A Correlation Attack on Full SNOW-V and SNOW-Vi

Zhen Shi, Chenhui Jin, Jiyan Zhang, Ting Cui, Lin Ding and Yu Jin

Information Engneering University, Zhengzhou, China

@Eurocrypt 2022

Contents

1. The SNOW-V and SNOW-Vi Stream Ciphers
2. Linear approximation of SNOW-V
3. Automatic search of linear approximation trails of SNOW-V
4. Evaluating a special type of binary linear approximations
5. A correlation attack on SNOW-V/SNOW-Vi

The SNOW-V and SNOW-Vi Stream Ciphers

- The stream cipher SNOW-V is a new member of the SNOW family. It was proposed for the 5G mobile communication system by Ekdahl et al. [1]
- To achieve the strong security requirements by the 3GPP standardization organization for the 5G system, it's need for SNOW-V to provide a 256-bit security level with a 256-bit key.
- Recently, a faster variant of SNOW-V, called SNOW-Vi, was proposed at ACM WiSec 2021. [2]

---

[1] [EJMY19] 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)

[2] [EMJY21] 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)

The SNOW-V and SNOW-Vi Stream Ciphers

- The SNOW-V stream cipher consists of a LFSR part and a FSM part.
- SNOW-Vi is exactly the same as SNOW-V, with some differences in the LFSR part and the tap choice of T2.

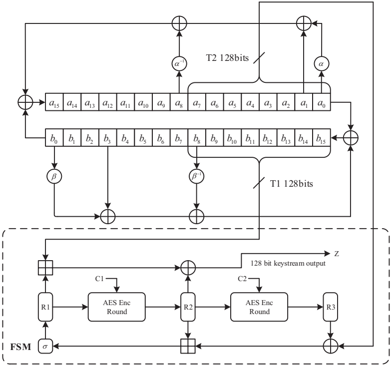

Fig. 1. An overall schematic of SNOW-V

Summary of the Attacks on SNOW-V and SNOW-Vi

Attack type	Version	Round	Time	Data
Distinguishing attack [YJM22]	SNOW-V$_{\oplus}$	full	2^{303}	-
Correlation attack [GZ21]	SNOW-V$_{\sigma 0}$	full	$2^{251.93}$	$2^{103.83}$
Correlation attack [GZ21]	SNOW-V$_{\boxplus_{32}, \boxplus_8}$	full	$2^{377.01}$	$2^{253.73}$
Guess and determine [CDM20]	SNOW-V	full	2^{512}	7
Guess and determine [JLH20]	SNOW-V	full	2^{406}	7
Guess and determine [YJM22]	SNOW-V	full	2^{378}	8
Differential attack [HII+21]	SNOW-V	4	$2^{153.97}$	$2^{26.96}$
Differential attack [HII+21]	SNOW-Vi	4	$2^{233.99}$	$2^{7.94}$

No results on full SNOW-V or SNOW-Vi are faster than exhaustive key search.

SNOW-V$_{\oplus}$, SNOW-V$_{\sigma 0}$ and SNOW-V$_{\boxplus_{32}, \boxplus_8}$ are reduced variants of SNOW-V

Linear approximation of SNOW-V

- Motivation: find biased binary approximations of SNOW-V which only relate to the output words and LFSR states.

$$(\alpha, \beta, \gamma, l, m, n, h) \cdot (z_{t-1}, z_t, z_{t+1}, T_1^{(t-1)}, T_1^{(t)}, T_1^{(t+1)}, T_2^{(t)}) \stackrel{\rho}{=} 0$$

- Our methods: convert the linear approximations based on three consecutive outputs of SNOW-V into those of a composite function equivalently.

A simple observation of the four taps at three consecutive clocks:

$$T_2^{(t)} = T_1^{(t+1)} \oplus \beta * {T_1}^{(t-1)} \oplus \beta^{-1} * {T_1}^{(t)} \oplus ({T_1}^{(t-1)} >> 48) \oplus ({T_1}^{(t)} << 80).$$

Rewritten as

$$L(T_1^{(t-1)}, T_1^{(t)}) = T_1^{(t+1)} \oplus T_2^{(t)},$$

where L is a linear mapping recording the linear relationship above.

