Cryptanalysis Results on Achterbahn Stream Cipher by Combinatorial Equivalence Approach

Hope4Sec Team

Hope4Sec, Tallinn, Estonia
contact@hope4sec.eu

November 5, 2023

Abstract

This white paper presents the evaluation results on Achterbahn cryptanalysis based on *Combinatorial Equivalence* (CE) Approach. This approach enables to outperform existing attacks with a rather limited amount of ciphertext. With 2^{38} bits of ciphertext, Achterbahn-128's cryptanalysis has a complexity of $\mathcal{O}(2^{87})$ to retrieve the 351 internal state bits once the key setup is completed. As for Achterbachn-80, only 2^{35} bits of ciphertext data are required with an attack complexity of $\mathcal{O}(2^{53})$ to retrieve the 297 internal state bits.

Keywords: Combinatorial Equivalence Cryptanalysis, Nonlinear Shift Register, Stream Cipher, Algebraic Attack, Bent function, Achterbahn.

1 Introduction

Stream ciphers [8] are one of the main cryptographic primitives used in symmetric cryptography. At the origin, the first stream ciphers were built with Linear Feedback Shift Registers (LFSRs), where linearity is meant in the register update function while the combining function is meant to be non-linear to break the intrinsic linear properties of the sequences produced by the register. The most common design is the combiner generator, depicted in Figure 1 to which most of other the stream cipher designs can be reduced.

Most stream ciphers use combined or non-linearly filtered LFSRs [8]. However, their security has been questioned over the years. The natural evolution of these systems is towards the use of Non Linear Feedback Shift Registers (NLFSRs) [5] and Boolean functions with good cryptographic properties according to the general structure in Figure 1. If for the latter a substantial body of knowledge exists [**?**] – with however a large number of problems still open as soon as the number of variables exceeds ten – the study of NLFSRs is still in its infancy. To date, for example, no NLFSRs are known in maximum period for lengths greater than 31 [1] when looking for rather simple, sparse feedback

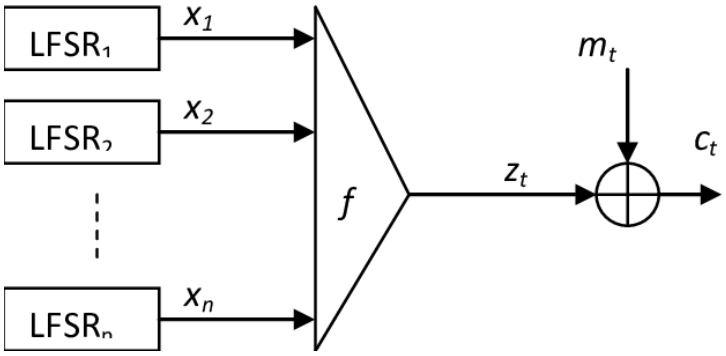

Figure 1: General Structure of LFSR-based Combiner

polynomials (a desirable property when dealing with implementation aspects). But the clever combination of these registers can lead to systems with effective security. The two best examples are Achterbahn [4] and Trivium [3].

Attacking this new class of systems remains an open problem [9]. For real systems, no effective cryptanalysis is known that could seriously question their security. Algebraic or statistical attacks by correlation are no longer applicable. It is therefore necessary to consider a radically different approach that is neither algebraic nor statistical.

We are working on such an approach. It is combinatorial in nature. The principle is to rewrite the system to produce a combinatorial equivalent system[1] and to translate the key search by a known combinatorial problem, of a complexity that is reachable in practice (data and time) and lower than the complexity of known attacks in the algebraic or statistical domains. We have named this approach CE (*Combinatorial Equivalence*). This attack has been successfully tested on the concept cipher *Cipherbent6* [6].

We consider real-life cryptanalysis only. We denote it *Effective Cryptanalysis*.

Definition 1.1 (Effective Cryptanalysis)**.** *A cryptanalysis is called effective whenever it can be performed:*

- *In a limited time that allows the cryptanalysis to be played a finite number of times.*
- *On realistic input data sizes that are compatible with the operational reality of the use cases.*

For instance, any attack requiring data without changing keys far beyond the crypto-periods in use in the industrial or governmental world is not effective. Moreover known plaintext attack are realistic as long as the amount of data does not exceed a few kilobits. Any attack whose time complexity is greater than 2^{70} is not effective nowadays. However, to evaluate the effective security of a system, the size of the required input data N takes precedence over the time complexity T: a system which can be broken with a relatively

[1]This approach is also called *Equivalent Combinatorial Rewriting*.

small value of N despite a high complexity T is in practice considered less secure than a system requiring a much higher value of N even for a significantly lower complexity. It is not so much impossible to increase computing power as it is to have enough input data (at least conceptually) at least if standard cryptographic policies are applied (in particular concerning crypto-periods).

