Evaluating pseudorandomness and superpseudorandomness of the iterative scheme to build SPN block cipher

Bui Cuong Nguyen, Tuan Anh Nguyen

***Abstract*— In this paper, the iterative scheme, namely the $\mathcal{V}$-scheme, is proposed constructing block ciphers. Then, the pseudorandomness and superpseudorandomness of this scheme are evaluated by using the Patarin's H-coefficient technique. In particular, the pseudorandomness of $\mathcal{V}$-scheme is achieved in the case that the number of round is at least 3, and $\mathcal{V}$-scheme is superpseudorandomness in the case that the number of round is greater than or equal 5. However, we have not yet evaluated superpseudorandomness of this scheme when the round is 4.**

***Tóm tắt*— Trong bài báo này, chúng tôi đưa ra lược đồ lặp gọi là lược đồ $\mathcal{V}$ dùng để xây dựng mã khối. Sau đó, đưa ra các kết quả đánh giá tính giả ngẫu nhiên và siêu giả ngẫu nhiên của lược đồ này được đưa ra dựa trên kỹ thuật hệ số H của Patarin. Trong đó, tính giả ngẫu nhiên của lược đồ đạt được khi số vòng của lược đồ là lớn hơn hoặc bằng 3. Đối với tính siêu giả ngẫu nhiên, chúng tôi đã chứng minh lược đồ đạt được khi số vòng lớn hơn hoặc bằng 5; còn khi số vòng bằng 4 chúng tôi chưa giải quyết được trong bài báo này.**

***Keywords*: block cipher structure, pseudorandomness; superpseudorandomness; H-coefficient technique.**

***Từ khóa*: cấu trúc mã khối, giả ngẫu nhiên; siêu giả ngẫu nhiên; kỹ thuật hệ số H.**

I. INTRODUCTION

In order to construct a secure block cipher, the scheme of block cipher structure plays an important role. Cryptographic designers usually choose a scheme based on structures such as SPN, Feistel, ARX,… and evaluate security of these scheme by their pseudorandomness and superpseudorandomness [1-5] which are described in [6]. The pseudorandomness and superpseudorandomness of schemes will ensure that an attacker which have unbounded (but finite) computation capabilities, can not distinguish the scheme from a perfect random function (permutation) with a non-negligible probability. In this model, a block cipher is considered as a random function (or a random permutation) associated with a randomly selected key. In [8], Henri Gilbert and Marine Minier stated that the strongest security requirement one can put on a f random function or permutation representing a key dependent cryptographic function is that f be undistinguishable with a non-negligible success probability from a perfect random function f^* or permutation c^*, even if a probabilistic testing algorithm $\mathcal{A}$ of unlimited power is used for that purpose.

Related results. The pseudorandomness and superpseudorandomness of a block cipher structure have been attracting research attention in the cryptography community. In 1988, Luby and Rackoff proposed the formal definitions of pseudorandomness and superpseudorandomness of block ciphers in [6]. In addition, they demonstrated that the 3-round Feistel structure is pseudorandomness and 4-round Feistel structure is superpseudorandomness. Patarin presented the H-coefficient technique and used it to prove these two results (see [7]). In [8], Gilbert and Minier used a simpler but rather effective approach based on Patarin's two main theorems to evaluate the pseudorandomness and superpseudorandomness for L and R schemes. In addition, at the SAC conference in 2009, Patarin systematized his theorems and formally introduced the H-coefficient technique to evaluate the secure of some block cipher schemes (see [7]). Hence, the H-coefficient technique is indeed an effective method for evaluating the secure of some encryption schemes and it is improved continuously (see [9]). For the SPN structure, the results of pseudorandomness and superpseudorandomness are actually attracting research attention in the world [10, 11].

This manuscript is received on January 10, 2018. It is commented on January 30, 2018 and is accepted on March 13, 2018 by the first reviewer. It is commented on February 1, 2018 and is accepted on March 8, 2017 by the second reviewer.

However, the approach of these results are based on the assumption that S-boxes are random permutation and diffusion layer is not specific in the evaluation model which makes it is difficult to evaluate.

Our contribution. In this paper, we considered the $\mathcal{V}$ scheme for constructing a SPN block cipher where the pseudorandomness and superpseudorandomness are evaluated in detail based on the H-coefficient technique. Specifically, the pseudorandom distinguishers with a non-negligible probability for 1-round and 2-round of scheme are given. Then, the theoretical result represented that 3-round $\mathcal{V}$-scheme is pseudorandomness. Finally, the superpseudorandomness of $\mathcal{V}$-scheme is considered.

Outline. This paper organized as follows: Section 2 represents some notations, security models and methods using Patarin's H-coefficient technique. Section 3 describes the iterative scheme considered in this paper. Section 4 and 5 respectively show the evaluation results of the pseudorandomness and superpseudorandomness of our scheme. Finally, some conclusions and an open problem are given.

II. PRELIMINARIES

A. Notations

Through this paper we are using the following notation: I_n denotes the $\mathbb{Z}_2^n$, $F_{n,m}$ denotes the set of functions from I_n into I_m, F_n denotes the set of functions from I_n into I_n, P_n denotes the set of permutations on I_n: thus $|F_{n,m}| = 2^{m \cdot 2^n}$.

B. The security model

First, we represent the definition of a pseudorandom distinguisher as follows:

Definition 1 ([12]). *Let $n, m > 1$. A pseudorandom distinguisher is a deterministic algorithm $\mathcal{A}$ with unbounded (but finite) computation capabilities, which given a function $F: I_n \to I_m$ can query it by asking values $x \in I_n$ of which it obtains the image $y = F(x)$. Depending on the answers $y \in I_m$ it obtains, $\mathcal{A}$ output either 0 or 1.*

A random function of $F_{n,m}$ is defined as a random variable f of $F_{n,m}$ and can be view as a probabilitiy distribution $(\Pr[f = \phi])_{\phi \in F_{n,m}}$ over $F_{n,m}$. A random function (a random permutation, respective) is a function (permutation) which is randomly chosen from $F_{n,m}(P_n)$ with a fixed probability. Thus, we have the definition of a perfect random function (perfect random permutation) as follows:

Definition 2 ([8]). *We define a perfect random function f^* of $F_{n,m}$ as a uniformly drawn element of $F_{n,m}$. In other words, f^* is associated with the uniform probability distribution over $F_{n,m}$. We define a c^* perfect random permutation on I_n as a uniformly drawn element of P_n. In other words, c^* is associated with the uniform probability distribution over P_n.*

Next, we define the advantage of a distinguisher $\mathcal{A}$ in distinguishing a random function F from a perfect random function F^*:

Definition 3 ([12]). *Let F be a random function, F^* be a perfect random function. The advantage a pseudorandom distinguisher $\mathcal{A}$ has in distinguishing F from F^* is:*

$$\mathrm{Adv}_{\mathcal{A}} := \left|\Pr[\mathcal{A}^F = 1] - \Pr\left[\mathcal{A}^{F^*} = 1\right]\right| \quad (1)$$

