Chapter

Implementing Symmetric Cryptography Using Sequence of Semi-Bent Functions

Samed Bajrić

Abstract

Symmetric cryptography is a cornerstone of everyday digital security, where two parties must share a common key to communicate. The most common primitives in symmetric cryptography are stream ciphers and block ciphers that guarantee confidentiality of communications and hash functions for integrity. Thus, for securing our everyday life communication, it is necessary to be convinced by the security level provided by all the symmetric-key cryptographic primitives. The most important part of a stream cipher is the key stream generator, which provides the overall security for stream ciphers. Nonlinear Boolean functions were preferred for a long time to construct the key stream generator. In order to resist several known attacks, many requirements have been proposed on the Boolean functions. Attacks against the cryptosystems have forced deep research on Boolean function to allow us a more secure encryption. In this work we describe all main requirements for constructing of cryptographically significant Boolean functions. Moreover, we provide a construction of Boolean functions (semi-bent Boolean functions) which can be used in the construction of orthogonal variable spreading factor codes used in code division multiple access (CDMA) systems as well as in certain cryptographic applications.

Keywords: symmetric cryptography, Boolean functions, Walsh spectrum, nonlinearity, resiliency, (fast) algebraic attack

1. Introduction

Cryptography has become a branch of information theory and is used within a mathematical approach to study the transmission of information from place to place. In a modern society, exchange and storage of information in an efficient, reliable, and secure manner are of fundamental importance. Applications of cryptography are present in many aspects of our society, and they include authentication and encryption (bank cards, wireless telephone, e-commerce), access control (car lock systems, ski lifts), and payment (prepaid telephone cards, e-cash). Behind all the previously mentioned applications, an underlying cryptographic system has to satisfy a number of security goals. Some important aspects in information security are data confidentiality, data integrity, authentication, and non-repudiation, and some of these goals will be elaborated later in the framework of Boolean

functions. Therefore, cryptography is evermore important for business and industry as well as for society at large.

A classic example of a cryptosystem is depicted in **Figure 1**. Such a cryptosystem primitive is also called symmetric-key encryption algorithm, since the transmitted message (plaintext) is encrypted (into ciphertext) and decrypted with the same secret key which is shared between both sender and recipient. Symmetric cryptography is best introduced with an easy-to-understand problem: There are two users, Alice and Bob, who want to communicate over an insecure channel. The actual problem starts with the bad guy, Oscar, who has access to the channel, for instance, by hacking into an Internet router or by listening to the radio signals of a Wi-Fi communication. This type of unauthorized listening is called eavesdropping. Obviously, there are many situations in which Alice and Bob would prefer to communicate without Oscar listening. For instance, if Alice and Bob represent two offices of a car manufacturer, and they are transmitting documents containing the business strategy for the introduction of new car models in the next few years, these documents should not get into the hands of their competitors or of foreign intelligence agencies for that matter. In this situation, symmetric cryptography offers a powerful solution: Alice encrypts her message m using a symmetric algorithm, yielding the ciphertext c. Bob receives the ciphertext and decrypts the message. Decryption is, thus, the inverse process of encryption. What is the advantage? If we have a strong encryption algorithm, the ciphertext will look like random bits to Oscar and will contain no information whatsoever that is useful to him.

Symmetric-key cryptography comprises two large families of cryptographic primitives, namely, block and stream ciphers (see **Figure 2**). Since both block and stream ciphers provide significant performance improvement compared to public-key encryption techniques, they are commonly used as encryption schemes in practice. However, the design rules for these two primitives are quite different.

In general, symmetric-key cryptography is much more computationally efficient than public-key cryptography (approximately 1000 faster), and it requires shorter key length to ensure the same level of security. On the other hand, every pair of users that wants to communicate using symmetric encryption must share a common secret key. If n users want to ensure a pairwise secure communication, a total of $\frac{n(n-1)}{2}$ secret keys need to be exchanged, and every user must store and keep safe $n-1$ different secret keys, which is in many cases highly impractical. In comparison, public-key cryptography offers a functionality of only keeping a single private key secret.

The security of symmetric cryptosystems is strongly influenced by Boolean functions. They are often used as nonlinear combining functions in stream ciphers based on linear feedback shift register. Those functions allow making the relationship between the plaintext and the ciphertext as complex as possible. More precisely, a bit of the ciphertext is obtained from a bit of the plaintext by adding

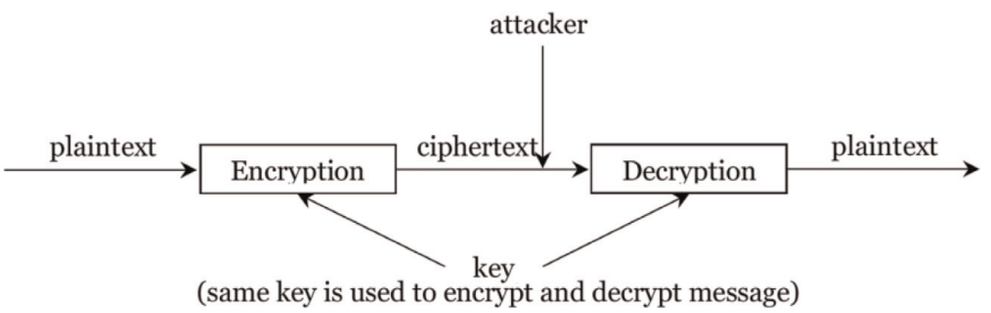

Figure 1.
Model of classic cryptosystem.