Walsh spectrum

- Walsh spectrum of a function f:

$$W_{(f)}(\alpha \rightarrow \beta) = \frac{1}{2^n} \sum_{x \in F_2^n} (-1)^{\beta \cdot f(x) \oplus \alpha \cdot x}$$

- Walsh spectrum of composite function $g \circ f$:

$$W_{(g \circ f)}(\alpha \rightarrow \beta) = \sum_{\gamma \in F_2^n} W_{(f)}(\alpha \rightarrow \gamma) W_{(g)}(\gamma \rightarrow \beta)$$

Linear approximation of SNOW-V

The keystream outputs in three consecutive clocks can be expressed by

$$\begin{aligned} z_{t-1} &= (T_1^{(t-1)} \boxplus E^{-1}(R_2)) \oplus E^{-1}(R_3), \\ z_t &= (T_1^{(t)} \boxplus R_1) \oplus R_2, \\ z_{t+1} &= (T_1^{(t+1)} \boxplus \sigma(R_2 \boxplus (R_3 \oplus T_2^{(t)}))) \oplus E(R_1). \end{aligned}$$

Let $\alpha, \beta, \gamma, l, m, n, h$ be 128-bit masks.The following equation show a nonzero correlation ρ when the masks take certain values:

$$\begin{aligned} & (\alpha, \beta, \gamma, l, m, n, h) \cdot (z_{t-1}, z_t, z_{t+1}, T_1^{(t-1)}, T_1^{(t)}, T_1^{(t+1)}, T_2^{(t)}) \\ = \; & \alpha \cdot (E^{-1}(R_2) \boxplus T_1^{(t-1)}) \oplus \beta \cdot R_2 \oplus \gamma \cdot (\sigma(R_2 \boxplus (R_3 \oplus T_2^{(t)})) \boxplus T_1^{(t+1)}) \\ & \oplus \alpha \cdot E^{-1}(R_3) \oplus \beta \cdot (R_1 \boxplus T_1^{(t)}) \oplus \gamma \cdot E(R_1) \\ & \oplus l \cdot T_1^{(t-1)} \oplus m \cdot T_1^{(t)} \oplus n \cdot T_1^{(t+1)} \oplus h \cdot T_2^{(t)} \\ \overset{\rho}{=} \; & 0. \end{aligned}$$

Divide the linear approximation into 6 sub-functions

Let

$$F(x, y, z, u, v, w) := (f_6 \circ f_5 \circ f_4 \circ f_3 \circ f_2 \circ f_1)(x, y, z, u, v, w),$$

where

$$\begin{aligned}
f_1(x, y, z, u, v, w) &= (x \boxminus v, y, z, u, L(z, u) \oplus v, w), \\
f_2(x, y, z, u, v, w) &= ((\sigma^{-1}(x) \boxminus y) \oplus v, y, z, u, v, w), \\
f_3(x, y, z, u, v, w) &= (E^{-1}(x), E^{-1}(y), z, u, v, w), \\
f_4(x, y, z, u, v, w) &= (x, (y \boxplus z), u, v, w), \\
f_5(x, y, z, u, v) &= (x, y, z, u, E^{-1}(v)), \\
f_6(x, y, z, u, v) &= (x, y, u, (z \boxplus v)).
\end{aligned}$$

It is clear that the composite function F has 6-word input and 4-word output.

Divide the linear approximation into 6 sub-functions

Theorem 1. Assume that $R_1, R_2, R_3, T_1^{(t-1)}, T_1^{(t)}, T_1^{(t+1)}$ are independent and uniform distributed. For the binary linear approximation of F

$$(\gamma, \beta, l, m, n, \gamma) \xrightarrow[\rho_F]{F} (\alpha, \alpha, h, \beta),$$

we have $\rho = \rho_F$.

Theorem 1 indicates:

- we can convert the problem of computing the correlation into that of searching for linear approximations of F equivalently.
- by using the properties of Walsh spectrum of composite functions, we can evaluate the approximations of SNOW-V by measuring the linear trails directly.

