Permutation Block Coding for Biometrical Authentication *

Vladimir B. Balakirsky
Institute for Experimental Mathematics
Essen, Germany
e-mail: v_b_balakirsky@rambler.ru

Anahit R. Ghazaryan
Institute for Experimental Mathematics
Essen, Germany
e-mail: a_ghazaryan@rambler.ru

A. J. Han Vinck
Institute for Experimental Mathematics
Essen, Germany
e-mail: vinck@iem.uni-due.de

ABSTRACT

We propose a permutation block coding scheme for biometrical authentication where the permutation stored in the database is chosen according to a non–uniform probability distribution constructed on the basis of the code (contrary to the known proposal published in [1]). As a result, we show that the scheme can be designed in a way that it has a so–called perfect algorithmic secrecy and present examples of schemes with this property.

Keywords

Encoding, Decoding, Authentication, Biometrics

1. BIOMETRICAL AUTHENTICATION

1.1 Coding approaches to a general biometrical authentication problem

The biometrical authentication problem can be presented as designing codes for the schemes in Figure 1. Let $\mathcal{B}$ and $\mathcal{C}$ be subsets of binary vectors of length n. The set $\mathcal{B}$ is the set of biometric vectors, and the probability distribution

$$(\Pr_{\text{bio}}\{ B = \mathbf{b} \}, \mathbf{b} \in \{0,1\}^n)$$

is known. The entries of the set $\mathcal{C}$ are codewords assigned by the designer. The encoding is the transformation of a pair $(\mathbf{x}, \mathbf{b})$, where the vector $\mathbf{b}$ is generated by the source and $\mathbf{x}$ is chosen according to a uniform probability distribution over the code $\mathcal{C}$, to another binary vector $\mathbf{y}$. This vector and the value of $\text{Hash}(\mathbf{x})$ are stored in the database under the name of the person whose biometrical characteristics are expressed by the vector $\mathbf{b}$, where Hash is a fixed "one–way" hash function. Having received a vector $\mathbf{b}'$ and the name of the person, the decoder reads $(\mathbf{y}, \text{Hash}(\mathbf{x}))$ from the database and decodes the codeword as the vector $\hat{\mathbf{x}} \in \mathcal{C}$. If $\text{Hash}(\hat{\mathbf{x}}) = \text{Hash}(\mathbf{x})$, then the identity claim is accepted; otherwise, the identity claim is rejected.

One of possible realizations of the scheme in Figure 1 is an additive block coding scheme [2]. In this case, the biometric vector $\mathbf{b}$ is considered as an additive noise that corrupts the transmitted codeword $\mathbf{x}$ and $\mathbf{y} = \mathbf{x} \oplus \mathbf{b}$. The decoding is based on the observations [2], [3]:

$$\left.\begin{array}{l} \mathbf{y} = \mathbf{x} \oplus \mathbf{b} \\ \mathbf{b}' = \mathbf{b} \oplus \mathbf{e} \end{array}\right\} \Rightarrow \left\{\begin{array}{l} \mathbf{x} \oplus \mathbf{e} = \mathbf{y} \oplus \mathbf{b}', \\ \mathbf{x} \oplus \mathbf{b} = \mathbf{y}. \end{array}\right.$$

Thus, the verifier analyzes the outcomes of transmission of the codeword $\mathbf{x}$ over two parallel channels, called the observation channel, $\mathbf{x} \to \mathbf{x} \oplus \mathbf{e}$, and the biometric channel, $\mathbf{x} \to \mathbf{x} \oplus \mathbf{b}$, while the attacker analyzes the output of the biometric channel.