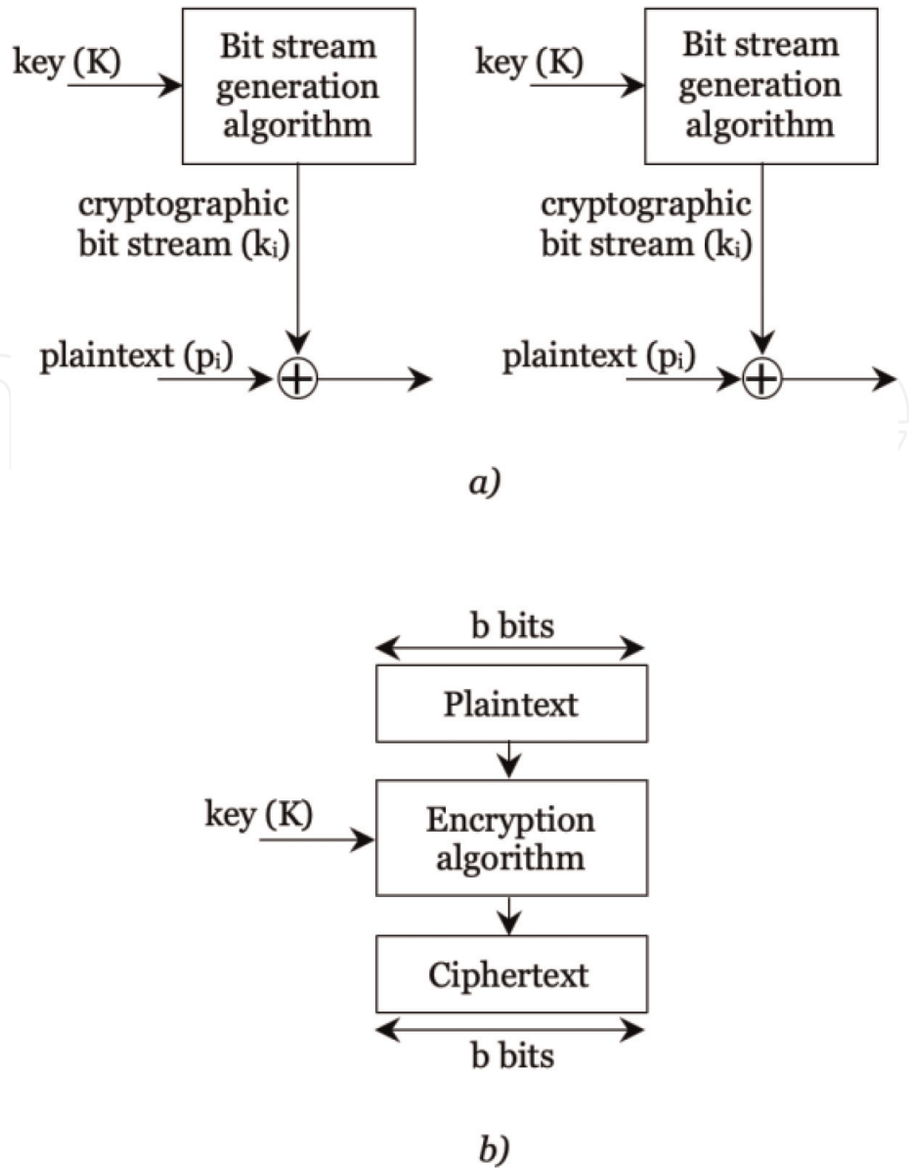

Figure 2.
Symmetric-key encryption schemes. (a) Stream cipher using algorithmic bit stream generator. (b) Block cipher.

bitwise a key digit (the output of the Boolean function) whose dependence upon the LFSR entries (the secret information) is nonlinear. Thus, the security of such cryptosystems deeply relies on the choice of the Boolean function because the complexity of the relationship between the plaintext and the ciphertext depends entirely on the Boolean function. Indeed, some properties of the Boolean function can be exploited to gain access to the contents of encrypted messages, even if the key is unknown. Therefore, Boolean functions need to have some important characteristics that are called security criteria to resist several types of attacks (see Section 3). Furthermore, the research fields of Boolean functions regarding the cryptography include the design and implementation, the properties of Boolean functions, the construction and counting of Boolean functions with certain properties, the trade-off between different properties, and the properties according to new attacks.

A special class of Boolean functions defined as semi-bent function has been introduced in 1994, by scientists Chee, Lee, and Kim [1]. The motivation for their study is firstly related to their use in cryptography (in the design of cryptographic functions). Indeed, semi-bent functions can be balanced and resilient. They also possess various desirable characteristics such as low autocorrelation, a maximal

nonlinearity among balanced plateaued functions, but they cannot have high algebraic degree. In terms of linear feedback shift-register synthesis, they are usually generated by certain power polynomials over a finite field and in addition are characterized by a low cross-correlation and high nonlinearity. Besides their practical use in cryptography, they are also widely used in code division multiple access (CDMA) communication systems for sequence design [2, 3]. In this context, families of maximum length linear feedback shift-register sequences having three-valued cross-correlation are used. Such sequences have received a lot of attention since the late 1960s and can be generated by a semi-bent function. Even though a lot of work has been done on semi-bent functions, there are a few generic methods of constructing semi-bent functions that can be found in the literature. The classification of these functions is still elusive, especially their construction are challenging problems. Some open problems and an overview of the known construction related on semi-bent functions can be found in the book of Mesnager [4]. The rest of this chapter is organized as follows. In Section 2 the essential background on Boolean functions is given. Some main requirements for constructing significant Boolean function are given in Section 3. An infinity class of semi-bent function specified by employing some sufficient conditions is given in Section 4. Some concluding remarks are given in Section 5.

2. Useful definitions and terms

Let $\mathbb{F}_2^n$ denote the n-dimensional vector space over the prime field $\mathbb{F}_2$. Let $x = (x_1, ..., x_n)$ be a vector over $\mathbb{F}_2$ of length n.

A *Boolean function* $f(x_1, ..., x_n)$ *in n-variables* is an arbitrary function from $\mathbb{F}_2^n$ to $\mathbb{F}_2$. It can also be interpreted as the output column of its *truth table*, i.e., a binary string of length 2^n,

$$f = [f(0, 0, ..., 0), f(1, 0, ..., 0), ..., f(1, 1, ..., 1)]. \tag{1}$$

An n-variable function f is said to be *balanced* if its output column in the truth table contains equal number of 1's and 0's.

