Improved Linear Approximations of SNOW-V and SNOW-Vi

Zhen Shi, Chenhui Jin, and Yu Jin

Information Engineering University, Zhengzhou 450000, China
shizhenieu@126.com, jinchenhui@126.com

Abstract. in this paper, we improve the results of linear approximation of SNOW-V and SNOW-Vi. We optimized the automatic search program by replacing the S-box part with accurate characterizations of the Walsh spectral of S-boxes, which results in a series of trails with higher correlations. On the basis of existing results, we investigate the common features of linear approximation trails with high correlation, and search for more trails by exhausting free masks. By summing up the correlations of trails with the same input and output masks, we get closer to the real correlation. As a result, we get a linear approximation with a correlation $-2^{-47.76}$,which results in a correlation attack on SNOW-V and SNOW-Vi with a time complexity $2^{246.53}$, data complexity $2^{237.5}$ and memory complexity $2^{238.77}$.

Key words: SNOW-V; SNOW-Vi; Cryptanalysis, Linear Approximation; Automatic Search.

1 Introduction

SNOW-V[2] is a new member of SNOW family stream ciphers. SNOW-V has greatly expanded the internal state of the original structure of SNOW 3G, and was announced to satisfy the 256-bit security level requirement for 5G from 3GPP. SNOW-Vi[3] is another version of SNOW-V, which can be implemented faster while avoiding the linear relation between four taps of the LFSR part. The schematic of SNOW-V is depicted in Fig1. The LFSR part of SNOW-V is a circular structure consisting of two LFSRs, and the size of each register in FSM part increases to 128 bits. For SNOW-Vi, besides the field and update transformation of the LFSR, the tap is replaced as well. We refer to the design reports [2, 3] for more details.

One recent attempt to analyze these two stream ciphers focus on the linear approximation of SNOW-V and SNOW-Vi and presents a correlation attack applicable to both stream ciphers[5]. In [5], the linear approximation of SNOW-V based on a compose function technique is proposed, and the linear trails are searched out using the automatic search technique. With the binary approximation with correlation $2^{-49.54}$, a correlation attack was mounted with an expected time complexity $2^{248.81}$, memory complexity 2^{240} and 2^{240} keystream words of a

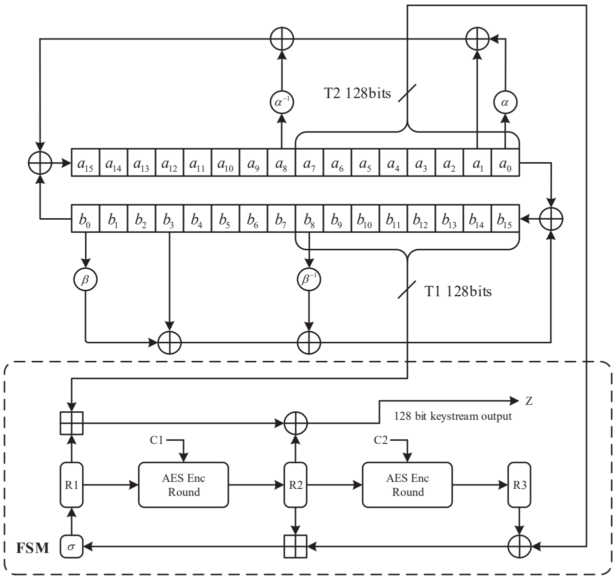
T2 128bits
α^{-1}
α
a_{15} a_{14} a_{13} a_{12} a_{11} a_{10} a_9 a_8 a_7 a_6 a_5 a_4 a_3 a_2 a_1 a_0
b_0 b_1 b_2 b_3 b_4 b_5 b_6 b_7 b_8 b_9 b_{10} b_{11} b_{12} b_{13} b_{14} b_{15}
β
β^{-1}
T1 128bits
Z
128 bit keystream output
C1
C2
R1
AES Enc Round
R2
AES Enc Round
R3
FSM
σ

Fig. 1. The keystream generation phase of the SNOW-V stream cipher

pair of key and IV obtained. In this paper, we adopt the accurate characterization of the Walsh spectral of S-boxes, and search out several linear approximation trails with higher correlation. Based on that, we investigate the common features of these linear approximation trails, and try to get closer to the real correlation of an approximation. For convenience, we use the same notations as in [5].

