Improved Boolean Combining Functions for Achterbahn

Berndt M. Gammel, Rainer Göttfert, and Oliver Kniffler

Infineon Technologies AG
P. O. Box 80 09 49
D-81609 Munich, Germany

e-mail:
berndt.gammel@infineon.com
rainer.goettfert@infineon.com
oliver.kniffler@infineon.com

October 14, 2005

Abstract

The Boolean combining function in the 80-bit-key stream cipher Achterbahn [1] is weak. Its major weakness consists of the fact that by setting two specific variables to zero, the function becomes linear. Its second weakness consists of the fact that it can be approximated by a linear function which agrees with the Boolean combining function with probability 3/4. By exploiting the first weakness, Johansson, Meier, and Muller [2] managed to break the reduced version of Achterbahn with complexity of 2^{56} steps and the full version of Achterbahn with complexity of 2^{73} steps. By exploiting the second weakness, they found a distinguishing attack which needs to process 2^{64} keystream bits. However, both weaknesses can be removed by simply adding three monomials to the initial Boolean combining function of Achterbahn. For the revised Boolean combining function—and with everything else of the algorithm left unchanged—the complexity of the first attack described in [2] becomes 2^{85} steps for the reduced version of Achterbahn and 2^{111} steps for the full version of Achterbahn. The complexity of the second attack described in [2] is raised to 2^{128} for both versions.

1 Introduction

Achterbahn [1] is a binary additive stream cipher designed for 80-bit-key security. The core of Achterbahn consists of eight binary nonlinear feedback shift registers (NLFSR's) of lengths between 22 and 31. The output sequences of the eight NLFSR's are combined by the Boolean function

$$R(x_1, \ldots, x_8) = x_1 + x_2 + x_3 + x_4 + x_5x_7 + x_6x_7 + x_6x_8 + x_5x_6x_7 + x_6x_7x_8. \qquad (1)$$

Although R is balanced and 4th-order correlation immune, it has the weakness that the entire nonlinear part of R vanishes if the two variables x_5 and x_6 are both set to zero. This weakness of R was exploited by Johansson, Meier, and Muller [2] in a clever cryptanalytic attack which we will call the *JMM1-attack* in the following. The JMM1-attack breaks the reduced version of Achterbahn with complexity of 2^{56} steps and the full version of Achterbahn with complexity of 2^{73} steps. The complexities are related to the shift register lengths as follows. The lengths of the NLFSR's that correspond to the input variables x_5 and x_6 are 27 and 28, respectively. The complexity of the JMM1-attack against the reduced version of Achterbahn is

$$2^{27+28+1} = 2^{56}.$$

In the full version of Achterbahn, the variability of the key-dependent shift register output functions must also be taken into account. That is, in the above formula, we have to add the maximum degrees of the filter polynomials (see [1]) to the shift register lengths. For the two shift registers of lengths 27 and 28, the maximum degrees of the filter polynomials are 8 and 9, respectively. This yields

$$2^{(27+8)+(28+9)+1} = 2^{73},$$

which is the complexity of the JMM1-attack against the full version of Achterbahn.

The second weakness of the function R in (1) consists of the fact that it agrees with each of the two linear approximations $l = x_1+x_2+x_3+x_4+x_6$ and $l' = x_1+x_2+x_3+x_4+x_7$ with probability 3/4. This weakness was exploited by Johansson, Meier, and Muller [2] in a distinguishing attack. We will refer to this attack as the *JMM2-attack* in the following. The JMM2-attack can distinguish the keystream of the reduced and the full version of Achterbahn from a true random sequence by processing 2^{64} keystream bits. The number $n = 2^{64}$ can be derived from the number e of variables in the given linear approximations and the probability p of agreement (or disagreement) between the Boolean function R and its linear approximation. Here we have $p = 3/4$ and $e = 5$. Let $\Delta = |p - 1/2| = |3/4 - 1/2| = 1/4$, then

$$n = \left(\frac{1}{2\Delta}\right)^{2^{e+1}} = 2^{64}. \tag{2}$$

2 Improved Boolean combining functions

The two weaknesses in Achterbahn's initial combining function R can easily be removed. To illustrate that point, we discuss two examples for improved Boolean combining functions R' and R''. Each function blasts the complexity of both the JMM1-attack and the JMM2-attack far beyond the complexity of exhaustive key search. In addition to the two functions discussed in this note, there are several other Boolean combining functions that would avert the two attacks on Achterbahn described in [2].

2.1 First example of an improved combining function

The simplest way to simultaneously remove both weaknesses in Achterbahn's initial combining function R is achieved by adding three quadratic terms to it. The revised Boolean combining function R' is given by

$$R' = R + x_5x_6 + x_5x_8 + x_7x_8, \tag{3}$$

where R is given in (1). The function R' is again balanced and 4th-order correlation immune. It contains all six quadratic terms that can be formed from the four variables x_5, x_6, x_7, and x_8. As a consequence, one must set at least three of the eight variables of R' to zero before R' becomes linear.