Distinguisher for Distinguishing Attack

From the linear relation

$$L(T_1^{(t-1)}, T_1^{(t)}) = T_1^{(t+1)} \oplus T_2^{(t)},$$

we know that

$$l \cdot T_1^{(t-1)} \oplus m \cdot T_1^{(t)} \oplus n \cdot T_1^{(t+1)} \oplus h \cdot T_2^{(t)} = 0,$$

when $n\boldsymbol{L} = h\boldsymbol{L} = (l||m)$ holds, in which $||$ represents the cascading operation. Then Equation (1) shall become

$$\begin{aligned} & (\alpha, \beta, \gamma, l, m, n, h) \cdot (z_{t-1}, z_t, z_{t+1}, T_1^{(t-1)}, T_1^{(t)}, T_1^{(t+1)}, T_2^{(t)}) \\ = \quad & \alpha \cdot z_{t-1} \oplus \beta \cdot z_t \oplus \gamma \cdot z_{t+1} \stackrel{\rho}{=} 0, \end{aligned}$$

which contains only the output words z_{t-1}, z_t, z_{t+1}.

Distinguisher for Correlation Attack

When

$$l \cdot T_1^{(t-1)} \oplus m \cdot T_1^{(t)} \oplus n \cdot T_1^{(t+1)} \oplus h \cdot T_2^{(t)} \neq 0,$$

we will get a distinguisher for correlation attack which can be used to recover the initial state of the LFSR.

The linear approximation trail of F

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

Distinguisher for Correlation Attack

- The correlation can be evaluated as

$$\rho(a,b,c,d,q) = \rho_A(a, n \oplus d \to \gamma)\rho_A(b \oplus \beta, d \oplus h \to a\boldsymbol{\sigma})\rho_E(\alpha \to d \oplus h)$$
$$\rho_E(c \to b)\rho_A(e, c \to \alpha)\rho_E(q \to \gamma)\rho_A(f, q \to \beta),$$

with the constraint $d\boldsymbol{L} = (e \oplus l)||(f \oplus m)$.

- The accurate correlation of a binary linear approximation with the input and output masks defined by parameters $\alpha, \beta, \gamma, l, m, n, h$ of F should be computed by

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

which means exhausting the intermediate masks a, b, c, d, q will obtain the accurate correlation.

Automatic search of linear approximation trails of SNOW-V

STP-based automatic search model

The model of the linear approximation mainly contains three substitution layers and four layers of addition modulo 2^{32} operations as the nonlinear part.

Addition modulo 2^{32}. Denoting the output mask as u, the input masks as v, w, the i-th bit of Boolean vector x as x_i, the model to obtain a valid linear approximation is

$$\begin{aligned}
&z_{n-1} = 0, \\
&z_j = z_{j+1} \oplus u_{j+1} \oplus v_{j+1} \oplus w_{j+1} (0 \leq j < n-1), \\
&z_i \geq u_i \oplus v_i (0 \leq i < n), \\
&z_i \geq u_i \oplus w_i,
\end{aligned}$$

where z is a dummy variable.

Automatic search of linear approximation trails of SNOW-V

Addition modulo 2^{32} Denoting the output mask as u, the input masks as v, w, the i-th bit of Boolean vector x as x_i, the model to obtain a valid linear approximation is

$$z_{n-1} = 0; z_j = z_{j+1} \oplus u_{j+1} \oplus v_{j+1} \oplus w_{j+1} (0 \leq j < n-1);$$
$$z_i \geq u_i \oplus v_i (0 \leq i < n); z_i \geq u_i \oplus w_i;$$

where z is a dummy variable.

The correlation of the linear approximation is not zero if and only if z satisfies the model and $\rho_A(v, w \to u) = (-1)^{(u \oplus v)\cdot(u \oplus w)} 2^{-wt(z)}$.

Replace the absolute correlation of the j-th modular addition $2^{-wt(z^{(j)})}$ with $Z^{(j)} = 10^t wt(z^{(j)})$ to keep consistent with the accuracy of the correlation of S-boxes.