Pseudorandom distinguishers as defined above are allowed to make encryption queries only. Superpseurandom distinguishers are allowed to make decryption queries:

Definition 4 ([12]). *Let $N > 1$. A superpseudorandom distinguisher is a deterministic algorithm $\mathcal{A}$ with unbounded (but finite) computation capabilities, which can query a given permutation $C \in P_N$ by providing it with values $x \in I_N$ of which it obtains to its choosing either the image $y = C(x)$, or the inverse image $y = C^{-1}(x)$. Depending on the answers $y \in I_N$ it obtains, $\mathcal{A}$ outputs either 0 or 1.*

The advantage of a superpseudorandom distinguisher in distinguishing a random permutation C from a perfect random permutation C^* is defined similary to the case of pseudorandom distinguishers. In this paper, the random functions we want to distinguish from the perfect random ones are built by embedding perfect random functions $f_1^*, \ldots, f_t^*$ into a structure ϕ. The domain and range of $f_1^*, \ldots, f_t^*$ have variable size; it is smaller than the size of the domain and range of $\phi(f_1^*, \ldots, f_t^*)$. The such structure ϕ is sometimes called function (or permutation) generator. A function generator ϕ is said pseudorandom if for all pseudorandom distinguishers A of which the number of queries q is polynomial in N (block size), the advantage remains negligible (for N big enough). More formally:

Definition 5 ([12]). *A function generator ϕ is pseudorandom if for all polynomials $P(N), Q(N)$, there is an integer $N_0 \in N$ such that: $\forall N \geq N_0$, for all pseudorandom distinguishers $\mathcal{A}$ allowed to make $q \leq Q(N)$ queries,*

$$\mathrm{Adv}_{\mathcal{A}}(\phi(f_1^*, \dots, f_t^*), F^*) \leq \frac{1}{P(N)}.$$

Superpseudorandom permutations generators are defined similarly with respect to superpseudorandom distinguishers.

C. H-coefficient technique

In this section, we represent two Patarin's main theorem which were used to prove pseudorandomness and superpseudorandomness of structures based on the Luby-Rackoff model. This is very useful method to receive the advantage of a distinguisher has in distinguishing a random function (permutation) from a perfect random function (permutation).

K denotes the set of all t-tuples $(f_1, \dots, f_t)$ with $f_i \in P_n, 1 \leq i \leq t$. Let $G: K \to P_N$ be a permutation generator, here we have $N = 2n$.

Definition 6 ([9]**)**. *Let q be an integer (q is number of queries). Let $X = (X_i)_{1 \leq i \leq q}$ be a sequence of pairwise distinct elements of I_N. Let $Y = (Y_i)_{1 \leq i \leq q}$ be a sequence of elements of I_N. We denote by $H(X, Y)$ or simply by H if the context of the X_i, Y_i is clear, the number of $(f_1, \dots, f_t) \in K$ such that:*

$$\forall i, 1 \leq i \leq q, G(f_1, \dots, f_t)(X_i) = Y_i.$$

We denote $\mathcal{X}$ be a subset of I_N^q obtain all q-tuples $X = (X_1, \dots, X_q), X_i \in I_N, \forall i \neq j: X_i \neq X_j$.

Next, we consider the advantage of the pseudorandom distinguisher, allowed to make encryption queries only, the superpseudorandom distinguisher which allowed to make both encryption and decryption queries. These advantage were mention in [9] by Patarin (Theorem 3.4, Theorem 3.5). However, in order to evaluate our scheme, we represent two variants of these above theorems as follows:

Theorem 1 ([9]**)**. *Let α and β be real numbers, $\alpha, \beta > 0$. Let E be a subset of I_N^q such that $|E| \geq 2^{Nq} \cdot (1 - \beta)$. If:*

(1) For all $X \in \mathcal{X}$ and for all $Y \in E$ we have:

$$H(X, Y) \geq \frac{|K|}{2^{Nq}}(1 - \alpha);$$

Then

(2) For every pseudorandom distinguishers $\mathcal{A}$ allowed to make q encryption queries, we have:

$$Adv_{\mathcal{A}}(G(f_1, \dots, f_t), f^*) \leq \alpha + \beta$$

where $Adv_{\mathcal{A}}(G(f_1, \dots, f_t), f^*)$ denotes the advantage to distinguish $G(f_1, \dots, f_t)$ ($(f_1, \dots, f_t)$ is uniformly chosen from K) from a perfect random function $f^* \in F_N$.

Theorem 2 (*[9]***)**. *Let $\epsilon > 0$ be a real number. If:*

(1) For all $X \in \mathcal{X}$ and for all $Y \in \mathcal{X}$ we have:

$$H(X, Y) \geq \frac{|K|}{2^{Nq}}(1 - \epsilon)$$

Then

(2) For every superpseudorandom distinguishers $\mathcal{A}$ allowed to make q encryption and decryption queries we have:

$$Adv_{\mathcal{A}}(G(c_1, \dots, c_t), c^*) \leq \epsilon + \frac{q(q-1)}{2 \cdot 2^N}$$

where $Adv_{\mathcal{A}}(G(c_1, \dots, c_t), c^*)$ denotes the advantage to distinguish $G(c_1, \dots, c_t)$ ($(c_1, \dots, c_t)$ is uniformly chosen from K) from perfect random permutation $c^* \in P_N$.

III. THE DESCRIPTION OF THE SCHEME

In this section, we propose an iterative scheme, called $\mathcal{V}$-scheme, which used to construct a $2n$-bit permutation from n-bit permutations. The 1-round $\mathcal{V}$-scheme is described as follows:

$$\phi(c_1^*, c_0^*)(\langle a, b \rangle) = \langle c_0^*(b), c_1^*(a) \oplus c_0^*(b) \rangle.$$

Then, r-round of this scheme is the composition of r function 1-round. Thus, the $2r$ n-bit permutations $c_0^*, \dots, c_{2r-1}^*$ make a $2n$-bit permutation as follows:

$$\phi(c_0^*, \dots, c_{2r-1}^*) = \phi(c_{2r-1}^*, c_{2r-2}^*) \circ \dots \circ \phi(c_1^*, c_0^*).$$

We can use this scheme to build a SPN block-cipher by choosing specific cryptographic elements. For example, the permutations $c_{2r-1}^*, c_{2r-2}^*, \dots, c_0^*$ are expressed by the combination of the XOR key addition, the substitution transformation, the linear transformation on two semi-blocks (n-bit) as decribed in Fig 2 (the dashed parts). Then, we have a SPN block cipher with the round function transformed a $2n$-bit block $X = (X_1, X_0)$ to a $2n$-bit block $Y = (Y_1, Y_0)$ as follows:

- The XOR key addition uses round $2n$-bit key $K = (K_1, K_0)$ where K_i is of n-bit length.
- The nonlinear layer S contains $2k$ w-bit S-boxes (such that $kw = n$) can simply be represented by a transformation $S(X) = \left(S_{2k-1}(x_{2k-1}), \dots, S_0(x_0)\right)$ with $X = x_{2k-1} \| x_{2k-2} \| \dots \| x_1 \| x_0$ where x_i is the w-bit word.
- The linear layer P performs linear transformation through the linear transformation n-bit P_0, P_1 as follows $P(X) = \left(P_0(X_0), P_0(X_0) \oplus P_1(X_1)\right)$.

