAN. 2002. Evolving Boolean functions satisfying multiple criteria. *In INDOCRYPT 2002, Lecture Notes in Computer Science*, Vol. 2551, Springer Verlag, Berlin, pp. 246–259.
- CUSICK, T. W., and P. STĂNICĂ. 2002. Fast evaluation, weights and nonlinearity of rotation-symmetric functions. *Discrete Mathematics*, **258**(1–3):289–301.
- DING, C., G. XIAO and W. SHAN. 1991. The stability theory of stream ciphers. *In Lecture Notes in Computer Science*, Vol. 561, Springer-Verlag, Berlin.
- FILLIOL, E., C. FONTAINE, and D. VIANNE. 2001. A new fast block cipher design: COS ciphers. *In Proceedings of the International Symposium on Information Theory 2001*.
- KIRKPATRICK, S., JR., C. D. GELATT, and M. P. VECCHI. 1983. Optimization by simulated annealing. *Science*, **220**(4598):671–680.
- MAITRA, S., and E. PASALIC. 2002. Further constructions of resilient Boolean functions with very high nonlinearity. *IEEE Transactions on Information Theory*, **48**(7):1825–1834.
- MAITRA, S., and P. SARKAR. 2002. Cryptographically significant Boolean functions with five-valued walsh spectra. *Theoretical Computer Science*, **276**(1 and 2):133–146.
- MAITY, S., and T. JOHANSSON. 2002. Construction of cryptographically important Boolean functions. *In INDOCRYPT 2002, Lecture Notes in Computer Science*, Vol. 2551, Springer-Verlag, Berlin, pp. 234–245.
- MATSUI, M. 1993. Linear cryptanalysis method for DES cipher. *In Advances in Cryptology—EuroCrypt’93, Lecture Notes in Computer Science*, Vol. 765, Springer-Verlag, Berlin, pp. 386–397.
- MILLAN, W., A. CLARK, and E. DAWSON. 1997. An effective genetic algorithm for finding highly nonlinear Boolean functions. *In First International Conference on Information and Communications Security, Lecture Notes in Computer Science*, Vol. 1334, Springer-Verlag, Berlin, pp. 149–158.
- MILLAN, W., A. CLARK, and E. DAWSON. 1998. Heuristic design of cryptographically strong balanced Boolean functions. *In Advances in Cryptology EUROCRYPT’98, LNCS*, Vol. 1403, Springer-Verlag, Berlin, pp. 489–499.

³A special page containing many online papers concerning optimization and cryptology can be found at <http://www.cs.york.ac.uk/security/Security/LibraryPages/NatureInspired.html>