Any Boolean function has a unique representation as a multivariate polynomial over Galois field of two elements, called *algebraic normal form* (ANF),

$$f(x_1, ..., x_n) = a_0 + \sum_{1 \le i \le n} a_i x_i + \sum_{1 \le i < j \le n} a_{ij} x_i x_j + ... + a_{12...n} x_1 x_2 ... x_n \tag{2}$$

where the coefficients $a_0, a_{ij}, ..., a_{12...n}$ belong to $\{0, 1\}$.

The *algebraic degree*, denoted by $\deg(f)$, is the number of variables in the highest order monomial with nonzero coefficient. A Boolean function with $\deg(f) \le 1$ is said to be *affine*, and the set of all n-variable affine functions is denoted by $\mathcal{A}_n$. An affine function with the constant term equal to zero is called a *linear* function.

The *nonlinearity* of an n-variable function f is $N_f = \min_{g \in \mathcal{A}_n} d(f, g)$, which measures the minimum distance between f and all n-variable affine functions.

Many properties of Boolean function can be deduced from its Walsh spectra. The *Walsh transform* of $f(x)$ in point $\omega \in \mathbb{F}_2^n$ is an integer-valued function over $\mathbb{F}_2^n$ defined by

$$W_f(\omega) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x) + x \cdot \omega}, \tag{3}$$

where $x \cdot \omega = x_1\omega_1 + ... + x_n\omega_n$ is the inner product of two vectors over $\mathbb{F}_2^n$. The set $\{W_f(\omega) : \omega \in \mathbb{F}_2^n\}$ is called the Walsh spectrum of f.

A Boolean function $f(x)$ is called *plateaued* if its Walsh spectrum only takes three values, 0 and $\pm 2^{\lambda}$, where λ is some positive integer.

Two Boolean functions $f(x), g(x)$ are said to be a pair of *disjoint spectra* functions if

$$W_f(\omega) \cdot W_g(\omega) = 0. \tag{4}$$

for all $\omega \in \mathbb{F}_2^n$.

In terms of Walsh spectra, the nonlinearity of f is given by

$$N_f = 2^{n-1} - \frac{1}{2} \max_{\omega \in \mathbb{F}_2^n} \left| W_f(\omega) \right|. \tag{5}$$

A function is balanced if and only if $W_f(0) = 0$, i.e., $\#\{x | f(x) = 0\} = \#\{x | f(x) = 1\}$.

An n-variable Boolean function f is said to be *bent* if its Walsh transform takes only two values $\pm 2^{\frac{n}{2}}$. Moreover, f is said to be *semi-bent* function if for all $\omega \in \mathbb{F}_2^n$

$$W_f(\omega) \in \begin{cases} \left\{0, \pm 2^{\frac{n+1}{2}}\right\}, \textit{if } n \textit{ is odd} \\ \left\{0, \pm 2^{\frac{n+2}{2}}\right\}, \textit{if } n \textit{ is even} \end{cases}. \tag{6}$$

The *derivative* of $f(x)$ at $a \in \mathbb{F}_2^n$, denoted by $D_a f(x)$, is a Boolean function defined by $D_a f(x) = f(x + a) + f(x)$, for all $x \in \mathbb{F}_2^n$. The notion of the derivative of a Boolean function is extended to higher orders as follows.

Suppose $\{a_1, ..., a_k\}$ is a basis of a k dimensional subspace V of $\mathbb{F}_2^n$. The k-th derivative of f with respect to V, denoted by $D_V f(x)$, is a Boolean function defined by

$$D_V f(x) = D_{a_k} D_{a_{k-1}} ... D_{a_1} f(x), \tag{7}$$

for all $x \in \mathbb{F}_2^n$.

3. Cryptographic requirements for constructing Boolean functions

One of the fundamental research topics in cryptography is the construction of cryptographically significant Boolean functions, that is, a function which possesses some of the following properties:

1. High ***algebraic degree*** aims to increase the linear complexity in ciphers. Using Boolean functions of high degree in block ciphers leads to more complicated systems of equations describing the cipher and hence makes cryptanalysis of the cipher more difficult. All cryptosystems using Boolean functions for confusion can be attacked if the functions have relatively low algebraic degree, i.e., the Berlekamp-Massey attack [5] or the Ronjom-Helleseth attack [6] can be applied. Note that the algebraic degree of a Boolean function in n-variables is at most n.

2. In order to prevent the system from leaking statistical dependence between the input and output, the concept of ***balancedness*** implies that a given Boolean function outputs equally many zeros and ones over all possible input values. To avoid distinguishing attacks [7], cryptographic function must be balanced.

Note that the algebraic degree of a Boolean balanced function in n-variables is at most $n-1$.

3. High ***nonlinearity*** is one of the most important properties in the design of symmetric-key cryptosystems, since it directly affects the resistance of the cipher to majority of cryptanalytic techniques. The nonlinearity simply measures the Hamming distance to the set of all affine functions. Therefore, a high nonlinearity implies a better resistance to affine approximation attacks [8]. According to the definition of nonlinearity, all affine functions have zero nonlinearity. On the other hand, a Boolean function having nonzero nonlinearity implies the function is not affine. Thus, the nonlinearity of a nonlinear Boolean function cannot exceed 2^{n-1}. On an even size Boolean space, there is a class of Boolean functions, called *bent* functions, that have maximum nonlinearity ($2^{n-1}-2^{\frac{n}{2}-1}$). In general, it is not an easy problem to identify all Boolean functions with high nonlinearity. However, this problem has been completely solved for quadratic Boolean functions (Boolean functions with the algebraic degree 2).

4. In order to avoid correlation attack [9], the concept of ***correlation immune*** of order m implies that any sub-function deduced from a given Boolean function by fixing at most m inputs has the same output distribution as a given Boolean function. Correlation immune has long been recognized as one of the critical indicators of nonlinear combining functions of shift registers in stream generators. Moreover, if a balanced Boolean function f is correlation immune of order m, then f is said to be *m-resilient*. When used in stream cipher systems, a Boolean function is required to have high nonlinearity and resiliency for protection against correlation attacks. It is actually very difficult to find a balanced Boolean function which has a high correlation immunity order and at the same time has a high nonlinearity.