In conclusion, the output of the round function will be obtained by the substitution and permutation transformation as follows $Y = P\left(S(X \oplus K)\right)$.

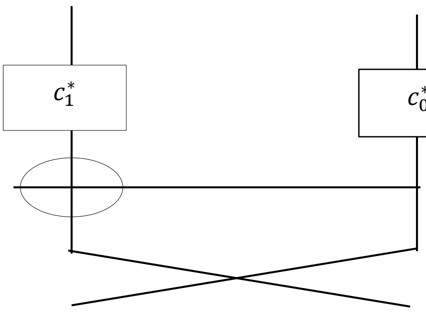


Fig 1. One-round $\mathcal{V}$-scheme

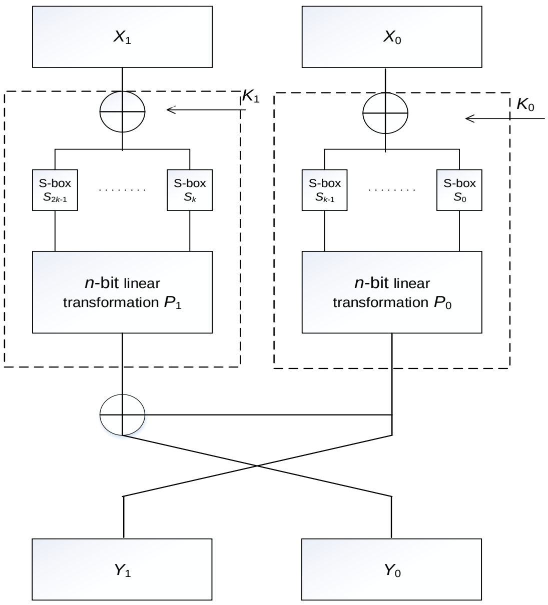


Fig 2. Round function of SPN block cipher with $2n$-bit block size, built from the proposed scheme

In the following sections, we will evaluate the pseudorandomness and superpseudorandomness of the $\mathcal{V}$-scheme.

IV. THE PSEUDORANDOMNESS OF THE $\mathcal{V}$-SCHEME

Fact 1. 1-round and 2-round $\mathcal{V}$-scheme are not pseudorandom.

Proof. *1-round.* Let $\mathcal{A}_1$ be a distinguisher, it operates as follows:

1. $\mathcal{A}_1$ chooses two values $X_1 = (a, b), X_2 = (a', b) \in I_{2n}$.
2. $\mathcal{A}_1$ queries into any f over F_{2n} to obtain $Y_1 = f(X_1) = (c, d)$ and $Y_2 = f(X_2) = (c', d')$.
3. $\mathcal{A}_1$ checks either $c = c'$ or not.
4. If $c = c'$ then $\mathcal{A}_1$ return 1 else returns 0.

Let p_1^* be the probability which $\mathcal{A}_1$ returns 1 when f is a perfect random function. Thus, $p_1^* = 2^{-n}$. Let p_1 be the probability which $\mathcal{A}_1$ returns 1 when $f = \phi(c_0^*, c_1^*)$ (1-round $\mathcal{V}$ scheme). Thus, $p_1 = 1$ because $c = c_0^*(b) = c'$. So, we have the advantage the pseudorandom distinguisher $\mathcal{A}_1$ is $\mathrm{Adv}_{\mathcal{A}_1}(f, f^*) = |p_1 - p_1^*| = 1 - 2^{-n}$. Thus, 1-round $\mathcal{V}$ scheme is not pseudorandom.

2-round. Let $\mathcal{A}_2$ be a distinguisher, it operates as follows:

1. $\mathcal{A}$ chooses two values $X_1 = (a, b), X_2 = (a', b) \in I_{2n}$.
2. $\mathcal{A}_2$ queries into any f over F_{2n} to obtain $Y_1 = f(X_1) = (c, d)$ and $Y_2 = f(X_2) = (c', d')$.
3. $\mathcal{A}_2$ checks either $c \oplus d = c' \oplus d'$ or not.
4. If $c \oplus d = c' \oplus d'$ then $\mathcal{A}_2$ return 1 else returns 0.

Let p_1^* be the probability which $\mathcal{A}_2$ returns 1 when f is a perfect random function. Thus, $p_1^* = 2^{-n}$. Let p_1 be the probability which $\mathcal{A}_2$ returns 1 when $f = \phi(c_0^*, c_1^*, c_2^*, c_3^*)$ (2-round $\mathcal{V}$ scheme). We have $p_1 = 1$ because of $c \oplus d = c_3^*\left(c_0^*(b)\right) = c' \oplus d'$. So the advantage the pseudorandom distinguisher $\mathcal{A}_2$ is $\mathrm{Adv}_{\mathcal{A}_2}(f, f^*) = |p_1 - p_1^*| = 1 - 2^{-n}$. Thus, 2-round $\mathcal{V}$ scheme is not pseudorandom □

When the number of round of $\mathcal{V}$-scheme is greater two, using H-coefficient technique we have the following result:

Proposition 1. Let $n > 0$ be an integer. Let $c_0^*, \dots, c_{2r-1}^* \in P_n$ are $2r$ $(r \geq 3)$ perfect random permutations and $f^* \in F_{2n}$ is a perfect random function. Let $f = \phi(c_0^*, \dots, c_{2r-1}^*)$ denotes the random permutation associated with the r-round $\mathcal{V}$-scheme. For any pseudorandom distinguisher $\mathcal{A}$ allowed to make q encryption queries, we have:

$$\mathrm{Adv}_{\mathcal{A}}(f, f^*) \leq r \cdot \frac{q(q-1)}{2^n}.$$

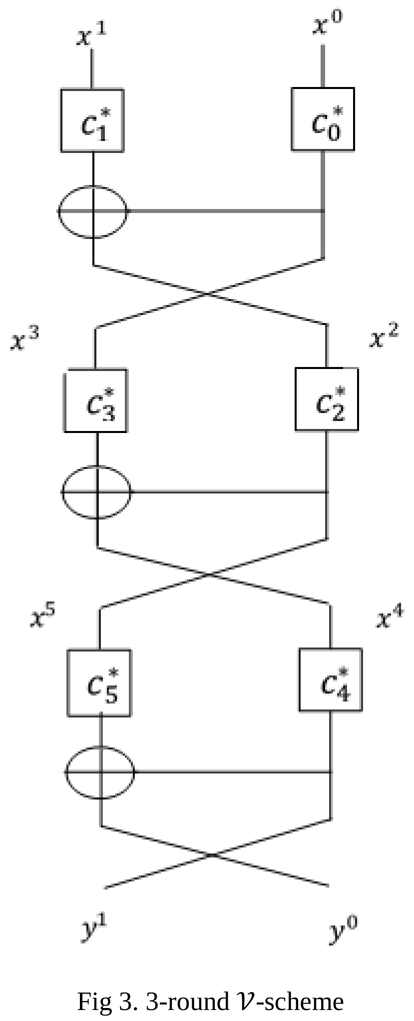


Fig 3. 3-round $\mathcal{V}$-scheme

Proof. In order to prove this proposition, we need some notations. $I^{\neq}$ denotes the subset of I_n^q consisting of all the q-tuples of pairwise distinct I^n. For $x = (x_1, \dots, x_q), y = (y_1, \dots, y_q) \in I_n^q$ we denote $x \sim y$ means that $\forall i, j, x_i = x_j$ if and only if $y_i = y_j$. Let $\mathcal{X} = \{X = (X_1, \dots, X_q), X_i = (x_i^1, x_i^0) \in I_{2n}, \forall i \neq j, X_i \neq X_j\}$, $x^t = (x_i^t)_{i=1..q} \in I_n^q$ and $y^t = (y_i^t)_{i=1..q} \in I_n^q$. Let (x^{2t+1}, x^{2t}) are intermediate variables at round $t \leq r$.

This proposition will be proven by using Theorem 1. It means that, we will construct a set E and find numbers α and β.

Firstly, we consider the set E:

$$E = \{Y = (Y_1, \dots, Y_q), Y_i = (y_i^1, y_i^0), y^1 \in I^{\neq}, y^1 \oplus y^0 \in I^{\neq}\}.$$

Secondly, we establish a lower bound on $|E|$ to find β. We have

$$\begin{aligned}
|E| &= |I_{2n}|^q \cdot (1 - \Pr[(y^1 \notin I^{\neq}) \vee (y^0 \oplus y^1 \notin I^{\neq})]) \\
&\geq |I_{2n}|^q \cdot \left(1 - \sum_{1 \leq i < j \leq q} \Pr[y_i^1 = y_j^1] - \sum_{1 \leq i < j \leq q} \Pr[y_i^0 \oplus y_i^1 = y_j^0 \oplus y_j^1]\right) \\
&= |I_{2n}|^q \left(1 - \sum_{1 \leq i < j \leq q} \sum_{s \in I_n} (\Pr[y_i^1 = s] \cdot \Pr[y_j^1 = s]) - \sum_{1 \leq i < j \leq q} \sum_{t \in I_n} (\Pr[y_i^0 \oplus y_i^1 = t] \cdot \Pr[y_j^0 \oplus y_j^1 = t])\right) \\
&\geq |I_{2n}|^q \cdot \left(1 - 2 \cdot \frac{q(q-1)}{2} \cdot 2^{-n}\right).
\end{aligned}$$

We can take $\beta = \frac{q(q-1)}{2^n}$.

Thirdly, in order to find β, we will establish a lower bound on the number of permutation $f = (c_0^*, \dots, c_{2r-1}^*)$ such that $f(X) = Y$ for all $X \in \mathcal{X}, Y \in E$. Now, we evaluate for three-round case, then we will generalize for the case $r > 3$. This mean that we find a lower bound on the number of permutations $(c_0^*, c_1^*, c_2^*, c_3^*, c_4^*, c_5^*)$ such that $f(X) = Y, \forall X \in \mathcal{X}, \forall Y \in E$ or:

$$\forall i, 1 \leq i \leq q,$$

$$\begin{cases} y_i^1 = c_4^*\left(c_3^*\left(c_0^*(x_i^0)\right) \oplus c_2^*\left(c_1^*(x_i^1) \oplus c_0^*(x_i^0)\right)\right) \\ y_i^1 \oplus y_i^0 = c_5^*\left(c_2^*\left(c_1^*(x_i^1) \oplus c_0^*(x_i^0)\right)\right) \end{cases}$$

The number of permutations c_0^* such that $c_0^*(x^0) = x^3$ with $x^3 \sim x^0$ is $|P_n|$. We have $y^0 \oplus y^1 \in I^{\neq} \Rightarrow x^5 \in I^{\neq} \Rightarrow x^2 \in I^{\neq}$ so permutation c_1^* must satisfy $c_1^*(x^1) \oplus c_0^*(x^0) \in I^{\neq}$. In order to establish it we first evaluate the number of permutations c_1^* such that:

$$c_1^*(x_i^1) \oplus c_0^*(x_i^0) = c_1^*(x_j^1) \oplus c_0^*(x_j^0) \qquad (1)$$

with $1 \leq i < j \leq q$.

- If $x_i^0 = x_j^0$ then there are no permutation realizes (1).
- If $x_i^0 \neq x_j^0, x_i^1 = x_j^1$ then there are no permutation realizes (1).
- If $x_i^0 \neq x_j^0, x_i^1 \neq x_j^1$ then $c_1^*(x_j^1)$ is determined by the value of $c_1^*(x_i^1)$. So the number of permutations c_1^* satisfy (1) is $\frac{|P_n|}{2^n - 1} \leq \frac{|P_n|}{2^{n-1}}$.

This mean that there are at most $\frac{q(q-1)|P_n|}{2^n}$ permutations c_1^* such that $\exists (i,j), 1 \leq i < j \leq q$ such that (1). (Because there are $\frac{q(q-1)}{2}$ pairs (i,j) such that $1 \leq i < j \leq q$). Thus, we have at least $|P_n|\left(1 - \frac{q(q-1)}{2^n}\right)$ permutations c_1^* satisfy $c_1^*(x^1) \oplus c_0^*(x^0) \in I^{\neq}$. The number of permutations c_2^* such that $c_2^*(x^2) = x^5$ for some $x^5 \in I^{\neq}$ is $|P_n|$. We have $x^4 \in I^{\neq}$ since $y^1 \in I^{\neq}$, so permutation c_3^* must satisfy

$$c_3^*\left(c_0^*(x^0)\right) \oplus c_2^*\left(c_1^*(x^1) \oplus c_0^*(x^0)\right) \in I^{\neq} \qquad (2)$$

By the similar way above, there are at least $|P_n|\left(1 - \frac{q(q-1)}{2^n}\right)$ permutations c_3^* satisfy (2). It is easy to see that the number of permutations c_4^* such that $c_4^*(x^4) = y^1$ with $y^1 \in I^{\neq}$ is $|P_n| \cdot \frac{(2^n - q)!}{2^n!}$. Similarly, there are $|P_n| \cdot \frac{(2^n - q)!}{2^n!}$ permutations c_5^* such that $c_5^*(x^5) = y^1 \oplus y^0$ with $y^1 \oplus y^0 \in I^{\neq}$.