Notice that the additive block coding scheme can be viewed as a wiretap–type scheme in Figure 2 where the legitimate recever also observes the same vector as the attacker. The permutation block coding scheme is a modification where the sum modulo 2 in the link to the attacker is replaced by a stochastic mapping $f(\mathbf{x}, \mathbf{b})$, see Figure 3. Such a modification is possible when both the vector $\mathbf{x}$ and $\mathbf{b}$ have equal weights and $f(\mathbf{x}, \mathbf{b})$ stands for the binary representation of a permutation $\boldsymbol{\pi}$ that transforms the vector $\mathbf{x}$ to the vector $\mathbf{b}$. Formally, let $\{0,1\}_w^n$ denote the set consisting of binary vectors of the Hamming weight w. The permutation of components of some vector $\mathbf{x} = (x_1, \ldots, x_n) \in \{0,1\}_w^n$ is determined by a vector $\boldsymbol{\pi} \in \mathcal{P}$ in such a way that $\boldsymbol{\pi}(\mathbf{x}) = (x_{\pi_1}, \ldots, x_{\pi_n})$, where $\mathcal{P}$ is the set of all possible permutations of components of the vector $(1, \ldots, n)$. Given a vector $\mathbf{b} \in \{0,1\}_w^n$ and a permutation $\boldsymbol{\pi} \in \mathcal{P}$, let $\boldsymbol{\pi}^{-1} \in \mathcal{P}$ denote the inverse permutation, i.e., $\boldsymbol{\pi}^{-1}(\mathbf{b}) = (b_{i_1(\boldsymbol{\pi})}, \ldots, b_{i_n(\boldsymbol{\pi})})$, where $i_j(\boldsymbol{\pi}) \in \{1, \ldots, n\}$ is the index determined by the equation $\pi_{i_j(\boldsymbol{\pi})} = j$. For all vectors $\mathbf{x}, \mathbf{b} \in \{0,1\}_w^n$, let

$$\mathcal{P}(\mathbf{x} \to \mathbf{b}) \stackrel{\triangle}{=} \{ \boldsymbol{\pi} \in \mathcal{P} : \boldsymbol{\pi}(\mathbf{x}) = \mathbf{b} \} \tag{1}$$

denote the set of permutations that transform the vector $\mathbf{x}$ to the vector $\mathbf{b}$. For example, let $n = 4, k = 2$. The set $\{0,1\}_2^4$ consists of $\binom{4}{2} = 6$ binary vectors of length 4 having the weight 2 and $\mathcal{P}$ is the set consisting of $4! = 24$ permutations of components of the vector $(1, 2, 3, 4)$. For all $\mathbf{x}, \mathbf{b} \in \{0,1\}_2^4$, the set $\mathcal{P}(\mathbf{x} \to \mathbf{b})$ consists of $2!2! = 4$ permutations. In particular, $\mathcal{P}(1100 \to 1010) = \{ 1324, 1423, 2314, 2413 \}$.

Notice that if $\mathbf{b} = \boldsymbol{\pi}(\mathbf{x})$ and $\mathbf{b}' = \mathbf{b} \oplus \mathbf{e}$, then

$$\boldsymbol{\pi}^{-1}(\mathbf{b}') = \boldsymbol{\pi}^{-1}(\mathbf{b}) \oplus \boldsymbol{\pi}^{-1}(\mathbf{e}) = \mathbf{x} \oplus \boldsymbol{\pi}^{-1}(\mathbf{e})$$

and

$$\text{wt}(\boldsymbol{\pi}^{-1}(\mathbf{e})) = \text{wt}(\mathbf{e}), \tag{2}$$

where wt denotes the Hamming weight, i.e., the decoder observes "the transmitted codeword" $\mathbf{x}$ as $\mathbf{x} \oplus \boldsymbol{\pi}^{-1}(\mathbf{e})$. If the source generating the noise vectors is assumed to be a memoryless source, then (2) implies that the presence of the permutation $\boldsymbol{\pi}^{-1}$ does not affect the decoding strategy, and the scheme is equivalent to the one in Figure 3.

1.2 Description of the permutation block coding scheme

Suppose that $\mathcal{B}, \mathcal{C} \subset \{0,1\}^n$. For all vectors $\mathbf{x}, \mathbf{b} \in \{0,1\}_w^n$, let us introduce the probability distribution

$$\gamma_{\mathbf{x},\mathbf{b}} = (\gamma(\boldsymbol{\pi}|\mathbf{x}, \mathbf{b}), \boldsymbol{\pi} \in \mathcal{P})$$

in such a way that $\gamma(\boldsymbol{\pi}|\mathbf{x}, \mathbf{b})$ can be positive only if $\boldsymbol{\pi} \in \mathcal{P}(\mathbf{x} \to \mathbf{b})$. Let us also denote a uniform probability distribution over the set $\mathcal{P}(\mathbf{x} \to \mathbf{b})$ by $\overline{\gamma}_{\mathbf{x},\mathbf{b}}$.