2 Optimization of the automatic search program

In [5], the correlations of S-boxes that are not 1 nor 0 are treated as 2^{-3} firstly, then feed the real correlation to the trail obtained and get a linear approximation trail, which results in a non optimal search. Combining the methods in [1] and [4], we can optimize the search program to make it more accurate. Given the input mask $x = (x_7, x_6, ..., x_0)$ and output mask $y = (y_7, y_6, ..., y_0)$, the corresponding correlation is denoted by $c(x, y)$. Since the absolute correlations but 1 of the S-box has 8 values, we split the linear correlation table into multiple Boolean functions like in [1]. Here we construct 8 Boolean functions:

$$f_k(x,y) = \begin{cases} 1, & if|c(x,y)| = 4k/256; \\ 0, & if|c(x,y)| \neq 4k/256. \end{cases} \quad k = 1, 2, ..., 8$$

and convert them into a series of logical conditions using $LogicFriday$. With the constraint

$$f_1(x,y)|f_2(x,y)|...|f_8(x,y) = x_0|x_1|...|x_7|y_0|y_1|...|y_7$$

added, we have the observation that $f_k(x, y) = 1$ if and only if $|c(x, y)| = 4k/256$. As STP solver does not support the floating-point data type, we replace the absolute correlation $|c(x, y)|$ with [4]

$$s = -\left\lfloor 10^t \log_2 |c(x, y)| \right\rfloor = \sum\nolimits_{k=1}^{8} \left\lfloor 10^t f_k(x, y) \log_2(256/4k) \right\rfloor,$$

and replace the absolute correlation of modular addition and objective functions by 10^t times as well, in which t is the precision parameter. Thus we get a much more accurate search program, which leads to several linear approximation trails with absolute correlations higher than $2^{-49.54}$, which is presented in [5]. In Appendix A there are two of the trails we have searched out, shown with the same symbols as in [5]. It is obvious that the two trails are also fit for SNOW-Vi, for both of them satisfy $d = 0$.

3 More accurate evaluation of linear approximations

As shown in [5], the approximation of SNOW-V can be split into consecutive approximations to 6 functions and the linear approximation trails have the expression

$$(\gamma, \beta, l, m, n, \gamma) \xrightarrow[d\mathbf{L}=(e\oplus l)||(f\oplus m),\rho_A(\gamma\leftarrow a,n\oplus d)]{f_1} (a, \beta, e, f, d, \gamma) \xrightarrow[\rho_A(\boldsymbol{\sigma}^T a\leftarrow b\oplus\beta,d\oplus h)]{f_2}$$
$$(d \oplus h, b, e, f, h, \gamma) \xrightarrow[\rho_E(d\oplus h\leftarrow\alpha)\rho_E(b\leftarrow c)]{f_3} (\alpha, c, e, f, h, \gamma) \xrightarrow[\rho_A(\alpha\leftarrow e,c)]{f_4} (\alpha, \alpha, f, h, \gamma)$$
$$\xrightarrow[\rho_E(\gamma\leftarrow q)]{f_5} (\alpha, \alpha, f, h, q) \xrightarrow[\rho_A(\beta\leftarrow f,q)]{f_6} (\alpha, \alpha, h, \beta),$$

The accurate correlation of a linear approximation of SNOW-V should be computed as

$$c(\alpha, \beta, \gamma, l, m, n, h) = \sum_{a,b,c,d,q} \rho(a, b, c, d, q)$$

Therefore, we investigate the characteristics of the masks (a, b, c, d, q). As mentioned in Section 2, all the trails we've searched out follow $d = 0$. In fact, by constructing the search program with adding the constraint $d \neq 0$, we make sure experimentally that when α and β are zeros but their 12th bytes, there is no linear trail with the masks satisfying

$$l = \alpha, \quad m = \beta, \quad h = \gamma = 0x81ec5a80, 0, 0, 0, n = 0x81ec5a00, 0, 0, 0, \quad d \neq 0$$

Thus, the accurate correlation becomes

$$c(\alpha, \beta, \gamma, l, m, n, h) = \sum_{a,b,c,q} \rho(a, b, c, 0, q)$$

Now we analyze the intermediate masks (a, b, c, q) with the input and output constrained above. The correlation of f_5 is $\rho_5 = \rho_E(\gamma \leftarrow q)$ and only the 12th byte of $\mathbf{P}^T\gamma$ is nonzero, where $\mathbf{P}$ is the matrix of the linear transformation of AES round function. So we can deduce that only the 12th byte of q is nonzero from the assumption $\rho(a, b, c, 0, q) \neq 0$, i.e. q has the form $(X, 0, 0, 0)$, where X is a nonzero byte. In a similar way, the correlation of f_4 is $\rho_A(\alpha \leftarrow e, c)$ with $d = 0$, which implies $e = l = \alpha$, according to the properties of linear approximation of addition modulo 2^{32} we know that c has the same form as q. The correlation of f_3 is $\rho_E(h \leftarrow \alpha)\rho_E(b \leftarrow c)$, which indicates that $\mathbf{P}^T b$ has the form $(X, 0, 0, 0)$ as well. The correlation of f_2 is $\rho_A(\sigma^T a \leftarrow b \oplus \beta, h)$, due to the fact that the most significant bit of h is 1, it shows that the 127th bits of both a and b are 1.