From the above arguments, we have:

$$\begin{aligned} H &\geq |P_n| \cdot |P_n|\left(1 - \frac{q(q-1)}{2^n}\right) \cdot |P_n| \\ &\quad \cdot |P_n|\left(1 - \frac{q(q-1)}{2^n}\right) \cdot |P_n| \\ &\quad \cdot \frac{(2^n - q)!}{2^n!} \cdot |P_n| \cdot \frac{(2^n - q)!}{2^n!} \\ &= |P_n|^6 \cdot \left(1 - \frac{q(q-1)}{2^n}\right)^2 \cdot \left(\frac{(2^n - q)!}{2^n!}\right)^2 \\ &\geq |P_n|^6 \cdot \left(1 - \frac{2q(q-1)}{2^n}\right) \cdot \frac{1}{2^{2nq}}. \end{aligned}$$

Thus, $\alpha = 2 \cdot \frac{q(q-1)}{2^n}$.

Next, we will find α in the case $r > 3$. So we establish the number of permutation $f = (c_0^*, \dots, c_{2r-1}^*)$ such that $f(X) = Y$ for all $X \in \mathcal{X}, Y \in E$. We evaluate by the following way: we will establish the number of permutations c_{2t-1}^*, c_{2t-2}^* such that $x^{2t+1}, x^{2t} \in I^{\neq}$ with $2 \leq t \leq r - 1$. We can assume that $x^{2t-2} \in I^{\neq}$ (because after the first round we can choose permutation c_1^* such that $x^2 \in I^{\neq}$ as the 3 round case), since $x^{2t+1} = c_{2t-2}^*(x^{2t-2})$, so we have the number of permutations c_{2t-2}^* is $|P_n|$. The permutation c_{2t-1}^* must satisfy $c_{2t-1}^*(x^{2t-1}) \oplus x^{2t+1} \in I^{\neq}$ because of $x^{2t} = c_{2t-1}^*(x^{2t-1}) \oplus x^{2t+1}$. By the similar method in the 3-round case, we have the number of permutations c_{2t-1}^* is $|P_n|\left(1 - \frac{q(q-1)}{2^n}\right)$. We now only need evaluate the number of permutations in the first and last round. Luckly, it is like the 3-round case that we have done. We have the number of permutations $c_0^*, c_{2r-1}^*, c_{2r-2}^*$ are $|P_n|, |P_n| \cdot \frac{(2^n - q)!}{2^n!}, |P_n| \cdot \frac{(2^n - q)!}{2^n!}$ respectively; for c_1^* we have at least $|P_n|\left(1 - \frac{q(q-1)}{2^n}\right)$. Then, we have:

$$\begin{aligned} H &\geq |P_n|^{2r} \cdot \left(\frac{(2^n - q)!}{2^n!}\right)^2 \cdot \left(1 - \frac{q(q-1)}{2^n}\right)^{r-1} \\ &\geq |P_n|^{2r} \cdot \frac{1}{2^{2nq}} \\ &\quad \cdot \left(1 - (r-1) \cdot \frac{q(q-1)}{2^n}\right). \end{aligned}$$

Thus $\alpha = (r-1) \cdot \frac{q(q-1)}{2^n}$.

Appling Theorem 1 with $\beta = \frac{q(q-1)}{2^n}$ and $\alpha = (r-1) \cdot \frac{q(q-1)}{2^n}$ we have the Proposition 1 □

The pseudorandomness of this scheme in Proposition 1 is still achieved when the perfect random permutations $2r$-tuples $(c_0^*, \cdots, c_{2r-1}^*)$ are replaced by the perfect random functions $2r$-tuples $(f_0^*, \cdots, f_{2r-1}^*)$.

IV. THE SUPPERPSEUDORANDOMNESS OF THE $\mathcal{V}$-SCHEME

Since 1-round and 2-round $\mathcal{V}$-scheme are not pseudorandom so they are not superpseudorandom. For 3-round $\mathcal{V}$-scheme, we have following fact:

Fact 2. 3-round $\mathcal{V}$-scheme is not superpseudorandom.

Proof. Let $\mathcal{A}_3$ be a distinguisher, it operates as folllows:

1. $\mathcal{A}_3$ chooses two values $Y_1 = (c, d), Y_2 = (c', d') \in I_{2n}$ such that $c \oplus d = c' \oplus d'$.
2. $\mathcal{A}_3$ queries Y_1, Y_2 into f^{-1} to obtain (a, b) and (a', b').
3. $\mathcal{A}_3$ queries (a, b') and (a', b) into f to obtain (s, t) and (s', t').
4. $\mathcal{A}_3$ checks either $s \oplus t = s' \oplus t'$ or not. If $s \oplus t = s' \oplus t'$ then $\mathcal{A}_3$ returns 1 else returns 0.

Let p_3^* be the probability which $\mathcal{A}_3$ returns 1 when f is perfect random permutation over P_{2n}. Thus, $p_3^* = 2^{-n}$. Let p_3 be the probability which $\mathcal{A}_3$ return 1 when $f = \phi(c_0^*, \dots, c_5^*)$. Thus $p_3 = 1$. Indeed:

Because of $c \oplus d = c' \oplus d'$, $c_5^*\left(c_2^*\left(c_1^*(a) \oplus c_0^*(b)\right)\right) = c_5^*\left(c_2^*\left(c_1^*(a') \oplus c_0^*(b')\right)\right)$; it mean that $c_1^*(a) \oplus c_0^*(b) = c_1^*(a') \oplus c_0^*(b')$. So we have $c_1^*(a) \oplus c_0^*(b') = c_1^*(a') \oplus c_0^*(b)$; it mean that $c_5^*\left(c_2^*\left(c_1^*(a) \oplus c_0^*(b')\right)\right) = c_5^*\left(c_2^*\left(c_1^*(a') \oplus c_0^*(b)\right)\right)$. This mean that $s \oplus t = s' \oplus t'$. So the advantage the superpseudorandom distinguisher $\mathcal{A}_3$ is $\text{Adv}_{\mathcal{A}_3}(f, f^*) = |p_3 - p_3^*| = 1 - 2^{-n}$. Thus, 3-round $\mathcal{V}$ scheme is not superpseudorandom □

For the number round is greater than 4, we have following proposition:

Proposition 2. Let $n > 0$ be an integer. Let $c_0^*, \dots, c_{2r-1}^* \in P_n$ are $2r$ $(r \geq 5)$ perfect random permutation and $f^* \in P_{2n}$ is a perfect random permutation. Let $f = \phi(c_0^*, \dots, c_{2r-1}^*)$ denotes the permutation associated with the r-round $\mathcal{V}$-scheme. For any superpseudorandom distinguisher $\mathcal{A}$ allowed to make q encryption and decryption queries we have:

$$\text{Adv}_{\mathcal{A}}(f, f^*) \leq \frac{(r-1)q(q-1)}{2^n} + \frac{q(q-1)}{2.2^{2n}}.$$

Proof. In order to prove this proposition, we need some notations. $I^{\neq}$ denotes the subset of I_n^q consisting of all the q-tuples of pairwise distinct I^n. For $x = (x_1, \dots, x_q), y = (y_1, \dots, y_q) \in I_n^q$ we denotes $x \sim y$ means that $\forall i, j, x_i = x_j$ if and only if $y_i = y_j$. For $X \in \mathcal{X}$ we denote $X = (X_1, \dots, X_q), X_i = (x_i^1, x_i^0)$ and $x^t = (x_i^t)_{i=1..q}$.