- MILLAN, W., A. CLARK, and E. DAWSON. 1999. Boolean function design using hill climbing methods. *In* Fourth Australasian Conference on Information, Security and Privacy, April 1999. Lecture Notes in Computer Science, Vol. 1587, Springer-Verlag, Berlin, pp. 1–11.
- PASALIC, E., S. MAITRA, T. JOHANSSON, and P. SARKAR. 2001. New constructions of resilient and correlation immune Boolean functions achieving upper bound on nonlinearity. *In* Workshop on Coding and Cryptography—WCC 2001, Paris, January 8–12, 2001. Electronic Notes in Discrete Mathematics, Vol. 6, Elsevier Science, Amsterdam.
- PIEPZYK, J., and C. X. QU. 1999. Fast hashing and rotation-symmetric functions. *Journal of Universal Computer Science*, 5(1):20–31.
- SARKAR, P., and S. MAITRA. 2000a. Construction of nonlinear Boolean functions with important cryptographic properties. *In* Advances in Cryptology—EUROCRYPT 2000, May 2002. Lecture Notes in Computer Science, Vol. 1807, Springer-Verlag, Berlin, pp. 485–506.
- SARKAR, P., and S. MAITRA. 2000b. Nonlinearity bounds and constuction of resilient Boolean functions. *In* Advances in Cryptology—Crypto 2000, Berlin. Lecture Notes in Computer Science, Vol. 1880, *Edited by* M. Bellare. Springer-Verlag, Berlin, pp. 515–532.
- STĂNICĂ, P., and S. MAITRA. 2002. Rotation symmetric Boolean functions—Count and cryptographic properties. *In* R. C. Bose Centenary Symposium on Discrete Mathematics and Applications, December 2002, Indian Statistical Institute, Calcutta and Technical Report of Cryptology Research Group, Indian Statistical Institute, Technical Report Number CRG/2002/10, November 18, 2002. Also available at http://www.isical.ac.in/~crg/tech_reports.html.
- STĂNICĂ, P., S. MAITRA, and J. A. CLARK. 2004. Results on rotation symmetric bent and correlation immune Boolean functions. *In* Fast Software Encryption 2004, *Edited by* B. Roy and S. Maitra. February 2004, Delhi.
- TARANNIKOV, Y. V. 2000. On resilient Boolean functions with maximum possible nonlinearity. *In* Progress in Cryptology—INDOCRYPT 2000, Lecture Notes in Computer Science, Vol. 1977, Springer-Verlag, Berlin, pp. 19–30.
- ZHANG, X.-M., and Y. ZHENG. 1995. GAC—The criterion for global avalanche characteristics of cryptographic functions. *Journal of Universal Computer Science*, 1(5):316–333.
- ZHENG, Y., and X. M. ZHANG. 2000. Improved upper bound on the nonlinearity of high order correlation immune functions. *In* Selected Areas in Cryptography—SAC 2000, Lecture Notes in Computer Science, Vol. 1012, Springer-Verlag, Berlin, pp. 264–274.
- XIAO, G. Z., and J. L. MASSEY. 1988. A spectral characterization of correlation immune combining functions. *IEEE Transactions on Information Theory*, 34(3):569–571.

APPENDIX A: DESCRIPTION OF SIMULATED ANNEALING

Kirkpatrick, Gelatt and Vecchi (1983) proposed *simulated annealing*, a new search technique inspired by the cooling processes of molten metals. It merges hillclimbing with the probabilistic acceptance of nonimproving moves. The basic algorithm is shown in Figure 2. The search starts at some initial state $S := S_0$. There is a control parameter T known as the temperature. This starts “high” at T_0 and is gradually lowered. At each temperature, a number moves in inner loop (MIL) of moves to new states are attempted. A candidate state Y is randomly selected from the neighborhood $N(S)$ of the current state. The change in value, δ , of f is calculated. If it improves the value of $f(S)$ (i.e., if $\delta < 0$ for a minimization problem) then a move to that state is taken ($S = Y$); if not, then it is taken with some probability. The worse a move is, the less likely it is to be accepted. The lower the temperature T , the less likely is a worsening move to be accepted. Probabilistic acceptance is determined by generating a random value U in the range $(0, \dots, 1)$ and performing the indicated comparison. Initially the temperature is high and virtually any move is accepted. As the temperature is lowered it becomes ever more difficult to accept worsening moves. Eventually, only improving moves

```

S := S0
T := T0
repeat
{
  for(int i = 0; i < MIL; i++)
  {
    Select Y ∈ N(S)
    δ = f(Y) - f(S)
    if (δ < 0) then
      S = Y
    else
      Generate U := U(0, 1)
      if (U < exp(-δ/T)) then S := Y
  }
  T = T × α
}
until stopping criterion is met

```

FIGURE 2. Basic simulated annealing for minimization problems.

are allowed and the process becomes “frozen.” The algorithm terminates when the stopping criterion is met. Common stopping criteria, and the ones used for the work in this paper, are to stop the search after a fixed number MaxIL of inner loops have been executed, or else when some maximum number MUL of consecutive unproductive inner loops have been executed (i.e., without a single move having been accepted). Generally the best state achieved so far is also recorded (because the search may actually move out of it and subsequently be unable to find a state of similar quality). At the end of each inner loop the temperature is lowered. The simplest way of lowering the temperature is to multiply by a constant cooling factor α in the range $(0, \dots, 1)$; this is known as *geometric cooling*. The basic simulated annealing algorithm has proven remarkably effective over a range of problems.