The best strategy for the JMM1-attack would be to set x_5, x_6, and x_7 to zero. The lengths of the corresponding shift registers are 27, 28, and 29, respectively. The maximum degrees of the corresponding filter polynomials are 8, 9, and 9, respectively. It follows that the complexity of the JMM1-attack against the reduced version of Achterbahn with the revised Boolean function R' is

$$2^{27+28+29+1} = 2^{85}.$$

The complexity against the full version is

$$2^{(27+8)+(28+9)+(29+9)+1} = 2^{111}.$$

We now investigate the effect of the JMM2-attack on Achterbahn with the revised Boolean function R'. First, we observe that the two best approximations of R' are

$$l_j = x_1 + x_2 + x_3 + x_4 + x_j \qquad \text{with } j = 5 \text{ and } j = 8.$$

Each of these linear functions agrees with R' with probability $p = 5/8$. Thus $\Delta = |5/8 - 1/2| = 1/8$, $e = 5$, and

$$n = \left(\frac{1}{2\Delta}\right)^{2^{e+1}} = 2^{128}.$$

Therefore, 2^{128} keystream bits must be processed in order to distinguish the keystream of Achterbahn from a random sequence.

We summarize the complexities of the JMM1 and the JMM2 attack against the Achterbahn stream cipher with revised Boolean function R' in Table 1 below. We include the complexity of the classical correlation attack of Siegentahler [3] as well.

	Reduced Achterbahn	Full Achterbahn
JMM1	2^{85}	2^{111}
JMM2	2^{128}	2^{128}
Siegenthaler	2^{123}	2^{159}

Table 1: Complexities of attacks against Achterbahn with revised Boolean combining function R'

2.2 Second example of an improved combining function

Another way to get rid of the weaknesses in Achterbahn's initial combining function R is to replace it by

$$R'' = x_1 + x_2 + x_3 + \sum_{4 \le i < j \le 8} x_i x_j + \sum_{4 \le i < j < k \le 8} x_i x_j x_k + \sum_{4 \le i < j < k < l \le 8} x_i x_j x_k x_l. \tag{4}$$

The function R'' is balanced and 3rd-order correlation immune. We have to set at least four variables to zero in order that R'' becomes linear. The best strategy for the JMM1-attack would be to set x_4, x_5, x_6, and x_7 to zero. The lengths of the corresponding shift registers and the maximum degrees of the corresponding filter polynomials are 26, 27, 28, 29, and 8, 8, 9, 9, respectively. The complexity of the JMM1-attack against the reduced version of Achterbahn therefore is

$$2^{26+27+28+29+1} = 2^{111}.$$

The complexity of the JMM1-attack against the full version of Achterbahn is

$$2^{(26+8)+(27+8)+(28+9)+(29+9)+1} = 2^{145}.$$

We now investigate the effect of the JMM2-attack on Achterbahn with revised Boolean combining function R''. The best linear approximations of R'' are the five linear functions

$$l_j = x_1 + x_2 + x_3 + x_j \quad \text{with } j = 4, 5, 6, 7, 8.$$

Each of these linear functions agrees with R'' with probability $p = 9/16$. Thus $\Delta = |9/16 - 1/2| = 1/16$ and $e = 4$. Using again formula (2), we obtain

$$n = \left(\frac{1}{2\Delta}\right)^{2^{e+1}} = 2^{96}.$$

Therefore, 2^{96} keystream bits must be processed in order to distinguish the keystream of Achterbahn from a random sequence. We summarize the complexities of the attacks in Table 2 below.

	Reduced Achterbahn	Full Achterbahn
JMM1	2^{111}	2^{145}
JMM2	2^{96}	2^{96}
Siegenthaler	2^{96}	2^{124}

Table 2: Complexities of attacks against Achterbahn with revised Boolean combining function R''

3 Conclusions

Due to weaknesses of the Boolean combining function in the initial proposal [1] of the Achterbahn stream cipher, a cryptanalytic attack and a distinguishing attack was found [2]. However, the deficiencies of Achterbahn's initial combining function R can easily be removed. We discussed two improved combining functions R' and R'' which do not have the identified weaknesses. If the initial combining function R is replaced by either the function R' or the function R''—and if everything else of the algorithm is left unchanged—then both of the attacks described in [2] are less efficient than exhaustive key search. On the other hand, the improved Boolean combining functions R' and R'' are still small enough to fit on a T-shirt.

References

[1] B. Gammel, R. Göttfert, and O. Kniffler: The Achterbahn Stream Cipher, eSTREAM, ECRYPT Stream Cipher Project, Report 2005/002, 2005. http://www.ecrypt.eu.org/stream/ciphers/achterbahn/achterbahn.pdf.

[2] T. Johansson, W. Meier, and F. Muller: Cryptanalysis of Achterbahn, Sept. 2005. http://www.ecrypt.eu.org/stream/papersdir/064.pdf.

[3] T. Siegenthaler: Correlation-immunity of nonlinear combining functions for cryptographic applications, *IEEE Trans. Inform. Theory* **IT-30**, 776–780, 1984.