This proposition will be proven by using Theorem 2. It means that, we will find a number ϵ such that:

$$H(X, Y) \geq |P_n|^{2r} \cdot \frac{1}{2^{2nq}} \cdot (1 - \epsilon).$$

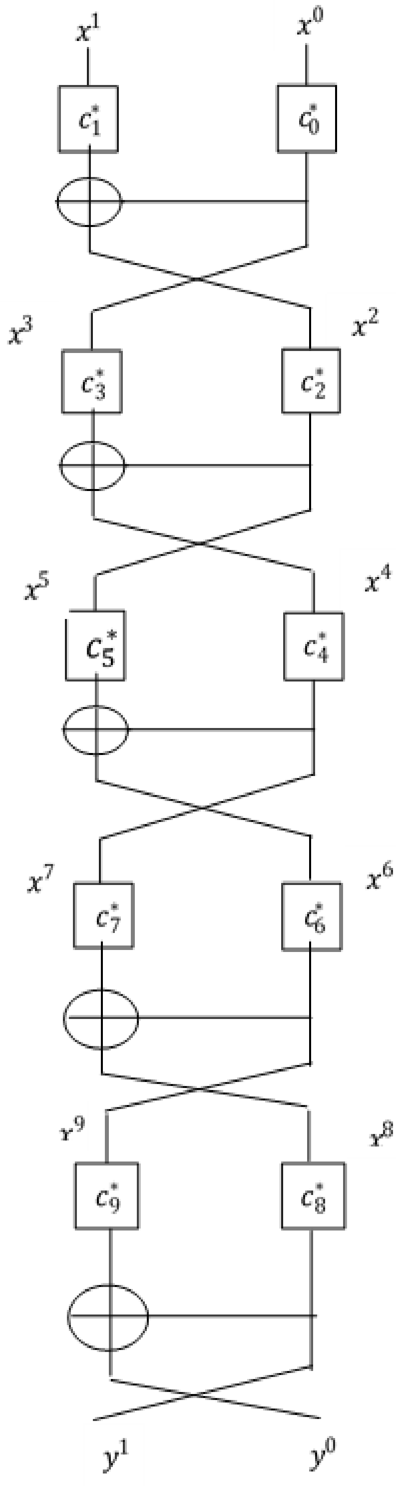


Fig 4. 5-round $\mathcal{V}$-scheme

In order to find ϵ, we establish a lower bound on the the number of permutations $f = (c_0^*, \dots, c_{2r-1}^*)$ such that $f(X) = Y$ for all $X, Y \in \mathcal{X}$. We evaluate for five-round case, then we will generalize for the case $r > 5$. This mean that we find a lower bound on the number of permutations $(c_0^*, \dots, c_9^*)$ such that $f(X) = Y, \forall X, Y \in \mathcal{X}$ or $\forall 1 \le i \le q$:

$$y_i^1 = c_8^*\Bigg(c_7^*\bigg(c_4^*\Big(c_3^*\big(c_0^*(x_i^0)\big) \oplus c_2^*\big(c_1^*(x_i^1) \oplus c_0^*(x_i^0)\big)\Big)\bigg) \oplus c_6^*\bigg(c_5^*\Big(c_2^*\big(c_1^*(x_i^1) \oplus c_0^*(x_i^0)\big)\Big) \oplus c_4^*\Big(c_3^*\big(c_0^*(x_i^0)\big) \oplus c_2^*\big(c_1^*(x_i^1) \oplus c_0^*(x_i^0)\big)\Big)\bigg)\Bigg)$$

$$y_i^1 \oplus y_i^0 = c_9^*\Bigg(c_6^*\bigg(c_5^*\big(c_1^*(x_i^1) \oplus c_0^*(x_i^0)\big) \oplus c_4^*\Big(c_3^*\big(c_0^*(x_i^0)\big) \oplus c_2^*\big(c_1^*(x_i^1) \oplus c_0^*(x_i^0)\big)\Big)\bigg)\Bigg).$$

The number of permutations c_0^* such that $c_0^*(x^0) = x^3$ with $x^3 \sim x^0$ is $|P_n|$. For such c_0^*, there are at least $|P_n|\left(1 - \frac{q(q-1)}{2^n}\right)$ permutations c_1^* such that $c_1^*(x^1) \oplus c_0^*(x^0) = x^2 \in I^{\neq}$. For $x^2 \in I^{\neq}$, the number of permutations c_2^* such that $c_2^*(x^2) = x^5 \in I^{\neq}$ is $|P_n|$. We take c_3^* such that $c_3^*(x^3) \oplus x^5 = x^4 \in I^{\neq}$. In order to establish it we first evaluate the number of permutations c_3^* such that $c_3^*(x_i^3) \oplus x_i^5 = c_3^*(x_j^3) \oplus x_j^5$ with $1 \le i < j \le q$. If $x_i^3 = x_j^3$ then there are no permutation this condition because of $x^5 \in I^{\neq}$. If $x_i^3 \neq x_j^3$ then $c_3^*(x_j^3)$ is determined by the value of $c_3^*(x_i^3)$. So, there are $\frac{|P_n|}{2^n - 1} \le \frac{|P_n|}{2^{n-1}}$ the number of permutations c_3^* such that $c_3^*(x_i^3) \oplus x_i^5 = c_3^*(x_j^3) \oplus x_j^5$. This mean that there are at most $\frac{q(q-1)|P_n|}{2^n}$ permutations c_3^* such that $\exists (i,j), 1 \le i < j \le q$ such that $c_3^*(x_i^3) \oplus x_i^5 = c_3^*(x_j^3) \oplus x_j^5$. (Because there are $\frac{q(q-1)}{2}$ pairs (i,j) such that $1 \le i < j \le q$).