We exhaust the 12th bytes of $\mathbf{P}^T b$, q and c. For a, we have the observation experimentally that all the trails with high correlation we have searched out follows $a = (0xX000000, 0, 0, 0)$, so we exhaust the most significant byte of a only. We did this for every trail with high correlation we have found. The two best results are shown in Appendix B. In fact, they are obtained with the two trails in Appendix A by the operation above, and we can see the interesting aggregation effect: the correlation of the first approximation is lower than the second, while the first linear trail has a higher correlation.

By the method above, we get much closer to the real correlation. With the second approximation we can make an improvement of the correlation attack with $B = 238$, $M \approx 2^{200}$, $N \approx 2^{237.5}$ and $p_s = 0.999992$. The time complexity presented in [5] and memory can be reduced to $2^{246.53}$ and $2^{238.77}$ respectively, with $2^{237.5}$ words obtained.

4 Conclusion

In this paper we optimize the automatic search program and investigate the complete linear approximation of SNOW-V and SNOW-Vi under certain masks. Taking the accurate characterization into account, we can make a more precise search taking a longer time, and search out linear trails with higher correlations. Then we try to restore the complete binary approximations based on the linear trails with high correlations. By exhausting the most probable free masks, we believe it is closer to the actual value. Applying this method we find two correlations higher than those in [5], which can result in an improved correlation attack to SNOW-V and SNOW-Vi.

References

1. Abdelkhalek, A., Sasaki, Y., Todo, Y., Tolba, M., Youssef, A.M.: MILP modeling for (large) S-boxes to optimize probability of differential characteristics. IACR Transactions on Symmetric Cryptology pp. 99–129 (2017)
2. Ekdahl, P., Johansson, T., Maximov, A., Yang, J.: A new SNOW stream cipher called SNOW-V. IACR Transactions on Symmetric Cryptology pp. 1–42 (2019)
3. Ekdahl, P., Maximov, A., Johansson, T., Yang, J.: SNOW-Vi: an extreme performance variant of SNOW-V for lower grade cpus. In: Proceedings of the 14th ACM Conference on Security and Privacy in Wireless and Mobile Networks. pp. 261–272 (2021)

4. Liu, Y., Liang, H., Li, M., Huang, L., Hu, K., Yang, C., Wang, M.: STP models of optimal differential and linear trail for S-box based ciphers. Science China Information Sciences **64**(5), 1–3 (2021)
5. Shi, Z., Jin, C., Zhang, J., Cui, T., Ding, L.: A correlation attack on full SNOW-V and SNOW-Vi. Cryptology ePrint Archive, Report 2021/1047 (2021), https://ia.cr/2021/1047

Appendix A

Two of linear approximation trails with correlations higher than $2^{-49.54}$:
The 1st trail: correlation = 2^{-48}

$$\begin{aligned}
&\alpha = c = l = 0xc, 0, 0, 0\\
&\beta = m = 0x80, 0, 0, 0\\
&\gamma = b = h = 0x81ec5a80, 0, 0, 0\\
&a = 0xc1000000, 0, 0, 0\\
&q = 0xa0, 0, 0, 0\\
&n = 0x81ec5a00, 0, 0, 0\\
&d = 0
\end{aligned}$$

The 2nd trail: correlation $\approx -2^{-49.063}$

$$\begin{aligned}
&\alpha = c = l = 0xd, 0, 0, 0\\
&\beta = m = 0x40, 0, 0, 0\\
&\gamma = b = h = 0x81ec5a80, 0, 0, 0\\
&a = 0xc1000000, 0, 0, 0\\
&q = 0x60, 0, 0, 0\\
&n = 0x81ec5a00, 0, 0, 0\\
&d = 0
\end{aligned}$$

Appendix B

Two linear approximation found:
The 1st approximation: correlation $\approx 2^{-48.065}$

$$\begin{aligned}
&\alpha = l = 0xc, 0, 0, 0\\
&\beta = m = 0x80, 0, 0, 0\\
&\gamma = h = 0x81ec5a80, 0, 0, 0\\
&n = 0x81ec5a00, 0, 0, 0
\end{aligned}$$

The 2nd approximation: correlation $\approx -2^{-47.760}$

$$\begin{aligned}
&\alpha = l = 0xd, 0, 0, 0\\
&\beta = m = 0x40, 0, 0, 0\\
&\gamma = h = 0x81ec5a80, 0, 0, 0\\
&n = 0x81ec5a00, 0, 0, 0
\end{aligned}$$