5. Optimal ***algebraic immunity*** aims to provide resistance against algebraic attack. The algebraic immunity is the minimum value of d such that a given Boolean function f or its complement $1+f$ admits an annihilator (a nonzero Boolean function g such that $fg=0$) of algebraic degree d. In ciphers, Boolean functions with high algebraic immunity should be used in order to avoid the application of algebraic cryptanalysis [10]. Recall that algebraic attacks recover the secret key, or at least the initialization of the system, by solving a system of multivariate algebraic equations that describes a cipher. Although a high algebraic immunity is the necessary cryptographic requirement, it is not sufficient, because of a more general kind of attack introduced by Courtois [11] in 2003 called fast algebraic attack. It is well-known that maximum algebraic immunity of n-variable Boolean function is $\lceil \frac{n}{2} \rceil$. The problem of efficiently constructing balanced Boolean functions with optimal algebraic immunity is thus of great significance. Moreover, several examples of functions having optimal algebraic immunity could be found but no example of correlation immune Boolean function with optimal algebraic immunity.

However, the major problem in construction of cryptographically strong functions is that the multiple criteria mentioned above have to be satisfied at the same time, while there exist intrinsic trade-offs between them. Such properties allow the system designer to quantify the level of resistance of the system to attacks. Since the number of Boolean functions in n-variables is 2^{2^n}, an exhaustive search of functions

which satisfy some of the properties above is practically impossible (unless the input variable space n is quite small). Indeed, the difficulty precisely lies in finding the best trade-offs between all criteria and proposing concrete constructions of functions achieving them. Thus, bringing new construction methods of these functions is still a vivid research activity.

By (n, m, d, N_f) function we specify an n-variable, m-resilient Boolean function f, algebraic degree d, and nonlinearity N_f. Siegenthaler [9] proved that $m + d \leq n + 1$ if $m \leq n - 2$. The exact nature of trade-offs among order of correlation immunity, nonlinearity, and algebraic degree has also been investigated, for instance, ([12, 13]. Using the above bounds, one may naturally try to provide the construction of an (n, m, d, N_f) function for any given n and m while at the same time attempting to optimize d and N_f . This optimization can be efficiently done for a small number of variables $n \leq 5$, but even some interesting open problems for n>5, related to the existence of $(8, 1, 6, 116)$ and $(7, 2, 4, 56)$ functions, were settled using some sophisticated computer search and theoretical results [14]. The importance of finding these optimized functions in small number of variables lies in the fact that one can use these functions recursively to obtain new instances of optimal functions in larger number of variables. For instance, Tarannikov [15] has provided a construction technique of optimized resilient Boolean functions with maximum possible nonlinearity. Basically Tarannikov's construction is a recursive one, and using this technique and taking an (n, m, d, N_f) optimized function, such as the $(7, 2, 4, 56)$ function, one can generate a sequence of optimal plateaued $(7 + 3i, 2 + 2i, 4 + i, 2^{7+3i-1} - 2^{2+2i+1})$ functions, $(10, 4, 5, 480)$, $(13, 6, 6, 3968)$, $(16, 8, 7, 32256)$, *etc*. A modified version of Tarannikov's construction was presented in [16]. A construction of Boolean functions with maximum nonlinearity and small order of resiliency has also been considered in [17]. Later, Carlet [18] proposed a general framework for these iterative concatenation methods, unifying most of these techniques into a single method called "indirect sum." This construction leads to a multiple branching infinite tree of functions, but in order to employ this construction in the design of optimal plateaued functions in an iterative manner, there are certain conditions imposed on the initial pairs of disjoint spectra functions.

A recursive construction method of optimal plateaued functions (the functions of the form $(n, m, n - m - 1, 2^{n-1} - 2^{m+1})$ and for $m > \frac{n}{2} - 2$) is given in [19]. The iteration once again employs a $(7, 2, 4, 56)$ function, whose 6-variable sub-functions have disjoint spectra, to construct a sequence of $(7 + 4i, 2 + 3i, 4 + i, 2^{7+4i-1} - 2^{2+3i+1})$ optimal plateaued functions (whose $(7 + 4i - 1)$-variable sub-functions are again disjoint spectra functions). Nevertheless, this iterative method generates the functions with relatively large order of resiliency ($(11, 5, 5, 964)$, $(15, 8, 6, 15872)$, $(19, 11, 7, 258048)$, *etc*.), and in addition it only gives one infinite sequence of optimal plateaued functions. For instance, in the first step of iteration, an optimal plateaued $(11, 5, 5, 964)$ function is generated whose 10-variable sub-functions are again disjoint spectra functions (two $(10, 5, 4, 452)$ disjoint spectra functions), thus leaving some open slots concerning the construction of optimal plateaued functions when $n = 8, 9, 10$. On the other hand, a modified Tarannikov construction has a slightly different effect, since the resiliency is increased by two at each step of iteration (but the degree is also increased by one) and the iteration step is three instead of four. Still, optimal plateaued functions cannot be generated for $n = 8$ or $n = 9$ using the particular $(7, 2, 4, 56)$ function.

The idea of employing a set of disjoint spectra functions in construction of highly nonlinear resilient functions was firstly elaborated in [16]. Later, the sets of disjoint spectra functions were successfully used in constructions of almost optimal

resilient functions. The generalized Maiorana-McFarland (GMM) construction method for obtaining the almost optimal resilient functions has been proposed in [20]. Namely, this construction generates the functions with relatively large number of variables and small order of resiliency. The resulting functions cannot be viewed as a pair of disjoint spectra almost optimal resilient functions. Recently, Zhang and Pasalic used GMM technique to obtain the strictly optimal resilient functions with high nonlinearity and good algebraic properties [21]. The design of some balanced functions that also achieve currently best known nonlinearity can be found in [22]. Although these construction methods achieve currently the best nonlinearity for a given function, these methods are only efficient for relatively large input space of variables.