*This work was partially supported by the DFG.

Processing of a given biometric vector $\mathbf{b}$ at the enrollment stage is organized as follows (see Figure 4):

- *choose a codeword* $\mathbf{x}$ *according to a uniform probability distribution over the code* $\mathcal{C}$ *and compute the value of* $\mathrm{Hash}(\mathbf{x})$;
- *given a pair of vectors* $(\mathbf{x}, \mathbf{b}) \in \{0,1\}_w^n \times \{0,1\}_w^n$, *choose a permutation* $\boldsymbol{\pi} \in \mathcal{P}$ *according to the probability distribution* $\gamma_{\mathbf{x},\mathbf{b}}$;
- *store* $(\mathrm{Hash}(\mathbf{x}), \boldsymbol{\pi})$ *in the database.*

The stored data can be used for authentication of a person, whose biometric characteristics are given by the vector $\mathbf{b}'$, as follows (see Figure 5):

- *read the data* $(\mathrm{Hash}(\mathbf{x}), \boldsymbol{\pi})$ *associated with the claimed person from the database;*
- *apply the inverse permutation* $\boldsymbol{\pi}^{-1}$ *to the vector* $\mathbf{b}'$ *and decode the codeword given a received vector* $\boldsymbol{\pi}^{-1}(\mathbf{b}')$ *as* $\hat{\mathbf{x}}_{\boldsymbol{\pi}}$;
- *if* $\mathrm{Hash}(\hat{\mathbf{x}}_{\boldsymbol{\pi}}) = \mathrm{Hash}(\mathbf{x})$, *then accept the identity claim; if* $\mathrm{Hash}(\hat{\mathbf{x}}_{\boldsymbol{\pi}}) \neq \mathrm{Hash}(\mathbf{x})$, *then reject the identity claim.*

1.3 Secrecy of the permutation block coding

In general case, the attacker applies a fixed function $\psi : \mathcal{P} \to \mathcal{C}$ to the permutation $\boldsymbol{\pi}$ stored in the database and submits the vector $\mathbf{b}' = \boldsymbol{\pi}(\psi(\boldsymbol{\pi}))$ to the verifier. Thus, the probability of successful attack is expressed as

$$\Lambda_{\mathcal{C},\gamma}(\psi) = \frac{1}{M} \sum_{\mathbf{x} \in \mathcal{C}} \sum_{\mathbf{b}} \Pr_{\mathrm{bio}}\{B = \mathbf{b}\} \sum_{\boldsymbol{\pi} \in \mathcal{P}} \gamma(\boldsymbol{\pi}|\mathbf{x}, \mathbf{b}) \chi\{\, \psi(\boldsymbol{\pi}) = \mathbf{x}\,\},$$

where $M = |\mathcal{C}|$ and χ denotes the indicator function: $\chi\{\mathcal{S}\} = 1$ is the statement $\mathcal{S}$ is true and $\chi\{\mathcal{S}\} = 0$ otherwise. Notice that the vector $\mathbf{x} \in \{0,1\}_w^n$ and the permutation $\boldsymbol{\pi} \in \mathcal{P}$ uniquely determine the vector $\mathbf{b}^0 \in \{0,1\}_w^n$ such that $\boldsymbol{\pi} \in \mathcal{P}(\mathbf{x} \to \mathbf{b}^0)$. Namely, $\mathbf{b}^0 = \boldsymbol{\pi}(\mathbf{x})$, and the sum at the right-hand side contains at most one non–zero term.