Thus, we have at least $|P_n|\left(1 - \frac{q(q-1)}{2^n}\right)$ permutations c_3^* satisfy $c_3^*(x^3) \oplus x^5 = x^4 \in I^{\neq}$. The number of permutations c_4^* such that $c_4^*(x^4) = x^7 \in I^{\neq}$ is $|P_n|$. We have permutation c_5^* must satisfy $c_5^*(x^5) \oplus x^7 = x^6 \sim y^1 \oplus y^0$ because of $y^1 \oplus y^0 = c_9^*\left(c_6^*(c_5^*(x^5) \oplus x^7)\right)$. Let k be the number of the distinct values $y_i^1 \oplus y_i^0$. Thus, we have $\alpha = \frac{2^n!}{(2^n - k)!}$ values of x^6 such that $x^6 \sim y^1 \oplus y^0$. We will take a loose estimate the number of permutations c_5^* by adding the condition $x^6 \oplus x^7 \in I^{\neq}$. For a fix $x^7 \in I^{\neq}$, we will establish the number c_6^* as defined above such that $x^6 \oplus x^7 \in I^{\neq}$. In order to establish it we first evaluate $S_{i,j}$ which denotes the number of values x^6 such that $x_i^6 \oplus x_i^7 = x_j^6 \oplus x_j^7$ with $1 \le i < j \le q$ (note that x^6 satisfy $x^6 \sim y^1 \oplus y^0$). If $y_i^1 \oplus y_i^0 = y_j^1 \oplus y_j^0$ then $x_i^6 = x_j^6$, so there are not values x^6 satisfy above condition. If $y_i^1 \oplus y_i^0 = y_j^1 \oplus y_j^0$ then x_j^6 is determined by the expression $x_j^6 = x_i^6 \oplus x_i^7 \oplus x_j^7$. Thus, there are $2^n(2^n - 2) \cdots (2^n - k + 1) = \frac{\alpha}{2^n - 1} \le \frac{\alpha}{2^{n-1}}$ elements of $S_{i,j}$. This mean that there are at most $\frac{q(q-1)\alpha}{2^n}$ values x^6 such that $\exists (i,j), 1 \le i < j \le q$ such that $x_i^6 \oplus x_i^7 = x_j^6 \oplus x_j^7$. Thus, for a fix $x^7 \in I^{\neq}$, there are at least $\alpha\left(1 - \frac{q(q-1)}{2^n}\right)$ values x^6 such that $x^6 \sim y^1 \oplus y^0$ and $x^6 \oplus x^7 \in I^{\neq}$. For the x^6 as defined above we have $\Pr[c_5^*(x^5) \oplus x^7 = x^6] = \frac{(2^n - q)!}{2^n!}$ because of $x^5 \in I^{\neq}$. Thus, the number of permutations c_5^* which such that $c_5^*(x^5) \oplus x^7 = x^6 \sim y^1 \oplus y^0$ is greater than $|P_n| \cdot \frac{(2^n - q)!}{2^n!} \cdot \alpha\left(1 - \frac{q(q-1)}{2^n}\right)$. Next, there are $|P_n|$ permutations c_6^* such that $c_6^*(x^6) = x^9 \sim y^0 \oplus y^1$. The permutation c_7^* must satisfy $c_7^*(x^7) \oplus x^9 = x^8 \sim y^1$ because of $y^1 = c_8^*(x^8) = c_8^*(c_7^*(x^7) \oplus x^9)$. Let t be the number of distinct values y_i^1. By the similar method above, there are at least $|P_n| \cdot \frac{(2^n - q)!}{2^n!} \cdot \beta\left(1 - \frac{q(q-1)}{2^n}\right)$ permutations c_7^* such that $c_7^*(x^7) \oplus x^9 = x^8 \sim y^1$ with $\beta = \frac{2^n!}{(2^n - t)!}$. The number of permutations c_8^* such that $c_8^*(x^8) = y^1$ with $x^8 \sim y^1$ is $\frac{|P_n|}{\beta}$. The number of permutations c_9^* such that $c_9^*(x^9) = y^0 \oplus y^1$ with $x^9 \sim y^0 \oplus y^1$ is $\frac{|P_n|}{\alpha}$. From the above arguments, we have:

$$H(X,Y) \geq |P_n| \cdot |P_n| \left(1 - \frac{q(q-1)}{2^n}\right) \cdot |P_n|$$
$$\cdot |P_n| \left(1 - \frac{q(q-1)}{2^n}\right) \cdot |P_n| \cdot |P_n|$$
$$\cdot \frac{(2^n - q)!}{2^n!} \cdot \alpha \left(1 - \frac{q(q-1)}{2^n}\right)$$
$$\cdot |P_n| \cdot |P_n| \cdot \frac{(2^n - q)!}{2^n!}$$
$$\cdot \beta \left(1 - \frac{q(q-1)}{2^n}\right) \cdot \frac{|P_n|}{\beta} \cdot \frac{|P_n|}{\alpha}$$
$$= |P_n|^{10} \left(1 - \frac{q(q-1)}{2^n}\right)^4 \cdot \left(\frac{(2^n - q)!}{2^n!}\right)^2$$
$$\geq |P_n|^{10} \left(1 - \frac{4q(q-1)}{2^n}\right) \cdot \frac{1}{2^{2nq}}.$$

Thus, $\epsilon = \frac{4q(q-1)}{2^n}$.

Next, we will find ϵ in the case $r > 5$. So we will establish the number of permutations $f = (c_0^*, \dots, c_{2r-1}^*)$ such that $f(X) = Y$ for all $X, Y \in \mathcal{X}$. We will evaluate the number of permutations c_{2t-1}^*, c_{2t-2}^* such that $x^{2t+1}, x^{2t} \in I^{\neq}$ with $2 \leq t \leq r-3$. We can assume that $x^{2t-2} \in I^{\neq}$ (because after the first round we can choose permutation c_1^* such that $x^2 \in I^{\neq}$ as the 5 round case), since $x^{2t+1} = c_{2t-2}^*(x^{2t-2})$, so we have the number of permutations c_{2t-2}^* is $|P_n|$. The permutation c_{2t-1}^* must satisfy $c_{2t-1}^*(x^{2t-1}) \oplus x^{2t+1} \in I^{\neq}$ because of $x^{2t} = c_{2t-1}^*(x^{2t-1}) \oplus x^{2t+1}$. By the similar method in 5-round case, we have the number of permutations c_{2t-1}^* is $|P_n|\left(1 - \frac{q(q-1)}{2^n}\right)$. Also, we have the number of permutations $c_0^*, c_{2r-6}^*, c_{2r-4}^*$ are equal $|P_n|$; there are at least $|P_n|\left(1 - \frac{q(q-1)}{2^n}\right), |P_n| \cdot \frac{(2^n-q)!}{2^n!} \cdot \alpha\left(1 - \frac{q(q-1)}{2^n}\right), \frac{(2^n-q)!}{2^n!} \cdot \beta\left(1 - \frac{q(q-1)}{2^n}\right)$ permutations $c_1^*, c_{2r-5}^*, c_{2r-3}^*$, respectively; the number of permutations c_{2r-1}^*, c_{2r-2}^* is $\frac{|P_n|}{\alpha}, \frac{|P_n|}{\beta}$, respectively, with $\alpha = \frac{2^n!}{(2^n-k)!}, \beta = \frac{2^n!}{(2^n-t)!}$ and k, t are the number of distinct value $y_i^1 \oplus y_i^0, y_i^1$. From above agruments we have:

$$H \geq |P_n|^{2r} \cdot \left(\frac{(2^n - q)!}{2^n!}\right)^2 \cdot \left(1 - \frac{q(q-1)}{2^n}\right)^{r-1}$$
$$\geq |P_n|^{2r} \cdot \frac{1}{2^{2nq}} \cdot \left(1 - (r-1) \cdot \frac{q(q-1)}{2^n}\right)$$

Thus, $\epsilon = (r-1) \cdot \frac{q(q-1)}{2^n}$.

