Discrete Applied Mathematics 174 (2014) 1–10

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Cryptographic properties of the hidden weighted bit function

Qichun Wang [a,*], Claude Carlet [b], Pantelimon Stănică [c], Chik How Tan [a]

[a] Temasek Laboratories, National University of Singapore, 117411, Singapore

[b] LAGA, Department of Mathematics, University of Paris 8 (and Paris 13 and CNRS), Saint-Denis cedex 02, France

[c] Department of Applied Mathematics, Naval Postgraduate School, Monterey, CA 93943-5216, USA

ARTICLE INFO

Article history:
Received 8 March 2013
Received in revised form 30 December 2013
Accepted 10 January 2014
Available online 31 January 2014

Keywords:
Hidden weighted bit function
Algebraic immunity
Nonlinearity
BDD-based attack

ABSTRACT

The hidden weighted bit function (HWBF), introduced by R. Bryant in IEEE Trans. Comp. 40 and revisited by D. Knuth in Vol. 4 of The Art of Computer Programming, is a function that seems to be the simplest one with exponential Binary Decision Diagram (BDD) size. This property is interesting from a cryptographic viewpoint since BDD-based attacks are receiving more attention in the cryptographic community. But, to be usable in stream ciphers, the functions must also satisfy all the other main criteria. In this paper, we investigate the cryptographic properties of the HWBF and prove that it is balanced, with optimum algebraic degree and satisfies the strict avalanche criterion. We calculate its exact nonlinearity and give a lower bound on its algebraic immunity. Moreover, we investigate its normality and its resistance against fast algebraic attacks. The HWBF is simple, can be implemented efficiently, has a high BDD size and rather good cryptographic properties, if we take into account that its number of variables can be much larger than for other functions with the same implementation efficiency. Therefore, the HWBF is a good candidate for being used in real ciphers. Indeed, contrary to the case of symmetric functions, which allow such fast implementation but also offer to the attacker some specific possibilities due to their symmetry, its structure is not suspected to be related to such dedicated attacks.

1. Introduction

To resist the main known attacks, Boolean functions used in stream ciphers should have good cryptographic properties: balancedness, high algebraic degree, high algebraic immunity, high nonlinearity and good immunity to fast algebraic attacks. Up to now, many classes of Boolean functions with high algebraic immunity have been introduced [1,5–8,13,14,21,22,27–30,35–37,41–43]. However, most of them do not satisfy all the necessary criteria and the few classes which do satisfy, are not very efficiently implementable; moreover, none of the papers studying these classes took BDD-based attacks into consideration.

BDD-based attacks were first introduced by Krause in 2002 [19]. They might be efficient against LFSR-based generators [19,20,33,34]. To resist BDD-based attacks, a Boolean function should have a high BDD size.

The hidden weighted bit function (HWBF) was proposed by Bryant [2]. It is an easily defined function that has an exponential BDD size, but has a VLSI implementation with low area-time complexity [2]. In [18], Knuth reproved Bryant's theorem stating that the HWBF has a large BDD size, regardless of how one reorders its variables. Therefore, the HWBF can resist BDD-based attacks and could be implemented efficiently. However, many other cryptographic properties of the HWBF were still unknown.

* Corresponding author. Tel.: +65 85552483; fax: +65 68726840.
E-mail addresses: qcwang@fudan.edu.cn, tslwq@nus.edu.sg (Q. Wang), claude.carlet@univ-paris8.fr (C. Carlet), pstanica@nps.edu (P. Stănică), tsltch@nus.edu.sg (C.H. Tan).

0166-218X/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.dam.2014.01.010

In this paper, we investigate the important cryptographic properties of this function and show that it is balanced, with optimum algebraic degree and satisfies the strict avalanche criterion. We calculate exactly its nonlinearity and give a lower bound on its algebraic immunity. These two parameters are not at an optimal level (but they are not low either). The function would then not be a good choice as a filter function (in a stream cipher) if it was implemented with a number of variables which is usual for other functions such as the Carlet–Feng function [7] (say, between 16 and 20 variables). But its very simple structure allows using it with many more variables (at least twice) and then the values of the nonlinearity and of the algebraic immunity allow good resistance to the main attacks while the function has still a much faster hardware implementation, which allows the stream cipher to be in the same time robust against the main known attacks and fast. This is also the case of some symmetric functions (whose output depend only on the Hamming weight of the input), but the specificity of symmetric functions represents a threat since it has the reputation of allowing dedicated attacks. The structure of the HWBF function is almost as simple as that of symmetric functions but the fact that, for a given Hamming weight different from 0 and n of the input, the output is non-constant (and is even almost balanced in the case of Hamming weights near $n/2$, that is, for most probable ones), the function represents a better tradeoff between robustness and speed. We also investigate the normality and give some computational results on the resistance of the HWBF against fast algebraic attacks, revealing that the HWBF displays good behavior against fast algebraic attacks.