The attacker has two simple possibilities: 1) fix a codeword $\mathbf{x}' \in \mathcal{C}$ and submit the vector $\mathbf{b}' = \boldsymbol{\pi}(\mathbf{x}')$; 2) submit the most likely biometric vector. In the first case, the attacker has to know the code $\mathcal{C}$ and the stored permutation $\boldsymbol{\pi}$. In the second case, he does not know these data and equivalent to an attacker, who does not have access to the database and ignorant about the code. One can easily see that the probabilities of successful attacks are equal to $1/M$ and Q^*_{bio}, respectively. Therefore the probability of successful attack under the maximum *a posteriori* probability decoding of the codeword is bounded from below as follows:

$$\Lambda_{\mathcal{C},\gamma}(\psi) \geq \max\left\{\frac{1}{M},\, Q^*_{\mathrm{bio}}\right\}.$$

where

$$Q^*_{\mathrm{bio}} \stackrel{\triangle}{=} \max_{\mathbf{b}} \Pr_{\mathrm{bio}}\{B = \mathbf{b}\}.$$

We will present examples of schemes where this inequality is tight.

2. EXAMPLES OF SPECIFIC SCHEMES

Let $n = 8$, $w = 4$, $M = 4$, and let the codewords be specified by the matrix

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \\ \mathbf{x}_4 \end{bmatrix} = \begin{bmatrix} 00110011 \\ 01010101 \\ 10101010 \\ 11001100 \end{bmatrix}.$$

Suppose also that the biometric vector processed at the enrollment stage is one of rows of the matrix

$$\begin{bmatrix} \mathbf{b}_1 \\ . \\ . \\ \mathbf{b}_6 \end{bmatrix} = \begin{bmatrix} 00001111 \\ 00110011 \\ 01010101 \\ 10101010 \\ 11001100 \\ 11110000 \end{bmatrix}.$$

Denote $\mathcal{C} = \{\, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$ and $\mathcal{B} = \{\, \mathbf{b}_1, \ldots, \mathbf{b}_6\}$.

Proposition *For all pairs of vectors* $(\mathbf{x}, \mathbf{b}) \in \mathcal{C} \times \mathcal{B}$,

$$|\, \mathcal{P}(\mathbf{x} \to \mathbf{b})\,| = (4!)^2 = 576 \tag{3}$$

and

$$|\, \mathcal{P}_{\mathcal{C}\to\mathcal{B}}(\mathbf{x} \to \mathbf{b})\,| = 4(2!)^4 = 64, \tag{4}$$

where $\mathcal{P}_{\mathcal{C}\to\mathcal{B}}(\mathbf{x} \to \mathbf{b})$ *denotes the set of permutations* $\boldsymbol{\pi} \in \mathcal{P}(\mathbf{x} \to \mathbf{b})$ *such that* $\boldsymbol{\pi}(\mathbf{x}') \in \mathcal{B}$ *for all* $\mathbf{x}' \in \mathcal{C}$.

Let us illustrate considerations by the following examples:

$$\begin{bmatrix} \boldsymbol{\pi}' \\ \boldsymbol{\pi}'(\mathbf{x}_1) \\ \boldsymbol{\pi}'(\mathbf{x}_2) \end{bmatrix} = \begin{bmatrix} 1 & 2 & 5 & 6 & 3 & 4 & 7 & 8 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix},$$

$$\begin{bmatrix} \boldsymbol{\pi}'' \\ \boldsymbol{\pi}''(\mathbf{x}_1) \\ \boldsymbol{\pi}''(\mathbf{x}_2) \end{bmatrix} = \begin{bmatrix} 1 & 2 & 6 & 5 & 3 & 4 & 7 & 8 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

The permutations $\boldsymbol{\pi}'$ and $\boldsymbol{\pi}''$ belong to the set $\mathcal{P}$. Furthermore, $\boldsymbol{\pi}'(\mathbf{x}_1) = \boldsymbol{\pi}''(\mathbf{x}_1) = \mathbf{b}_1$. However $\boldsymbol{\pi}'(\mathbf{x}_2) \in \mathcal{B}$, while $\boldsymbol{\pi}''(\mathbf{x}_2) \notin \mathcal{B}$. Suppose that $\boldsymbol{\pi}'$ is the permutation stored in the database. The attacker applies this permutation to all codewords of the code $\mathcal{C}$ and constructs the list $\boldsymbol{\pi}'(\mathbf{x}_1), \ldots, \boldsymbol{\pi}'(\mathbf{x}_4)$. All entries of this list are possible biometric vectors. If the permutation $\boldsymbol{\pi}''$ is stored in the database, then the list $\boldsymbol{\pi}'(\mathbf{x}_1), \ldots, \boldsymbol{\pi}'(\mathbf{x}_4)$ contains only 2 biometric vectors. The probability of successful attack is greater in the second case, and the permutation $\boldsymbol{\pi}'$ can be considered as a "bad" permutation.