APPENDIX B : TRUTH TABLES OF EVOLVED FUNCTIONS

We present the truth tables in Hex. For RSBFs on nine variables there are two classes of size 1 (containing respectively, the elements 000000000 and 111111111, 2 of size 3, and 56 of size 9. Simple mathematics reveals that for Walsh value of 32 there are 15 classes of size 9 and 1 of size 1, for Walsh value of -32 there are 13 of size 9 and 1 of size 3, and for Walsh value of 0 there are 28 of size 9, one of size 3, and 1 of size 1. All classes with elements of weight less than or equal to 3 can be fixed to zeroes. However, this search was unsuccessful but by removing the balance constraint, i.e., $W_f(\vec{0})$, we could easily obtain unbalanced third order *CI* functions with nonlinearity 240 and algebraic degree 5. This function has never previously been demonstrated. On the other hand, by allowing a spectrum to include values of $+16$ and -16 , we obtained a (9, 2, 6, 240) function.

The truth table of one of the evolved (9, 2, 6, 240) RSBF functions is given below. The value of Δ_f is 152. The annealing parameters are as follows: MIL = 1000, MaxIL = 800, MUL = 50 and $\alpha = 0.95$. The initial pattern had 14 classes with Walsh value 32, with value -32 , and 8 with value 16 and 8 with value -16 . All classes with weight ≤ 2 were fixed at 0. The singleton class containing 111111111 was fixed at value 32. The average time for each run was 22 min.

0576 7a2d 6bcc 0da6 6c8b b1a0 45b7 d87c 79b5 c0db 9a17 9c44 6027 de3a b681 3fa1
 3ec3 9a32 a150 b7ce d78d 476a 83e4 6171 6944 1c7a f3b8 1a9d 9a6c d447 4aeb 8917

The truth table of one of the evolved [9, 3, 5, 240] RSBF functions is given below. The value of Δ_f is 80. The annealing parameters are as follows: MIL = 1,000, MaxIL = 1,000, MUL = 50, and $\alpha = 0.98$. There were 13 successes from 50 runs. The average time for each run was 22 minutes. Later experiments (with MIL = 8,000, MUL = 200 and MaxIL = 2000) produced 8 successes from 10 runs. The average time for each run was 1 h 10 min.

177a 2e98 18a9 d2d1 43c5 9c92 b649 a647 640a b467 c7f4 8709 db38 24d6 8839 712e
 7c24 41cd 9b64 692b a06b be31 d42e 05d2 e29b 5a91 0965 e23c 85c5 1b86 3b52 5ce8

The truth table of an evolved bent RSBF on 10 variables is given below. The annealing parameters are as follows: MIL = 2,000, MaxIL = 1,000, MUL = 50 and $\alpha = 0.98$. The average time for each run was 55 min.

7ffa ab9d dc9a d6f2 b2a1 82d8 e238 bb5c cb19 9957 c11c b2c4 b958 0fc4 8fce 63e5
 e1ce 0283 8393 277b b006 12f0 de5d b020 df97 76c0 05be e425 c1ef b0a9 285a a932
 fc17 e4ec 410c 845e 955f 964a 5c3a 3a9b 9a01 153d 424c ef40 f7e8 36a3 9b41 0c14
 e7aa 922e 7f68 e445 4177 8ab8 f861 5926 b042 f8ab 9f04 c9d2 48c1 729c 89c6 1a49

The truth table in Hex (lowest index first) of an evolved (10, 3, 5, 480) RSBF is given below having $\Delta_f = 192$. The annealing parameters are as follows: MIL = 1,000, MaxIL = 1,000, MUL = 50 and $\alpha = 0.98$. In 10 runs one success was achieved. The average time for each run was 55 min.

69c6 a578 9c23 6ad5 92b5 1d4b 7dc8 a227 d31c 8b36 56a7 61ca 2ea3 e4d4 9958 593e
 b64b 16e1 91da 4f29 662c dc3a 2957 e1d8 19e8 d81f ac35 b664 c6d3 3385 6696 4ba9
 9a3c 618f 4779 f803 d616 e398 21ae 5cd6 6969 5cb1 a2e4 1f9c 58c7 276a ad53 a691
 4693 bdc1 e694 12ee 99e5 0a76 cb38 3d25 a579 a64a 5b0e 9473 3d28 d26d 70cb c996