The paper is organized as follows: Section2 presents the specification of the Achterbahn stream cipher [4].Then in Section **??** we summarizes the initial cryptanalysis results and performances we have obtained on Achterbahn with respect to the CE cryptanalysis. Finally, Section 4 presents the future work to develop the CE cryptanalysis.

Disclaimer: *this white paper is not basically a research article in the usual sense. The CE technique is not public and is reserved for the industrial and governmental world. This paper is intended to present our results. The proof of our results can be provided on request by a challenge approach (sending an output sequence and returning the key) when applicable or after suitable agreement.*

2 Achterbahn Specifications

Achterbahn algorithm [4] belongs to the class of stream ciphers in which Non Linear Shift Registers (NLFSR) are combined by a non linear Boolean function. The concept cipher Cipherbent6 [6] is a minimal element of this class.

Two versions of Achterbahn do exist depending on the maximal key size: 80-bit and 128-bit keys. The main parameters are summarized in Table 1.

	ACHTERBAHN-80	ACHTERBAHN-128
Max. key length	80 bit	128 bit
Max. IV length	80 bit	128 bit
Max. frame length	244	244
Internal state	297 bit	351 bit

Table 1: Achterbahn parameters

2.1 Achterbahn-128

ACHTERBAHN-128 accommodates all key lengths between 40 and 128 and all IV - lengths between 0 and 128 that are multiples of eight. It is a keystream generator, consisting of 13 binary nonlinear feedback shift registers (NLFSRs) whose polynomials are very dense. The length of register i is $L_i = 21 + i$ for $i = 0, 1, \ldots, 12$. These NLFSR polynomials are primitive in the sense that their periods T_i are equal to $2^{L_i} - 1$.

The sequence which is used as an input to the Boolean combining function is a shifted version of itself. The shift amount is a fixed but register-dependent value.

The Boolean combining function $F(x_0, x_1, \ldots, x_{12}) : \mathbb{F}_2^{13} \to \mathbb{F}_2$ has a very dense ANF and exhibits strong cryptographic properties:

- F is balanced.
- F has algebraic degree 4.
- F is correlation immune of order 8.
- F has nonlinearity 3 584.
- F has algebraic immunity 4.
- Each variable in its ANF appears in at least one monomial of degree 4 such that the shift register lengths corresponding to the variables in that monomial are pairwise relatively prime.

2.2 Achterbahn-80

ACHTERBAHN-80 can be used with key lengths 40, 48, 56, 64, 72, and 80. All IV -lengths between 0 and 80 can be used provided that the IV -length is divisible by eight. n the same way, the sequence which is used as an input to the Boolean combining function is a shifted version of itself. The shift amount is a fixed but register-dependent value.

Achterbahn-80 consists of 11 registers, which are the same ones as in the above case, except for the first and the last ones. The Boolean combining function G is a sub-function of F :

$$G(x_1, \ldots, x_{11}) = F(0, x_1, \ldots, x_{11}, 0)$$

G also has a very dense ANF and exhibits strong cryptographic properties:

- G is balanced.
- G has algebraic degree 4.
- G is correlation immune of order 6.
- F has nonlinearity 896.
- F has algebraic immunity 4.
- Each variable in its ANF appears in at least one monomial of degree 4 such that the shift register lengths corresponding to the variables in that monomial are pairwise relatively prime.

As we can see, Achterbahn-128 contains Achterbahn-80 as a substructure.

2.3 Achterbahn NLFSRs

Nonlinear linear shift registers (NLFSR) of respective length 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33 have all maximum period. Their feedback polynomial are all dense

and of high degree. For instance, NLFSR A_{12} has the following Algebraic Normal Form (ANF) for its feedback polynomial:

$$\begin{aligned} A_{12}(x_0, x_1, \ldots, x_{32}) \quad = \quad & x_0 \oplus x_2 \oplus x_7 \oplus x_9 \oplus x_{10} \oplus x_{15} \oplus x_{23} \oplus x_{25} \oplus x_{30} \oplus x_8 x_{15} \\ & \oplus x_{12} x_{16} \oplus x_{13} x_{15} \oplus x_{13} x_{25} \oplus x_1 x_8 x_{14} \oplus x_1 x_8 x_{18} \oplus x_8 x_{12} x_{16} \\ & \oplus x_8 x_{14} x_{18} \oplus x_8 x_{15} x_{16} \oplus x_8 x_{15} x_{17} \oplus x_{15} x_{17} x_{24} \oplus x_1 x_8 x_{14} x_{17} \\ & \oplus x_1 x_8 x_{17} x_{18} \oplus x_1 x_{14} x_{17} x_{24} \oplus x_1 x_{17} x_{18} x_{24} \oplus x_8 x_{12} x_{16} x_{17} \oplus \\ & x_8 x_{14} x_{17} x_{18} \oplus x_8 x_{15} x_{16} x_{17} \oplus x_{12} x_{16} x_{17} x_{24} \oplus x_{14} x_{17} x_{18} x_{24} \oplus \\ & x_{15} x_{16} x_{17} x_{24}. \end{aligned}$$

Other register polynomials can be found in [4].