The result of Proposition shows that the most of permutations are bad permutations. This fact leads to the statement that the uniform probability distribution over the set $\mathcal{P}(\mathbf{x} \to \mathbf{b})$, where $\mathbf{x}$ is the selected codeword and $\mathbf{b}$ is the biometric vector, brings a rather poor performance. Namely, suppose that the probability distribution over the set $\mathcal{B}$ is uniform, i.e.,

$$\Pr_{\mathrm{bio}}\{B = \mathbf{b}\} = 1/6,\ \mathbf{b} \in \mathcal{B}.$$

Let $\mathbf{x}$ be the codeword of the code $\mathcal{C}$ used at the enrollment stage. If $\gamma_{\mathbf{x},\mathbf{b}} = \overline{\gamma}_{\mathbf{x},\mathbf{b}}$, then the permutation is uniformly chosen from the set containing 576 entries. Only 64 of these permutations have the property that the set $\boldsymbol{\pi}(\mathbf{x})$, $\mathbf{x} \in \mathcal{C}$ contains 4 biometric vectors, and the probability of successful attack is equal to 1/4. For the other 512 permutations, the set $\boldsymbol{\pi}(\mathbf{x})$, $\mathbf{x} \in \mathcal{C}$, contains 2 biometric vectors, and the probability of successful attack is equal to 1/2. Thus

$$\Lambda_{\mathcal{C},\overline{\gamma}}(\psi^{\mathrm{MAP}}) = \frac{64}{576}(1/4) + \frac{512}{576}(1/2) = 17/36.$$

Let us assign $\gamma_{\mathbf{x},\mathbf{b}}$ as a uniform probability distribution over the set $\mathcal{P}_{\mathcal{C}\to\mathcal{B}}(\mathbf{x} \to \mathbf{b})$ consisting of 64 entries. In all cases, the list $\boldsymbol{\pi}(\mathbf{x})$, $\mathbf{x} \in \mathcal{C}$, contains 4 biometric vectors, and the probability of successful attack is equal to 1/4. As a result, the probability of successful attack is expressed as

$$\Lambda_{\mathcal{C},\gamma}(\psi^{\mathrm{MAP}}) = \frac{64}{64}(1/4) = 1/4,$$

which is approximately twice less than $\Lambda_{\mathcal{C},\overline{\gamma}}(\psi^{\mathrm{MAP}})$. Moreover, we obtain that the lower bound $1/M$ on the probability $\Lambda_{\mathcal{C},\gamma}(\psi^{\mathrm{MAP}})$ is attained with the equality.

Let us consider a non–uniform probability distribution over the set $\mathcal{B}$. Namely, let $a \in [1/4, 1/2]$ be a fixed parameter and let

$$\Pr_{\mathrm{bio}}\{B = \mathbf{b}\} = \begin{cases} a, & \text{if } \mathbf{b} \in \{00001111, 11110000\}, \\ 1/4 - a/2, & \text{if } \mathbf{b} \in \mathcal{B}\backslash\{00001111, 11110000\}. \end{cases}$$

Notice that the set $\mathcal{P}_{\mathcal{C}\to\mathcal{B}}(\mathbf{x}_1 \to \mathbf{b}_1)$ contains 32 permutations $\boldsymbol{\pi}$ such that

$$\{\boldsymbol{\pi}(\mathbf{x}_1), \boldsymbol{\pi}(\mathbf{x}_2), \boldsymbol{\pi}(\mathbf{x}_3), \boldsymbol{\pi}(\mathbf{x}_4)\} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5, \mathbf{b}_6\}$$

and 32 permutations $\boldsymbol{\pi}$ such that

$$\{\boldsymbol{\pi}(\mathbf{x}_1), \boldsymbol{\pi}(\mathbf{x}_2), \boldsymbol{\pi}(\mathbf{x}_3), \boldsymbol{\pi}(\mathbf{x}_4)\} = \{\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_4, \mathbf{b}_6\}.$$