Appling Theorem 2 with $\epsilon = (r-1) \cdot \frac{q(q-1)}{2^n}$ we have:

$$\mathrm{Adv}_{\mathcal{A}}(f, f^*) \leq \frac{(r-1)q(q-1)}{2^n} + \frac{q(q-1)}{2.2^{2n}} \square$$

For 4-round $\mathcal{V}$-scheme, we have not proved the superpseudorandomness as well not find a distinguisher to affirm that 4-round $\mathcal{V}$-scheme is not a superpseudorandom permutation. However, if we use Theorem 2 it is easy to see that this technique can not apply. Indeed, let $q = 2$ we choose $X = (X_1, X_2), Y = (Y_1, Y_2)$ with $X_1 = (a, b), X_2 = (a', b), Y_1 = (c, d), Y_2 = (e, f)$ and $a \neq a', c \oplus d = e \oplus f, Y_1 \neq Y_2$. We assume that $f = \phi(c_0^*, \dots, c_7^*)$ be a function such that $f(X) = Y$. As 3-round and 5-round, we use intermediate variables to establish easier. We have $x_1^3 = c_0^*(b) = x_2^3$ and $x_1^2 = c_1^*(a) \oplus x_1^3 \neq c_1^*(a') \oplus x_2^3 = x_2^2$, so $x_1^5 = c_2^*(x_1^2) \neq c_2^*(x_2^2) = x_2^5$. This mean that $x_1^4 = c_3^*(x_1^3) \oplus x_1^5 \neq c_3^*(x_2^3) \oplus x_2^5 = x_2^4$, so $x_1^7 = c_4^*(x_1^4) \neq c_4^*(x_2^4) = x_2^7$. Then, we have $c \oplus d = c_7^*(x_1^7) \neq c_7^*(x_2^7) = e \oplus f$ contradicted with the supposition that $c \oplus d = e \oplus f$. Thus, $H(X,Y) = 0$ this means that we can not apply Theorem 2 to establish the superpseudorandomness for 4-round $\mathcal{V}$-scheme.

VI. CONCLUSION

In this paper, we proposed the new scheme, called $\mathcal{V}$-scheme, and analyzed in detail the pseudorandomness and superpseudorandomness of this scheme. The theoretic results show that $\mathcal{V}$-scheme need at least three rounds to reach pseudorandomness, while the superpseudorandomness is achieved when the number of round is greater than or equal 5. However, we have not established the superpseudorandomness for 4-round $\mathcal{V}$-scheme. When we use Theorem 2 to evaluate 4-round $\mathcal{V}$-scheme, we realize that with $q = 2$ there are values X and Y such that $H = 0$. This mean that if we want to prove 4-round $\mathcal{V}$-scheme is superpseudorandom then we need a more effective approach. Thus, it is an open problem for future research directions. In summary, the results in this paper allow the designer to build block ciphers or cryptography primitives based on block ciphers resisted to generic attacks as chosen ciphertext attack, chosen plaintext attack. According to the $\mathcal{V}$-scheme, we can build 128-bit SPN block ciphers from 64-bit permutations, that are implemented effectively on the current 64-bit platform.

REFERENCES

[1]. Iwata, T. and K. Kurosawa. "On the Pseudorandomness of the AES Finalists-RC6 and Serpent". in International Workshop on Fast Software Encryption, Springer, 2000.

[2]. Wenling, W., F. Dengguo, and C. Hua. "Collision attack and pseudorandomness of reduced-round Camellia". in International Workshop on Selected Areas in Cryptography, Springer, 2004.

[3]. Kang, J.-S., et al. "Pseudorandomness of MISTY-type transformations and the block cipher KASUMI". in Australasian Conference on Information Security and Privacy, Springer, 2001.

[4]. Moriai, S. and S. Vaudenay. "On the pseudorandomness of top-level schemes of block ciphers". in International Conference on the Theory and Application of Cryptology and Information Security, Springer, 2000.

[5]. Iwata, T., T. Yagi, and K. Kurosawa. "On the pseudorandomness of KASUMI type permutations". in Australasian Conference on Information Security and Privacy, Springer, 2003.

[6]. Luby, M. and C. Rackoff, "How to construct pseudorandom permutations from pseudorandom functions". SIAM Journal on Computing, **17**(2): pp. 373-386, 1988..

[7]. Patarin, J. "The "coefficients H" technique". in Selected Areas in Cryptography,Springer, 2008.

[8]. Gilbert, H. and M. Minier. "New results on the pseudorandomness of some blockcipher constructions". in Fast Software Encryption, Springer, 2001.

[9]. Nachef, V., J. Patarin, and E. Volte, "Feistel Ciphers: Security Proofs and Cryptanalysis", Springer, 2017.

[10]. Dodis, Y., et al., "Provable Security of Substitution-Permutation Networks".

[11]. Miles, E. and E. Viola, "Substitution-permutation networks, pseudorandom functions, and natural proofs". Journal of the ACM (JACM), 62(6): pp. 46, 2015

[12]. Piret, G.-F., "Block ciphers: security proofs, cryptanalysis, design, and fault attacks", UCL, 2005.

ABOUT THE AUTHORS

MSc. Bui Cuong Nguyen

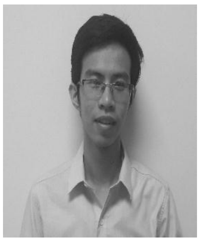
Workplace: Institute of Cryptography Science and Technology.

Email: nguyenbuicuong@gmail.com

The education process: has received a mathematical bachelor degree in Ha Noi National Universiy of Education, in 2004, and has received a mathematical master degree in Ha Noi University of Science, in 2008.

Research today: secret key cryptography,

.



BSc. Tuan Anh Nguyen

Workplace: Institute of Cryptography Science and Technology.

Email: tuananhnghixuan@gmail.com

The education process: has received a mathematical bachelor degree in Ha Noi National Universiy of Education, in 2016.

Research today: secret key cryptography.