4. A construction of semi-bent Boolean functions

As it is described in the previous section, in the design of cryptographic functions, there is a need to consider various nonlinear characteristics simultaneously. But some characteristics restrict each other. Bent functions, for example, have maximum nonlinearity and satisfy the propagation criteria with respect to every nonzero vector over the Boolean spaces on which they are defined. However, bent functions are not balanced and exist only on even size Boolean spaces. Furthermore, bent functions are not correlation immune, and they are not suitable for use in cryptosystems. Partially bent functions are highly nonlinear and can be balanced. However, except for bent functions, partially bent functions have nonzero linear structures that are cryptographically undesirable. For these reasons, people study other classes of Boolean functions to try to overcome the disadvantage of bent functions or partially bent functions. The class of plateaued Boolean functions is one candidate that is defined by a series of inequalities and examines the critical case of each inequality. Compared with other functions, plateaued functions may reach the upper bound on nonlinearity given by the inequalities.

In what follows we specify a simple generic method for deriving semi-bent functions. This method is deduced from two bent functions whose derivatives differ by a constant one. It should be noticed that there are strong connections behind the concepts of bentness and semi-bentness though many questions remain unanswered. In particular, it is not settled how the cardinality of the whole class of bent functions relates to the class of semi-bent functions. Most notably, it appears that certain classes of semi-bent functions derived in [23] defined for even n are not extendable to bent functions in $n+2$ variables. In [24] and recently in [25], a sufficient condition on two bent functions g and h used in the construction of semi-bent functions was given as the following theorem.

Theorem 1. *Let n be even, and suppose that f and g are two bent Boolean functions in n-variables. If there exists an $a \in \mathbb{F}_2^n$ such that $D_a f(x) = D_a g(x) + 1$, then the function*

$$h(x) = f(x) + g(x) + D_a f(x) + D_a[f(x)g(x)] \tag{8}$$

is a semi-bent function in even number of variables.

This condition immediately implies the possibility of constructing infinite classes of semi-bent functions using known classes of quadratic bent functions. Notice that all quadratic Boolean functions (including bent and semi-bent functions) are classified up to equivalence and any quadratic bent function is affine equivalent to the canonical form given by $\sum_{i=1}^{n/2} x_{2i-1}x_{2i}$.

One may define a Boolean function f with n even to be a quadratic bent function of the form $f(x) = \sum_{i=1}^{n} b_i x_i + \sum_{1 \le i < j \le n} c_{i,j} x_i x_j$ for suitably chosen $b_i, c_{i,j} \in \mathbb{F}_2$. Furthermore, let g be a Boolean function defined as $g(x) = f(x) + \sum_{i=1}^{n} \alpha_i x_i$, where $\alpha_i \in \mathbb{F}_2$. Then, if $a \in \mathbb{F}_2^n$ is such that $a \cdot \alpha = 1$, it can be shown that the function

$$h(x) = f(x) + g(x) + D_a f(x) + D_a[f(x)g(x)]$$

is a quadratic semi-bent Boolean function.

Another related approach, though without restriction on the degree of a single bent function used, is given by the following result.

Theorem 2. *Let f be bent Boolean function in even number of variables. For $a, \alpha \in \mathbb{F}_2^n$ such that $a \cdot \alpha = 1$ define the function g as either*

$$g(x) = \begin{cases} f(x) + \alpha \cdot x + d \\ f(x + a) + \alpha \cdot x + d \end{cases}, \tag{9}$$

where $d \in \mathbb{F}_2$. Then, the function

$$h(x) = f(x) + g(x) + D_a f(x) + D_a[f(x)g(x)]$$

is a semi-bent function.

Proof. Obviously, in both cases g is also a bent function, and if $g(x) = f(x) + \alpha x + d$, we have

$$\begin{aligned} D_a f(x) + D_a g(x) &= [f(x) + f(x + a)] + [g(x) + g(x + a)] \\ &= [f(x) + f(x + a)] + [f(x) + \alpha x + d + f(x + a) + \alpha x + a\alpha + d] \\ &= a \cdot \alpha = 1. \end{aligned}$$

A similar calculation gives that

$$D_a f(x) + D_a g(x) = 1 \text{ if } g(x) = f(x + a) + \alpha x + d.$$

By Theorem 1 we deduce that $h(x) = f(x) + g(x) + D_a f(x) + D_a[f(x)g(x)]$ is a semi-bent function. q.e.d.

This result enables us to construct, for even n, an infinite sequence of semi-bent functions from bent functions. It would be of interest to find other examples or classes of bent functions g_1, g_2, apart from using affine equivalent functions g_1 and g_2, satisfying $D_a g_1(x) = D_a g_2(x) + 1$. This appears to be a nontrivial task since apart from establishing the fact that the used bent functions are indeed affine inequivalent, at the same time, their derivatives need to satisfy the condition in Theorem 1.

Example 1. Let $f(x_1, x_2, x_3, x_4, x_5, x_6) = x_1x_3x_4 + x_2x_3x_4 + x_1x_5x_6 + x_2x_5x_6 + x_1x_2 + x_3x_5 + x_4x_6 + x_5x_6$ be a bent function of degree 3 over $\mathbb{F}_2^6$. Take $a = (0, 0, 1, 0, 0, 0)$ and $\alpha = (1, 0, 1, 0, 0, 0)$ such that $a \cdot \alpha = 1$. Define the function g as either

$$g(x) = \begin{cases} f(x) + x_1 + x_3 \\ f(x + a) + x_1 + x_3 \end{cases} = \begin{cases} f(x) + x_1 + x_3 \\ f(x) + x_1x_4 + x_2x_4 + x_1 + x_3 + x_5 \end{cases},$$

where $d = 0 \in \mathbb{F}_2$.