Automatic search of linear approximation trails of SNOW-V

8-bit S-box. Denote $c(x, y)$ as the correlation of an S-box with the input mask $x = (x_7, x_6, ..., x_0)$ and output mask $y = (y_7, y_6, ..., y_0)$. Since the nonzero absolute correlations of the S-box except 1 has 8 values, we split the linear correlation table into multiple Boolean functions. Here we construct 8 Boolean functions:

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

As the expressions longer than 256 characters are not supported by STP solver, $f(x, y)$ needs to be converted into a series of shorter constrains that are fully satisfied. Use *LogicFriday* to obtain the product-of-sum representation of a Boolean function.

Automatic search of linear approximation trails of SNOW-V

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 $|c(x,y)| = 4k/256$. STP solver does not support the floating-point data type, so we replace the absolute correlation of the i-th S-box $|c^{(i)}(x,y)|$

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

in which t is the precision parameter. Thus we get the absolute correlation of an S-box being accurate to t decimal places.

Automatic search of linear approximation trails of SNOW-V

Objective function

As there are 48 S-boxes and 16 modular additions taking part in the linear approximation, a trail can be evaluated by $\sum_{i=1}^{48} S^{(i)} + \sum_{j=1}^{16} Z^{(j)}$. A solution returned by the STP solver satisfying the constraint $\sum_{i=1}^{48} S^{(i)} + \sum_{j=1}^{16} Z^{(j)} < l$ represents a trail with the absolute correlation higher than $2^{-10^{-t}l}$.

The sign

After STP solver returns a linear approximation trail that satisfies all constraints, we verify the trail and determine its sign.

Automatic search of linear approximation trails of SNOW-V

Finding more trails

Assuming that the trail $(\alpha_0, \beta_0, \gamma_0, l_0, m_0, n_0, h_0, a_0, b_0, c_0, d_0, q_0)$ has been found, we can keep searching for other new solutions by introducing the additional constraints:

$$\alpha = \alpha_0, \beta = \beta_0, \gamma = \gamma_0, l = l_0, m = m_0, n = n_0, h = h_0,$$
$$(a \oplus a_0)|(b \oplus b_0)|(c \oplus c_0)|(d \oplus d_0)|(q \oplus q_0) \neq 0.$$

Different solutions can be generated one by one in this way, and the binary correlation gradually approaches its real value by summing up the correlations of linear trails.

Searching Results

The best result we have found is

$$\begin{aligned}
\alpha_1 &= l_1 = c_1 = 0xc, 0, 0, 0 \\
\beta_1 &= m_1 = 0x80, 0, 0, 0 \\
\gamma_1 &= h_1 = b_1 = 0x81ec5a80, 0, 0, 0 \\
n_1 &= 0x81ec5a00, 0, 0, 0 \\
a_1 &= 0xc1000000, 0, 0, 0 \\
q_1 &= 0xa0, 0, 0, 0 \\
d_1 &= 0, 0, 0, 0.
\end{aligned}$$

with the correlation 2^{-48} (The symbol '0' denotes 32-bit 0, and the leftmost 32-bit word is the most significant word.)

Searching Results

Another trail we will focus

$$\begin{aligned}
&\alpha_2 = l_2 = c_2 = 0xd, 0, 0, 0 \\
&\beta_2 = m_2 = 0x40, 0, 0, 0 \\
&\gamma_2 = h_2 = b_2 = 0x81ec5a80, 0, 0, 0 \\
&n_2 = 0x81ec5a00, 0, 0, 0 \\
&a_2 = 0xc1000000, 0, 0, 0 \\
&q_2 = 0x60, 0, 0, 0 \\
&d_2 = 0, 0, 0, 0.
\end{aligned}$$

with the correlation $-2^{-49.063}$.

Evaluating a special type of binary linear approximations

Observation

All the trails we have searched out with absolute correlation higher than 2^{-50} are of the form

$$\begin{aligned} \alpha &= l = 0x000000*, 0, 0, 0 \\ \beta &= m = 0x000000*, 0, 0, 0 \\ \gamma &= h = 0x81ec5a80, 0, 0, 0 \\ n &= 0x81ec5a00, 0, 0, 0. \end{aligned}$$

Thus, we can evaluate the accurate correlations of linear trails with

$$\alpha = l = 0x000000X, 0, 0, 0; \beta = m = 0x000000Y, 0, 0, 0;$$
$$\gamma = h = 0x81ec5a80, 0, 0, 0; n = 0x81ec5a00, 0, 0, 0,$$

by exhausting the intermediate masks a, b, c, d, q for given 8-bit words X and Y.

