Available online at www.sciencedirect.com
ScienceDirect
Information Sciences 178 (2008) 1903–1916
INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL
www.elsevier.com/locate/ins

WG: A family of stream ciphers with designed randomness properties

Yassir Nawaz *, Guang Gong

Department of Electrical and Computer Engineering, University of Waterloo, 200 University Avenue West, Waterloo, ON, Canada N2L 3G1

Received 18 July 2006; received in revised form 1 December 2007; accepted 3 December 2007

Abstract

In this paper we present a family of stream ciphers which generate a keystream with ideal two-level autocorrelation. The design also guarantees other randomness properties, i.e., balance, long period, ideal tuple distribution, and high and exact linear complexity. We discuss how these properties are achieved by the proposed design and show how to select various parameters to obtain an efficient stream cipher for the desired security level. We also show that the proposed generators are secure against time/memory/data tradeoff attacks, algebraic attacks and correlation attacks. Finally we present WG-128[1] as a concrete example of a WG stream cipher with a key size of 128 bits.

Keywords: WG; Welch–Gong; Stream cipher; Randomness properties; Two-level autocorrelation

1. Introduction

Traditionally many hardware oriented stream ciphers have been built using linear feedback shift registers (LFSRs) and Boolean functions with compact Algebraic Normal Forms (ANFs). This allowed for a very small and efficient implementation in hardware. However following the discovery of algebraic attacks [5,6,12,2], using Boolean functions with compact ANFs is no longer secure. One way to defeat the algebraic attacks is to use nonlinear feedback shift registers (NFSRs) or more generally, update the state of the stream cipher nonlinearly. Many hardware oriented stream ciphers in the ECRYPT stream cipher project have been designed according to this approach [8]. Whereas the tools to analyze the LFSR based stream cipher designs are well developed, very few theoretical results exist for the NFSRs. Therefore the security of such stream ciphers rely on the difficulty of analyzing the design itself. An alternative approach is to design a stream cipher in such a way that it is very easy to analyze. This allows the designers to explicitly prove various security properties of the design. The WG family of stream ciphers have been designed using this approach. To defeat the algebraic attacks on LFSR based stream ciphers, WG relies on nonlinear Boolean functions with large number of inputs, high degree and complex ANF

* Corresponding author. Tel.: +1 519 888 4567x2140; fax: +1 519 746 7260.
E-mail address: ynawaz@engmail.uwaterloo.ca (Y. Nawaz).

[1] WG-128 is a slightly modified version of the WG stream cipher which is a phase 2 candidate in profile 2 of the ECRYPT stream cipher project: eSTREAM. See http://www.ecrypt.eu.org/stream/

0020-0255/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2007.12.002

forms. To overcome the implementation complexity the Boolean function has been designed in polynomial form instead of ANF form. Implementation of such functions require small finite field multipliers. This results in a moderate increase in the hardware requirements, however the design is still practical.

A WG stream cipher consists of a WG keystream generator which produce a long pseudo-random keystream. The keystream is XORed with the plaintext to produce the ciphertext. From here onwards we will focus on the design and randomness properties of the WG keystream generators. The WG keystream generators use Welch–Gong (WG) transformations as the filtering functions. The WG transformations have very large ANFs and can be implemented in polynomial form using finite field arithmetic. These transformations correspond to the WG transformation sequences that were discovered by Golomb and coworkers [16]. In 2002, Gong and Youssef presented several cryptographic properties of WG transformation sequences. These properties make them a suitable candidate for cryptographic use.

The paper is organized as follows. In Section 2 we define the basic terms and notations. In Section 3 we first show that it is impractical to use WG transformation sequence generators from [10] as keystream generators in practical stream ciphers. We then examine the possibility of using a small WG transformation as a filtering function in a nonlinear filter to build a keystream generator. In Section 4 we show which cryptographic properties of the WG transformation sequences can be preserved in this design. For the properties that cannot be preserved completely we list the possible tradeoffs. Then in Section 5 we show how various parameters of the designs can be selected to design WG stream ciphers for various security levels and application requirements. Finally in Section 6 we present WG-128, as a concrete example of a WG keystream generator with a secret key of size 128 bits. Note that WG-128 is a slightly modified version of the WG stream cipher which is a phase 2 candidate in profile 2 of the ECRYPT stream cipher project: eSTREAM [8].

2. Preliminaries

In this section we define and explain the terms and notations that we will use to describe the WG keystream generators and their operation.

- $\mathbb{F}_2 = GF(2)$, finite field with 2 elements: 0 and 1.
- Let $n > m$, $\mathbb{F}_{2^n} = GF(2^n)$, i.e., extension field of $GF(2)$ with 2^n elements, and $\mathbb{F}_{2^m} = GF(2^m)$, i.e., extension field of $GF(2)$ with 2^m elements.
- $Tr_m^n(x) = \sum_{i=0}^{n/m-1} x^{2^{mi}}, x \in \mathbb{F}_{2^n}$, the trace function from $\mathbb{F}_{2^n} \to \mathbb{F}_{2^m}$.
- Polynomial basis of $\mathbb{F}_{2^m}$: Let α be the root of the primitive polynomial that generates $\mathbb{F}_{2^m}$. Then $\{1, \alpha, \alpha^2, \ldots, \alpha^{m-1}\}$ is the polynomial basis of $\mathbb{F}_{2^m}$ over $\mathbb{F}_2$.
- Normal basis of $\mathbb{F}_{2^m}$: Let γ be an element of $\mathbb{F}_{2^m}$ such that $\{\gamma, \gamma^{2^1}, \gamma^{2^2}, \ldots, \gamma^{2^{m-1}}\}$ is a basis of $\mathbb{F}_{2^m}$ over $\mathbb{F}_2$. Then $\{\gamma, \gamma^{2^1}, \gamma^{2^2}, \ldots, \gamma^{2^{m-1}}\}$ is a normal basis of $\mathbb{F}_{2^m}$ over $\mathbb{F}_2$.
- Any Boolean function has a unique representation as a multivariate polynomial over $\mathbb{F}_2$, called the *algebraic normal form* (ANF)

$$h(x_1, \ldots, x_n) = a_0 + \sum_{1 \leqslant i \leqslant n} a_i x_i + \sum_{1 \leqslant i < j \leqslant n} a_{i,j} x_i x_j + \cdots + a_{1,2,\ldots,n} x_1 x_2 \ldots x_n,$$

  where the coefficients $a_0, a_i, a_{i,j}, \ldots, a_{1,2,\ldots,n} \in \mathbb{F}_2$.
- The *algebraic degree*, $\deg(h)$, of a boolean function h is the number of variables in the highest order term with non-zero coefficient. A Boolean function is *affine* if there exists no term of degree >1 in the ANF and the set of all affine functions is denoted A_n.
- The *nonlinearity* of an n-variable function f is the minimum distance from the set of all n-variable affine functions, i.e.,

$$nl(h) = \min_{g \in A(n)} (d(h, g)).$$

- Let $X_1, X_2, \ldots, X_n$ be independent binary random variables with equal probability of 0 and 1. A Boolean function $f(x_1, x_2, \ldots, x_n)$ is said to be tth order correlation immune, if for each subset of t variables