Let us take $g(x) = f(x) + x_1 + x_3$. We have

$$D_a f(x) = f(x) + f(x+a) = f(x) + f(x) + x_1x_4 + x_2x_4 + x_5 = x_1x_4 + x_2x_4 + x_5,$$

so that

$$f(x) + g(x) + D_a f(x) = x_1x_4 + x_2x_4 + x_1 + x_3 + x_5.$$

Then, using the idempotent property of Boolean ring,

$$\begin{aligned} f(x)g(x) &= f(x)(f(x) + x_1 + x_3) = f(x)(1 + x_1 + x_3) \\ &= (x_1x_3x_4 + x_2x_3x_4 + x_1x_5x_6 + x_2x_5x_6 + x_1x_2 + x_3x_5 + x_4x_6 + x_5x_6)(1 + x_1 + x_5) \\ &= x_1x_2x_3x_4 + x_1x_2x_5x_6 + x_2x_3x_4x_5 + x_1x_2x_5 + x_1x_3x_4 + x_1x_3x_5 + x_1x_4x_6 \\ &\quad + x_2x_3x_4 + x_4x_5x_6 + x_4x_6. \\ f(x+a)g(x+a) &= f(x+a)(f(x+a) + x_1 + x_3 + 1) = f(x+a)(x_1 + x_3) \\ &= (f(x) + x_1x_4 + x_2x_4 + x_5)(x_1 + x_3) \\ &= f(x)(x_1 + x_3) + (x_1x_4 + x_2x_4 + x_5)(x_1 + x_3). \end{aligned}$$

After some simplification, we get

$$\begin{aligned} D_a[f(x)g(x)] &= f(x)g(x) + f(x+a)g(x+a) \\ &= f(x) + (x_1x_4 + x_2x_4 + x_5)(x_1 + x_3) \\ &= x_1x_5x_6 + x_2x_5x_6 + x_1x_2 + x_1x_4 + x_1x_5 + x_2x_4 + x_4x_6 + x_5x_6. \end{aligned}$$

Finally,

$$\begin{aligned} h(x) &= f(x) + g(x) + D_a f(x) + D_a[f(x)g(x)] \\ &= x_1x_5x_6 + x_2x_5x_6 + x_1x_2 + x_1x_5 + x_4x_6 + x_5x_6 + x_1 + x_3 + x_5. \end{aligned}$$

It is easy to compute the Walsh spectrum of function $h(x)$, i.e., $W_h(\omega) = \{0, \pm 16\}$, which means that $h(x)$ is a semi-bent function.

Notice that the standard derivation rule for multiplication does not apply for our definition of derivatives. Indeed, the derivative $D_a[f(x)g(x)] = f(x)g(x) + f(x+a)g(x+a)$ is different from $g(x)D_a f(x) + f(x)D_a g(x) = f(x+a)g(x) + f(x)g(x+a)$. Furthermore, using the fact that $D_a D_a f(x) = 0$ for any Boolean function f, we have

$$\begin{aligned} D_a h(x) &= h(x) + h(x+a) \\ &= f(x) + g(x) + D_a f(x) + D_a[f(x)g(x)] + f(x+a) + g(x+a) \\ &\quad + D_a f(x+a) + D_a[f(x+a)g(x+a)] \\ &= D_a f(x) + D_a g(x) + D_a D_a f(x) + D_a D_a[f(x)g(x)] \\ &= D_a f(x) + D_a g(x) = 1. \end{aligned}$$

Thus, the element a is always a linear structure of $h(x)$. Nevertheless, we show that under certain sufficient conditions, a is the only linear structure of $h(x)$. We have the following theorem.

Theorem 3. *Let h be defined as in Theorem 2, and assume that a bent function $f(x)$ is such that* $\deg(D_b f(x)) > 1$, *for any $b \in \mathbb{F}_2^n \setminus \{0\}$. Then h has a single linear structure, that is, $D_b h(x) = h(x) + h(x+b)$ is a constant function only for $b = a$.*

Proof. Assume that $g(x) = f(x+a) + \alpha x + d$. Without loss of generality, we can take $d = 0$. Then,

$$\begin{aligned} D_b f(x) + D_b g(x) &= [f(x) + f(x+b)] + [g(x) + g(x+b)] \\ &= [f(x) + f(x+b)] + [f(x+a) + \alpha(x+a) + d + f(x+a+b) + \alpha(x+a+b) + d] \\ &= D_b D_a f(x) + \alpha b, \end{aligned}$$

where $D_b D_a f(x) = f(x) + f(x+a) + f(x+b) + f(x+a+b)$, and therefore

$$D_b h(x) = D_b D_a f(x) + \alpha b + D_b D_a f(x) + D_b D_a [f(x)g(x)] = D_b D_a [f(x)g(x)] + \alpha b.$$

Hence, $D_b h(x)$ is constant if and only if $D_b D_a [f(x)g(x)]$ is constant. But,

$$\begin{aligned} D_b D_a [f(x)g(x)] &= D_b [f(x)g(x) + f(x+a)g(x+a)] \\ &= D_b [f(x)(f(x+a) + \alpha x) + f(x+a)(f(x) + \alpha(x+a))] \\ &= D_b [\alpha x(f(x) + f(x+a)) + \alpha a f(x+a)] \\ &= D_b [\alpha x(f(x) + f(x+a)) + f(x+a)] \\ &= \alpha x D_b D_a f(x) + \alpha b [f(x+b) + f(x+a+b)] + f(x+a) + f(x+a+b). \end{aligned}$$

Thus, if $\alpha b = 0$, then $D_b h(x)$ is constant if and only if

$$\begin{gathered} \alpha x D_b D_a f(x) = f(x+a) + f(x+a+b) \\ \alpha x [f(x) + f(x+a) + f(x+b) + f(x+a+b)] = f(x+a) + f(x+a+b) \\ (\alpha x + 1)[f(x+a) + f(x+a+b)] + \alpha x [f(x) + f(x+b)] = 0 \\ (\alpha x + 1) D_b f(x+a) + \alpha x D_b f(x) = 0 \\ \alpha x D_b f(x+a) + \alpha x D_b f(x) + D_b f(x+a) = 0. \end{gathered}$$

There are four possible cases:

1. $\alpha x D_b f(x+a) = \alpha x D_b f(x) = D_b f(x+a) = 0$, i.e.,
   $D_b f(x+a) = 0 \Leftrightarrow f(x+a) = f(x+a+b) \Rightarrow b = 0$. A contradiction.