Let us denote the subsets of these permutations by $\mathcal{P}'_{\mathcal{C}\to\mathcal{B}}(\mathbf{x}_1 \to \mathbf{b}_1)$ and $\mathcal{P}''_{\mathcal{C}\to\mathcal{B}}(\mathbf{x}_1 \to \mathbf{b}_1)$, respectively. Let
(a) $\gamma_{\mathbf{x}_1,\mathbf{b}_1}, \gamma_{\mathbf{x}_1,\mathbf{b}_6}$ be uniform probability distributions over the set $\mathcal{P}_{\mathcal{C}\to\mathcal{B}}(\mathbf{x}_1 \to \mathbf{b}_1)$;
(b) $\gamma_{\mathbf{x}_1,\mathbf{b}_2}, \gamma_{\mathbf{x}_1,\mathbf{b}_5}$ be uniform probability distributions over the set $\mathcal{P}'_{\mathcal{C}\to\mathcal{B}}(\mathbf{x}_1 \to \mathbf{b}_1)$;
(c) $\gamma_{\mathbf{x}_1,\mathbf{b}_3}, \gamma_{\mathbf{x}_1,\mathbf{b}_4}$ be uniform probability distributions over the set $\mathcal{P}''_{\mathcal{C}\to\mathcal{B}}(\mathbf{x}_1 \to \mathbf{b}_1)$.
If $\boldsymbol{\pi} \in \mathcal{P}'_{\mathcal{C}\to\mathcal{B}}(\mathbf{x}_1 \to \mathbf{b}_1)$, then the *a posteriori* probabilities associated with the biometric vectors $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_5, \mathbf{b}_6$ are equal to

$$\frac{1}{32}(a/2, 1/2 - a/2, 1/2 - a/2, a/2).$$

However $a/2 \geq 1/2 - a/2$, and the attacker outputs either the key codeword, which is mapped to the vector $\mathbf{b}_1$, or the key codeword, which is mapped to the vector $\mathbf{b}_6$. Similar considerations can be presented for the permutations belonging to the set $\mathcal{P}''_{\mathcal{C}\to\mathcal{B}}(\mathbf{x}_1 \to \mathbf{b}_1)$. As a result, we conclude that

$$\Lambda_{\mathcal{C},\gamma}(\psi^{\mathrm{MAP}}) = 64(a/64) = a,$$

i.e., the lower bound Q^*_{bio} on the probability $\Lambda_{\mathcal{C},\gamma}(\psi^{\mathrm{MAP}})$ is attained with the equality.

Let us consider the error–correcting capabilities of the verifier, who processes data of a legitimate user. Let P_w denote the probability that the vector $\mathbf{b}'$ differs from the vector $\mathbf{b}$ in w positions, $w = 0, \ldots, 8$. Then, assuming that the vectors $\mathbf{b}'$ are uniformly distributed over the set of vectors located at a fixed distance from the vector $\mathbf{b}$, we obtain that the probability of correct decoding for the code $\mathcal{C}$ and the threshold $T = 2$ is equal to $\tilde{\Lambda}^{(2)}_{\mathcal{C}} = P_0 + P_1 + (16/28)P_2$, since the decoder makes the correct decision for all error patterns of weight at most 1 and for 16 error patterns of weight 2 (the total number of error patterns of weight 2 is equal to 28). Suppose that the precessed biometric vectors are constructed as a concatenation of L vectors $\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(L)} \in \mathcal{B}$, i.e., the total length of the vector is equal to $8L$. Suppose also that the vectors $\mathbf{b}^{(1)}, \ldots, \mathbf{b}^{(L)}$ are independently generated according to a uniform probability distribution over the set $\mathcal{B}$. Let the verifier make the acceptance decision if and only if such a decision is made for all L entries. Then the probability of correct decision is equal to $(\tilde{\Lambda}^{(2)}_{\mathcal{C}})^L$. On the other hand, the probability of successful attack, when the probability distributions $\gamma_{\mathbf{x},\mathbf{b}}$ are used is equal to $(1/4)^L$.