The paper is organized as follows. In Section 2, the necessary background is established. We then investigate the cryptographic properties of the HWBF in Section 3. We end in Section 4 with conclusions.

2. Preliminaries

Let $\mathbb{F}_2^n$ be the n-dimensional vector space over the finite field $\mathbb{F}_2$. We denote by B_n the set of all n-variable Boolean functions, from $\mathbb{F}_2^n$ into $\mathbb{F}_2$.

Cosets of vector subspaces are also called *flats*. Let $f \in B_n$ and E be any flat. If the restriction of f to E, denoted by $f|_E$, is constant (respectively affine), then E is called a constant (respectively affine) flat for f.

Any Boolean function $f \in B_n$ can be uniquely represented as a multivariate polynomial in $\mathbb{F}_2[x_1, \ldots, x_n]$,

$$f(x_1, \ldots, x_n) = \sum_{K \subseteq \{1,2,\ldots,n\}} a_K \prod_{k \in K} x_k,$$

which is called its algebraic normal form (ANF). The algebraic degree of f, denoted by $\deg(f)$, is the number of variables in the highest order term with nonzero coefficient.

A Boolean function is *affine* if there exists no term of degree strictly greater than 1 in the ANF. The set of all affine functions is denoted by A_n.

Let

$$1_f = \{x \in \mathbb{F}_2^n | f(x) = 1\}, \qquad 0_f = \{x \in \mathbb{F}_2^n | f(x) = 0\},$$

be the support of a Boolean function f, respectively, its complement. The cardinality of 1_f is called the *Hamming weight* of f, and will be denoted by $wt(f)$. The *Hamming distance* between two functions f and g is the Hamming weight of $f + g$, and will be denoted by $d(f, g)$. We say that an n-variable Boolean function f is *balanced* if $wt(f) = 2^{n-1}$.

Let $f \in B_n$. The *nonlinearity* of f is its distance from the set of all n-variable affine functions, that is,

$$nl(f) = \min_{g \in A_n} d(f, g).$$

The nonlinearity of an n-variable Boolean function is bounded above by $2^{n-1} - 2^{n/2-1}$, and a function is said to be *bent* if it achieves this bound. Clearly, bent functions exist only for even n and it is known that the algebraic degree of a bent function is bounded above by $\frac{n}{2}$ [4,32]. The *r-order nonlinearity*, denoted by $nl_r(f)$, is its distance from the set of all n-variable functions of algebraic degrees at most r.

A Boolean function $f \in B_n$ is called *k-normal* (respectively, *k-weakly-normal*) if there exist a k-dimensional constant (respectively, affine) flat for f. If $k = \lceil \frac{n}{2} \rceil$, f is simply called a *normal* (respectively, *weakly-normal*) function.

For any $f \in B_n$, a nonzero function $g \in B_n$ is called an *annihilator* of f if fg (the function defined by $fg(x) = f(x)g(x)$) is null, and the *algebraic immunity* of f, denoted by $\mathcal{AI}(f)$, is the minimum value of d such that f or $f + 1$ admits an annihilator of degree d [24]. It is known that the algebraic immunity of an n-variable Boolean function is bounded above by $\lceil \frac{n}{2} \rceil$ [11].

To resist algebraic attacks, a Boolean function f should have a high algebraic immunity, which implies that the nonlinearity of f is also not very low since, according to Lobanov's bound [23]:

$$nl(f) \geq 2 \sum_{i=0}^{\mathcal{AI}(f)-2} \binom{n-1}{i}.$$

To resist fast algebraic attacks, a high algebraic immunity is not sufficient. If we can find g of low degree and h of algebraic degree not much larger than $n/2$ such that $fg = h$, then f is considered to be weak against fast algebraic attacks [10,17]. The higher order nonlinearities of a function with high (fast) algebraic immunity is also not very low [3,26,39].

The *Walsh transform* of a given function $f \in B_n$ is the integer-valued function over $\mathbb{F}_2^n$ defined by

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