Evaluating a special type of binary linear approximations

Property of the linear approximation of modular addition

The most significant nonzero bit of an input mask must be in the same position as that of the output mask.

Mask $\boldsymbol{c}$ and $\boldsymbol{q}$

- c and e have the form $(0x000000*, 0, 0, 0)$, i.e., c and e are zeros except for their 12-th bytes, with the assumption $\alpha = (0x000000X, 0, 0, 0)$ and $\rho_A(c, e \rightarrow \alpha) \neq 0$.
- In a similar way, we can deduce $q = (0x000000*, 0, 0, 0)$ by $\beta = (0x000000Y, 0, 0, 0)$ and $\rho_A(f, q \rightarrow \beta) \neq 0$, and so as f.
- we only need to exhaust at most 255 values of c and q respectively.

Evaluating a special type of binary linear approximations

Mask $\boldsymbol{d}$

As l, m, e, f are zeros except for their 12-th bytes, we can get $d\boldsymbol{L} = (0x000000*, 0, 0, 0, 0x000000*, 0, 0, 0)$ from the linear relation $d\boldsymbol{L} = (e \oplus l)||(f \oplus m)$.

- $(0, 0, 0, 0)$ is the unique solution of d.

Mask $\boldsymbol{b}$

- By $\rho_E(c \to b) \neq 0$ and $c = (0x000000*, 0, 0, 0)$, we know that $b\boldsymbol{P} = (0x000000*, 0, 0, 0)$, where $\boldsymbol{P}$ is the binary matrix of the linear transformation of AES round function. So there are 255 values for b to traverse as well.

Evaluating a special type of binary linear approximations

Mask $\boldsymbol{a}$

- a is constrained by both $\rho_A(a, n \oplus d \to \gamma) \neq 0$ and $\rho_A(b \oplus \beta, d \oplus h \to \boldsymbol{\sigma}^T a) \neq 0$.
- The first constraint means that the least significant three 32-bit words of a are zeros.
- The second indicates that the least significant three 32-bit words of $\boldsymbol{\sigma}^T a$ are zeros.
- Since the 15-th byte is the unique fixed point of σ among the 4 most significant bytes, we have $a = (0x * 000000, 0, 0, 0)$.

Evaluating a special type of binary linear approximations

- We only need to exhaust 4 bytes to get all the trails with nonzero correlation and reach the accurate correlation of the linear approximation by summing them up when $\alpha, \beta, \gamma, l, m, n, h$ are chosen.
- We could also traverse X and Y to find the optimal approximation of this type.

Based on the two trails that have been searched, we compute the correlations and get

$$c(\alpha_1, \beta_1, \gamma_1, l_1, m_1, n_1, h_1) = 2^{-48.06};$$
$$c(\alpha_2, \beta_2, \gamma_2, l_2, m_2, n_2, h_2) = -2^{-47.76}.$$

A correlation attack on SNOW-V

Assume that $u = (u_{511}, u_{510}, \cdots, u_0)$ and $\hat{u} = (\hat{u}_{511}, \hat{u}_{510}, \cdots, \hat{u}_0)$ are the initial state and guessed initial state respectively.