3 Achterbahn Cryptanalysis Results

The aim of our work is to provide a cryptanalysis of both versions of Achterbahn which is as much effective as possible. From a limited size output (known plaintext attack and ciphertext only attack), we intend to recover the 297/351-bit initialization at time instant t. Recovering the initial key K and IV from this initialization is considered as easy.

The cryptanalysis of Achterbahn is in two parts [6]:

1. Obtaining a combinatorial equivalent of Achterbahn is a **one-time operation** that takes from 24 hours (version 1) to 48 hours (version 2) and has overall complexity of $\mathcal{O}(2^{42})$.
2. From a given combinatorial equivalent, the cryptanalysis part has been evaluated and partial simulations have confirmed our evaluations. Table 2 summarizes the different cryptanalysis results. For ciphertext-only attack and without loss of generalities, we suppose the underlying language is a English with ISO 646-US encoding.

Achterbahn version	Attack Type	Complexity	Data	Best Known Attack
Achterbahn-80	Known Plaintext	$\mathcal{O}(2^{53})$	2^{31}	$\mathcal{O}(2^{67})/2^{53}$ [7]
	Ciphertext Only	$\mathcal{O}(2^{53})$	2^{35}	-
Achterbahn-128	Known Plaintext	$\mathcal{O}(2^{87})$	2^{35}	$\mathcal{O}(2^{104})/2^{56}$ [7]
	Ciphertext Only	$\mathcal{O}(2^{87})$	2^{38}	-

Table 2: Cryptanalysis results for Achterbahn's versions compared to best known attacks

Experiments (partial simulations) have been conducted on two AMD Ryzen Threadripper 2990 WX 32-Core Processor x64 (256 Mb of RAM, 36 Tb HDD each) and NVIDIA GPU RTX GeForce 4090.

4 Conclusion et future work

The results obtained for Achterbahn with respect to *Equivalent Combinatorial Rewriting* outperform the few known cryptanalysis. These results are in line with the spirit of Effective cryptanalyses (Definition 1.1).

Very interestingly, our study confirmed a number of interesting points:

- the Boolean function F exhibits very remarkable structural properties that go beyond the cryptographic properties usually considered. The choice of this function is inevitably motivated by particular considerations that the authors have not explained.
- However, whatever the remarkable quality of this function, the ECR approach could be applied successfully, but it appeared to us that it was more the (relatively modest) size of the registers that made our approach effective. Larger register sizes (at least double), possibly fewer in number, would probably have made the ECR approach more difficult to apply. This is what we want to check, for example on the Trivium system.

Future research work also includes the application of our rewriting approach to block ciphers. The nature of cryptographic primitives being different from those in stream ciphers, new combinatorial objects are to be defined.

References

[1] P. Dąbrowski, G. Labuzek, T. Rachwalik and J. Szmidt. Searching for Nonlinear Feedback Shift Registers with Parallel Computing. *Information Processing Letters*, Vol. 114, Issue 5, pp. 268–272, 2014.

[2] B. A. Davey and H. A. Priestley. *Introduction to Lattices and Order*. Cambridge University Press, 2002.

[3] C. De Cannière, and B. Preneel. "Trivium specifications". eSTREAM submitted papers, 2005. Last retrieved on March 20th, 2023 on `https://www.ecrypt.eu.org/stream/ciphers/trivium/trivium.pdf`

[4] B. M. Gammel, R. Göttfert, O. Kniffler. "ACHTERBAHN-128/80". Achterbahn home page, 2006. Last retrieved on April 25th, 2023 on `http://www.matpack.de/achterbahn/Gammel_Goettfert_Kniffler_Achterbahn-128-80.pdf`

[5] S. W. Golomb. *Shift Register Sequences*. Aegean Park Press, 1981.

[6] Hope4Sec Crypto Team. Non-linear Shift Register-based Bent Combiner Cryptanalysis Results. Available on `https://hope4sec.eu/pages/cryptology-evaluation.html`, June 5^{th}, 2023.

[7] M. Naya-Plasencia. Cryptanalysis of Achterbahn-128/80 with a New Keystream Limitation. Research in Cryptology: Second Western European Workshop, WEWoRC. Revised Selected Papers, Lecture Notes in Computer Science. Vol. 4945, Springer. pp. 142-–152, 2007. Available on `https://www.ecrypt.eu.org/stream/papersdir/2007/004.pdf`

[8] R. A. Rueppel. *Analysis and Design of Stream Ciphers*. Springer Berlin Heidelberg, 1986.

[9] G. Yao. *Transformation and Security Analysis of NLFSR-based Stream Ciphers*. PhD Thesis, University of Melbourne, 2021.