Analysis of the SMS4 Block Cipher

Fen Liu[1], Wen Ji[1], Lei Hu[1], Jintai Ding[2],
Shuwang Lv[1], Andrei Pyshkin[3,⋆], and Ralf-Philipp Weinmann[3]

[1] State Key Laboratory of Information Security,
Graduate School of Chinese Academy of Sciences,
Beijing 100049, China
[2] Department of Mathematical Sciences,
University of Cincinnati,
Cincinnati, OH, 45221, USA
[3] Fachbereich Informatik,
Technische Universität Darmstadt,
64289 Darmstadt, Germany

Abstract. SMS4 is a 128-bit block cipher used in the WAPI standard for providing data confidentiality in wireless networks. In this paper we investigate and explain the origin of the S-Box employed by the cipher, show that an embedded cipher similar to BES can be obtained for SMS4 and demonstrate the fragility of the cipher design by giving variants that exhibit 2^{64} weak keys.

We also show attacks on reduced round versions of the cipher. The best practical attack we found is an integral attack that works on 10 rounds out of 32 rounds with a complexity of 2^{18} operations; it can be extended to 13 rounds using round key guesses, resulting in a complexity of 2^{114} operations and a data complexity of 2^{16} chosen pairs.

Keywords: block ciphers, cryptanalysis, UFN, algebraic structure.

1 Introduction

The Wired Authentication and Privacy Infrastructure (WAPI) standard is an alternative to the security mechanisms for wireless networks that are specified in IEEE 802.11i. It has been submitted to the International Standards Organization (ISO) by the Chinese Standards Association (SAC). Although it was subsequently rejected by the ISO in favour of IEEE 802.11i, WAPI still is officially mandated for securing wireless networks within China.

For protecting data packets, the WAPI standard references a 128-bit block cipher called SMS4 which initially was kept secret. In January 2006, the specification of this block cipher however was declassified and published [6]. Other than a differential power attack [11] in a Chinese journal, no analysis of this cipher has appeared in the open literature.

This document sheds light on the design of this block cipher and present a preliminary analysis of its strength against cryptanalytic attacks.

⋆ Supported by a stipend of the Marga und Kurt-Möllgaard-Stiftung.

J. Pieprzyk, H. Ghodosi, and E. Dawson (Eds.): ACISP 2007, LNCS 4586, pp. 158–170, 2007.

In Section 2 we give a description of the SMS4 cipher. In Section 3 we show how the SMS4 S-Box can be derived algebraically and how an embedding of SMS4 similar to the Big Encryption System (BES) can be obtained. Section 4 describes an practical integral attack on a 10-round version of SMS4 that can be extended to a theoretical attack on 13 rounds. Our results in Section 5 demonstrate the fragility of SMS4; we show that modifications of the round constants can lead to a large subspace of weak keys. Finally, in Section 6 we conclude this paper and summarize our findings.

1.1 Notation

In the following, we agree on the conventions used throughout the rest of this paper.

Since all operations of the cipher are defined on either 8-bit, 32-bit or 128-bit quantities, we shall use the following terminology: 8-bit values will simply be called *bytes*, 32-bit values *words* and 128-bit values will be called *blocks*. Word and block values shall be considered to be in big-endian order, i.e. the most-significant bit is in the leftmost position when writing the value as a bitstring.

Let $w \lll r$ denote a cyclic shift of the word w by r positions to the left. Sometimes we will need to write down blocks or words in which certain bytes are unknown. In these cases the symbol $\star$ shall denote bytes with unknown values.

To concatenate multiple byte values into a word and multiple word values into a block, we define a vector of bytes or words to be equivalent to a word respectively block value. To access individual bit ranges of a value w we shall use the notation $w_{[i\ldots j]}$ to extract bits i to j, e.g. for $w \in \mathbb{Z}_{2^{32}}$ the expression $w_{[7\ldots 0]}$ denotes the lowestmost byte of the word value w.

2 Description of the SMS4 Block Cipher

In this section we will give a top-down description of the SMS4 block cipher.

SMS4 is a 32 round unbalanced Feistel network; both the block and the key size are 128 bits. Following the terminology of [10], the cipher is a homogeneous, complete, source-heavy (96:32) UFN with 8 cycles.

Let the internal state be denoted by $\mathcal{S} = (S_1, S_2, S_3, S_4)$ where $S_i \in GF(2)^{32}$. The round keys of the cipher shall be denoted by $K_i \in GF(2)^{32}$.

Define the linear diffusion function λ as

$$\begin{aligned}\lambda : GF(2)^{32} &\to GF(2)^{32} \\ x &\mapsto x \oplus (x \lll 2) \oplus (x \lll 10) \oplus (x \lll 18) \oplus (x \lll 24)\end{aligned}$$

and the brick-layer function γ applying an 8-bit S-Box to the input 4 times in parallel as:

$$\begin{aligned}\gamma : GF(2)^{32} &\to GF(2)^{32} \\ x &\mapsto (\rho(x_{[31\ldots 24]}), \rho(x_{[23\ldots 16]}), \rho(x_{[15\ldots 8]}), \rho(x_{[7\ldots 0]}))\end{aligned}$$