2. $\alpha x D_b f(x+a) = \alpha x D_b f(x) = 1 \wedge D_b f(x+a) = 0$, i.e.,
   $D_b f(x+a) = 0 \Rightarrow b = 0$. A contradiction.

3. $\alpha x D_b f(x+a) = 0 \wedge \alpha x D_b f(x) = D_b f(x+a) = 1$, i.e.,
   $D_b f(x+a) = 0 \Rightarrow b = 0$. A contradiction.

4. $\alpha x D_b f(x+a) = D_b f(x+a) = 1 \wedge \alpha x D_b f(x) = 0$, i.e.,
   $D_b f(x+a) = 0 \Rightarrow b = 0$. A contradiction.

On the other hand, if $\alpha b = 1$, then $D_b h(x)$ is constant if and only if

$$\begin{gathered} \alpha x D_b D_a f(x) = f(x+a) + f(x+b) \\ \alpha x [f(x) + f(x+a) + f(x+b) + f(x+a+b)] = f(x+a) + f(x+b) \\ (\alpha x + 1)[f(x+a) + f(x+b)] + \alpha x [f(x) + f(x+a+b)] = 0. \end{gathered}$$

It is obvious that $f(x+a)=f(x+b)$ is equivalent to $f(x)=f(x+a+b)$. Thus, the above equation is constant if and only if $f(x+a)=f(x+b)$, which implies that $a=b$. The sufficiency of this condition is obvious. For the necessity, we first observe that for $a\neq b$ the functions $f(x+a)+f(x+b)$ and $f(x)+f(x+a+b)$ being derivatives of a bent function f are both nonconstant. Then, assuming that

$$D_bD_af(x)=f(x)+f(x+a)+f(x+b)+f(x+a+b)=0,$$

it would imply that $f(x+a)+f(x+b)$ is constant, a contradiction. On the other hand, the function $\alpha x D_bD_af(x)$ cannot be balanced, unless $D_bD_af(x)=\alpha x$. Because of the assumption, $\deg(f(x+a)+f(x+b))>1$ and therefore cannot be equal to αx.

The proof for the case $g(x)=f(x)+\alpha x+d$ is similar as above, and it is omitted here. q.e.d.

Notice the condition in Theorem 3 that $\deg(D_bf(x))>1$ is sufficient but may not be necessary. An analysis of other cryptographic criteria appears to be difficult due to the dependency of h on the choice of a bent function f and the use of the derivative $D_a[f(x)g(x)]$ in its definition, which is illustrated in the following example.

Example 2. Let n be even and $f(x,y)=x\cdot y$, where $x,y\in\mathbb{F}_2^k$ is a bent function and belongs to the Maiorana-McFarland class. Then, defining $g(x,y)=f(x+a,y+b)+(\alpha,\beta)\cdot(x,y)$ for a nonzero $(a,b)\in\mathbb{F}_2^k\times\mathbb{F}_2^k$ such that $(\alpha,\beta)\cdot(a,b)=1$, we have

$$g(x,y)=x\cdot y+(\alpha+b)\cdot x+(a+\beta)\cdot y+a\cdot b,$$

which is clearly a bent function obtained by adding an affine function to f. Similarly,

$D_{(a,b)}f(x,y)=x\cdot b+a\cdot y+a\cdot b$, so that

$$f(x,y)+g(x,y)+D_{(a,b)}f(x,y)=\alpha\cdot x+\beta\cdot y.$$

Then, using the idempotent property of Boolean ring,

$$\begin{aligned}f(x,y)\cdot g(x,y)&=(x\cdot y)(x\cdot y+(\alpha+b)\cdot x+(a+\beta)\cdot y+a\cdot b)\\&=(1+a\cdot b)(x,y)+((\alpha+b)\cdot x+(a+\beta)\cdot y)(x\cdot y).\end{aligned}$$

Note that the first term is a quadratic function and the second term is cubic. After some simplifications we have

$$\begin{aligned}D_{(a,b)}[f(x,y)g(x,y)]&=x\cdot y+(b\cdot x+a\cdot y+a\cdot b)(1+a\cdot b+\alpha\cdot x+\alpha\cdot a+b\cdot x\\&\quad+a\cdot b+a\cdot y+\beta\cdot y+\beta\cdot b)\\&=x\cdot y+(b\cdot x+a\cdot y+a\cdot b)(\alpha\cdot x+b\cdot x+a\cdot y+\beta\cdot y+a\cdot b+\beta\cdot b)\\&=x\cdot y+(b\cdot x+a\cdot y+a\cdot b)((\alpha+b)\cdot x+(\beta+a)\cdot y+a\cdot b+\beta\cdot b).\end{aligned}$$

Finally,

$$\begin{aligned}h(x,y)&=f(x,y)+g(x,y)+D_{(a,b)}f(x,y)+D_{(a,b)}[f(x,y)g(x,y)]\\&=x\cdot y+(\alpha\cdot x+\beta\cdot y)(b\cdot x+a\cdot y+a\cdot b+1)+(b\cdot x+a\cdot y+a\cdot b)(1+\beta\cdot b).\end{aligned}$$

More precisely, it can be illustrated using Example 1.

Example 3. Let
$f(x_1, x_2, x_3, x_4, x_5, x_6) = x_1x_3x_4 + x_2x_3x_4 + x_1x_5x_6 + x_2x_5x_6 + x_1x_2 + x_3x_5 +$
$x_4x_6 + x_5x_6$ be a bent function of degree 3 over $\mathbb{F}_2^6$. Take $a = (0, 0, 1, 0, 0, 0)$ and $\alpha = (1, 0, 1, 0, 0, 0)$ such that $a \cdot \alpha = 1$. Define the function g as $g(x) = f(x) + x_1 + x_3$. By Example 1 we have

$$h(x) = x_1x_5x_6 + x_2x_5x_6 + x_1x_2 + x_1x_5 + x_4x_6 + x_5x_6 + x_1 + x_3 + x_5.$$