For the output of LFSR at clock t, there exists a $\Gamma_t \in \{0,1\}^{512}$ such that

$$\Gamma_t \cdot u = l \cdot T_1^{(t-1)} \oplus m \cdot T_1^{(t)} \oplus n \cdot T_1^{(t+1)} \oplus h \cdot T_2^{(t)},$$

we can construct a distinguisher with the form

$$\begin{aligned}\phi_t(\hat{u}) =& \alpha \cdot z_{t-1} \oplus \beta \cdot z_t \oplus \gamma \cdot z_{t+1} \oplus \Gamma_t \cdot \hat{u} \\ =& \alpha \cdot z_{t-1} \oplus \beta \cdot z_t \oplus \gamma \cdot z_{t+1} \\ & \oplus l \cdot T_1^{(t-1)} \oplus m \cdot T_1^{(t)} \oplus n \cdot T_1^{(t+1)} \oplus h \cdot T_2^{(t)} \oplus \Gamma_t \cdot (u \oplus \hat{u}).\end{aligned}$$

$\phi_t(\hat{u})$ will show the correlation $c = -2^{-47.76}$ when $\hat{u} = u$, otherwise $\phi_t(\hat{u})$ is uniform distributed.

A correlation attack on SNOW-V

Preprocessing stage

Let the most significant B bits of the 512-bit binary vector x be x^h, the least significant $512 - B$ bits be x^l and the number of keystream words produced by a pair of key and IV be N.
For $1 \leq i_1, i_2 \leq N$, we have

$$(\Gamma_{i_1} \oplus \Gamma_{i_2}) \cdot u = (\Gamma^h_{i_1} \oplus \Gamma^h_{i_2}) \cdot u^h \oplus (\Gamma^l_{i_1} \oplus \Gamma^l_{i_2}) \cdot u^l.$$

If $\Gamma^l_{i_1} = \Gamma^l_{i_2}$, the equation above is $(\Gamma_{i_1} \oplus \Gamma_{i_2}) \cdot u = (\Gamma^h_{i_1} \oplus \Gamma^h_{i_2}) \cdot u^h$.

A correlation attack on SNOW-V

Preprocessing stage

A parity check equation of B bits of the initial state u is

$$\alpha\cdot(z_{i_1-1}\oplus z_{i_2-1})\oplus\beta\cdot(z_{i_1}\oplus z_{i_2})\oplus\gamma\cdot(z_{i_1+1}\oplus z_{i_2+1})\oplus(\Gamma^h_{i_1}\oplus\Gamma^h_{i_2})\cdot u^h \overset{c^2}{=} 0,$$

if $\Gamma^l_{i_1}=\Gamma^l_{i_2}$ holds.

Since the probability $p(\Gamma^l_{i_1}=\Gamma^l_{i_2})=2^{-(512-B)}$, the expected number of parity check equations with $\Gamma^l_{i_1}=\Gamma^l_{i_2}$ among C_N^2 pairs of Γ_i is $M=C_N^2 2^{-(512-B)}\approx 2^{-(513-B)}N^2$. Thus $2^{-(513-B)}N^2$ parity check equations can be constructed in preprocessing stage on average.

A correlation attack on SNOW-V

Processing stage

- Among the M parity check equations we denote the j-th equation $(\alpha, \beta, \gamma) \cdot Z_j \oplus \delta_j \cdot u^h = 0$, where $Z_j = (z_{i_1-1} \oplus z_{i_2-1}, z_{i_1} \oplus z_{i_2}, z_{i_1+1} \oplus z_{i_2+1})$ and $\delta_j = (\Gamma^h_{i_1} \oplus \Gamma^h_{i_2})$.
- For each guessed B bits $\hat{u}^h \in \{0,1\}^B$ of the initial state u, we evaluate the parity checks, then get $T(\hat{u}^h) = \sum_{j=1}^{M} (-1)^{(\alpha,\beta,\gamma)\cdot Z_j \oplus \delta_j \cdot \hat{u}^h}$, and predict the $\hat{u}$ that maximizes $T(\hat{u}^h)$ as the correct one.
- For the remaining $512 - B$ bits, the above process can be repeated when the first B bits are known.

A correlation attack on SNOW-V

- A correlation attack with an expected time complexity $2^{246.53}$, a memory complexity $2^{238.77}$, and $2^{237.5}$ keystream words, which can recover the internal state of SNOW-V at the clock producing the first keystream word.
- **To the best of our knowledge, this is the first attack on full SNOW-V with the time complexity less than exhaustive attack.**

A correlation attack on SNOW-Vi

The six functions are

$$\begin{aligned}
&f_1(x,y,z,t,u,v,w) = ((x \boxminus u), y, z, t, v, w),\\
&f_2(x,y,z,u,v,w) = ((\sigma^{-1}(x) \boxminus y) \oplus v, y, z, u, w),\\
&f_3(x,y,z,u,v) = (E^{-1}(x), E^{-1}(y), z, u, v),\\
&f_4(x,y,z,u,v) = (x, (y \boxplus z), u, v),\\
&f_5(x,y,z,u) = (x, y, z, E^{-1}(u)),\\
&f_6(x,y,z,u) = (x, y, (z \boxplus u)).
\end{aligned}$$

The composite function becomes

$$F(x,y,z,t,u,v,w) = (f_6 \circ f_5 \circ f_4 \circ f_3 \circ f_2 \circ f_1)(x,y,z,t,u,v,w),$$

with 7 input words and 3 output words.

A correlation attack on SNOW-Vi

The linear approximation is

$$(\gamma, \beta, l, m, n, h, \gamma) \xrightarrow[\rho_A(a,n\to\gamma)]{f_1} (a, \beta, l, m, h, \gamma) \xrightarrow[\rho_A(b\oplus\beta,h\to a\boldsymbol{\sigma})]{f_2} (h, b, l, m, \gamma)$$
$$\xrightarrow[\rho_E(\alpha\to h)\rho_E(c\to b)]{f_3} (\alpha, c, l, m, \gamma) \xrightarrow[\rho_A(l,c\to\alpha)]{f_4} (\alpha, \alpha, m, \gamma) \xrightarrow[\rho_E(q\to\gamma)]{f_5} (\alpha, \alpha, m, q)$$
$$\xrightarrow[\rho_A(m,q\to\beta)]{f_6} (\alpha, \alpha, \beta),$$

and the correlation of a linear trail can be computed as

$$\rho'(a, b, c, q) = \rho_A(a, n \to \gamma)\rho_A(b \oplus \beta, h \to a\boldsymbol{\sigma})\rho_E(\alpha \to h)\rho_E(c \to b) \rho_A(l, c \to \alpha)\rho_E(q \to \gamma)\rho_A(m, q \to \beta),$$

while the correlation of a linear trail of SNOW-V is

$$\rho(a, b, c, d, q) = \rho_A(a, n \oplus d \to \gamma)\rho_A(b \oplus \beta, d \oplus h \to a\boldsymbol{\sigma})\rho_E(\alpha \to d \oplus h) \rho_E(c \to b)\rho_A(e, c \to \alpha)\rho_E(q \to \gamma)\rho_A(f, q \to \beta),$$

A correlation attack on SNOW-Vi

The link between SNOW-V and SNOW-Vi

For any trail of the linear approximation process of SNOW-Vi in the proposed special type, the linear trail of SNOW-V determined by the same parameters

$$(\alpha, \beta, \gamma, l, m, n, h, a, b, c, q)$$

with $d = 0$ has the same correlation as that of SNOW-Vi, i.e.,

$$\rho(a, b, c, 0, q) = \rho'(a, b, c, q).$$

The linear approximation trails of SNOW-Vi correspond one-to-one to the trails with $d = 0$ of SNOW-V, and the set consisting of all linear trails of SNOW-Vi is a subset of that of SNOW-V.

A correlation attack on SNOW-Vi

We could approximate SNOW-Vi with the same correlation $-2^{-47.76}$ of SNOW-V under

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

- Similarly, the correlation attack with time complexity $2^{246.53}$, memory complexity $2^{238.77}$ and $2^{237.5}$ keystream words is effective for SNOW-Vi as well.
- **This is also the first attack better than exhaustive key search on full SNOW-Vi.**

Conclusion

- Propose a method to convert the linear approximation of the FSM part of SNOW-V into that of a composite function equivalently.
- Based on this method, we present a full coverage automatic search of linear trails of SNOW-V, and find a binary distinguisher with the correlation $-2^{-47.76}$.
- Using the approximation we mount the first correlation attack on full SNOW-V with the time complexity less than exhaustive key search. For SNOW-Vi, we prove the correlation attack is effective as well.

Thanks for Your Attention!