Moreover, by Theorem *2* h has a single linear structure only for $b = a$. Indeed,

$$\begin{aligned} D_a h(x) &= h(x) + h(x + a) \\ &= x_1x_5x_6 + x_2x_5x_6 + x_1x_2 + x_1x_5 + x_4x_6 + x_5x_6 + x_1 + x_3 + x_5 + \\ &\quad + x_1x_5x_6 + x_2x_5x_6 + x_1x_2 + x_1x_5 + x_4x_6 + x_5x_6 + x_1 + x_3 + 1 + x_5 \\ &= 1. \end{aligned}$$

5. Conclusions

The need for the most possible secure cryptographic primitives in cipher systems is of great importance. In the case of stream ciphers, most of the reliability and security lies in the Boolean functions. For the cryptographic point of view to be good, a Boolean function should possess several cryptographic properties mentioned in this work. Very often such properties contradict each other. Therefore, the problem of constructing Boolean functions with stronger cryptographic properties is still a vivid research activity. We may also require new properties because attacks never stop. On the other hand, semi-bent functions are interesting for defending against the so-called soft output joint attack on pseudorandom generators, which are used in the IS-95 standard of code division multiple access technology. In this work we present an infinite sequence of semi-bent functions using known classes of quadratic bent functions. The construction of other classes of infinite sequences of semi-bent functions is an interesting research challenge.

Acknowledgements

This research was supported by the Slovenian Research Agency (research program P2-0037).

Author details

Samed Bajrić
Laboratory for Open Systems and Networks, Jožef Stefan Institute, Ljubljana, Slovenia

*Address all correspondence to: samed@e5.ijs.si

References

[1] Chee S, Lee S, Kim K. Semi-bent functions. In: Advances in Cryptology-ASIACRYPT94. 1994

[2] Ding C, Mesnager S, Tang C, Xiong M. Cyclic Bent Functions and their Applications in Codes, Codebooks, Designs, MUBs and Sequences. 2018. Available from: https://arxiv.org/pdf/1811.07725.pdf

[3] Hunt FH, Smith DH. The construction of orthogonal variable spreading factor codes from semi-bent functions. IEEE Transactions on Wireless Communications. 2012;**11**(8): 2970-2975

[4] Mesnager S. Bent Functions—Fundamentals and Results. Switzerland: Springer International Publishing; 2016

[5] Massey JL. Shift-register synthesis and BCH decoding. IEEE Transactions on Information Theory. 1969;**15**(1): 1222-1127

[6] Ronjom S, Helleseth T. A new attack on the filter generator. IEEE Transactions on Information Theory. 2007;**53**(5):1752-1758

[7] Andreeva E, Bogdanov A, Mennink B. Towards understanding the known-key security of block ciphers. In: International Workshop on Fast Software Encryption. Springer; 2013

[8] Liu J, Mesnager S, Chen L. On the nonlinearity of S-boxes and linear codes. Cryptography and Communications. 2016:345-361

[9] Siegenthaler T. Correlation-immunity of nonlinear combining functions for cryptographic applications. IEEE Transactions on Information Theory. 1984;**30**:776-780

[10] Tang D, Carlet C, Tang X. Highly nonlinear Boolean functions with optimal algebraic immunity and good behavior against fast algebraic attacks. In: Transactions on Information Theory; Institute of Electrical and Electronics Engineers. 2013. pp. 653-664

[11] Courtois N, Meier W. Algebraic attacks on stream ciphers with linear feedback. In: EUROCRYPT 2003. LNCS 2656. Springer; 2003. pp. 345-359

[12] Han G, Li X, Zhou Q, Zheng D, Li H. 1-resilient Boolean functions on even variables with almost perfect algebraic immunity. Security and Communication Networks. 2017;**2017**:9

[13] Li LY, Zhang WG. Construction of resilient Boolean functions with high nonlinearity and good algebraic degree. Security and Communication Networks. 2015:2909-2916

[14] Maitra S, Pasalic E. Further constructions of resilient Boolean functions with very high nonlinearity. IEEE Transactions on Information Theory. 2002;**48**(7):1825-1834

[15] Tarannikov Y. On resilient Boolean functions with maximal possible nonlinearity. In: Indocrypt 2000. LNCS 1977. Springer-Verlag; 2000. pp. 19-30

[16] Pasalic E, Johansson T, Maitra S, Sarkar P. New constructions of resilient and correlation immune Boolean functions achieving upper bounds on nonlinearity. In: Workshop on Coding and Cryptography. Elsevier Science; 2001. pp. 425-435

[17] Sarkar P, Maitra S. Construction of nonlinear Boolean functions with important cryptographic properties. In: Advances in Cryptology EUROCRYPT 2000. LNCS 1807. Springer-Verlag; 2000. pp. 485-506

[18] Carlet C. On the secondary constructions of resilient and bent

functions. In: Coding, Cryptography and Combinatorics. Basel: Birkahauser Verlag; 2004. pp. 3-28

[19] Gao SM, Zhao Y, Zhao Z. Walsh spectrum of cryptographically concatenating functions and its applications in constructing resilient Boolean functions. The Journal of Computer Information Systems. 2011; **7**(4):1074-1081

[20] Zhang W, Xiao G. Constructions of almost optimal resilient Boolean functions on large even number of variables. IEEE Transactions on Information Theory. 2009;**55**(12): 5822-5831

[21] Zhang W, Pasalic E. Improving the lower bound on the maximum nonlinearity of 1-resilient Boolean functions and designing functions satisfying all cryptographic criteria. Information Sciences. 2017;**376**:21-30

[22] Zhang F, Wei Y, Pasalic E, Xia S. Large sets of disjoint spectra plateaued functions inequivalent to partially linear functions. IEEE Transactions on Information Theory. 2018:2987-2999

[23] Carlet C, Mesnager S. On semibent Boolean functions. IEEE Transactions on Information Theory. 2012;**58**(5): 3287-3292

[24] Sun G, Wu C. Construction of semi-bent Boolean functions in even number of variables. Chinese Journal of Electronics. 2009;**18**(2):231-237

[25] Pasalic E, Gangopadhyay S, Zhang W, Bajric S. Design methods for semi-bent functions. Information Processing Letters. 2019:61-70