2.1 Proof of Proposition

Suppose that $(\mathbf{x}, \mathbf{b}) = (\mathbf{x}_1, \mathbf{b}_1)$, i.e., $\mathbf{x} = 00110011$ and $\mathbf{b} = 00001111$. Equality (3) immediately follows from the fact that both $\mathbf{x}$ and $\mathbf{b}$ contain 4 zeroes and 4 ones. Notice that $\boldsymbol{\pi}(\mathbf{x}_1) \in \mathcal{B}$ implies $\boldsymbol{\pi}(\mathbf{x}_4) \in \mathcal{B}$, and $\boldsymbol{\pi}(\mathbf{x}_2) \in \mathcal{B}$ implies $\boldsymbol{\pi}(\mathbf{x}_3) \in \mathcal{B}$. Therefore,

$$\mathcal{P}_{\mathcal{C}\to\mathcal{B}}(\mathbf{x} \to \mathbf{b}) = \{\, \boldsymbol{\pi} \in \mathcal{P}(\mathbf{x}_1 \to \mathbf{b}_1) : \boldsymbol{\pi}(\mathbf{x}_2) \in \mathcal{B} \,\}.$$

If $\boldsymbol{\pi} \in \mathcal{P}(\mathbf{x}_1 \to \mathbf{b}_1)$ and $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_8)$, then $(\pi_1, \ldots, \pi_4)$ is a permutation of components of the vector $(1, 2, 5, 6)$ and $(\pi_5, \ldots, \pi_8)$ is a permutation of components of the vector $(3, 4, 7, 8)$.

The condition $\pi(\mathbf{x}_2) \in \mathcal{B}$ is satisfied if and only if there is a vector $\mathbf{s} \in \{0011, 0101, 1010, 1100\}$ with the following property: the first 4 components of the vector $\boldsymbol{\pi}$ and the last 4 components of the vector $\boldsymbol{\pi}$ specify the permutations of columns of the matrices

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

such that the 2-nd rows are equal to the vector $\mathbf{s}$. There are 4 possible vectors $\mathbf{s}$ and each vector can be constructed by $(2!)^4$ permutations. Therefore, this observation proves (4) for $(\mathbf{x}, \mathbf{b}) = (\mathbf{x}_1, \mathbf{b}_1)$. By the symmetric properties of the sets $\mathcal{C}$ and $\mathcal{B}$, one can see that considerations above also prove the statement for any fixed pair $(\mathbf{x}, \mathbf{b}) \in \mathcal{C} \times \mathcal{B}$.

3. CONCLUSION

The best guess of the message generated by a source is the message having the maximum *a priori* probability. When some information about the message is stored in the database, the best guess becomes the message having the maximum *a posteriori* probability. If both guesses coincide and this is true for any message, which can be generated by the source, then we claim that the system has a perfect algorithmic secrecy: although the attacker has access to the database, the optimum guess is always the most likely message, and the probability of correct guess is equal to the probability that the source generated this message. Therefore, the difference between the entropies of the *a priori* and *a posteriori* probability distributions over the messages, which is usually considered as a measure of secrecy of the system, can be large, but the attacker cannot include his knowledge about the message into the guessing algorithm to increase the probability of correct guess. By a proper assignment the probability distribution over the set of permutations that transform two binary vectors of the same weight to each other, it is possible to design the permutation block coding schemes having this property.
The permutation scheme can be viewed as an algorithm with an additional randomization over the set of permutations introduced in the data processing as compared to the additive scheme. The outcomes of the random experiment are public, but the secrecy can be much higher.

REFERENCES

[1] Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," *Advances in Cryptography – EUROCRYPT. Lecture Notes in Computer Science*, no. 3027, pp. 523–540, 2004.

[2] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," *Proc. ACM Conf. Computer and Communication Security*, 1999.

[3] V. B. Balakirsky, A. R. Ghazaryan, and A. J. Han Vinck, "Performance of additive block coding schemes oriented to biometric authentication", *Proc. 29th Symposium on Information Theory in the Benelux*, L. Van de Perre, A. Dejonghe, V. Ramon (eds.), May 29–30, Leuven, Belgium, pp. 19–26, 2008.

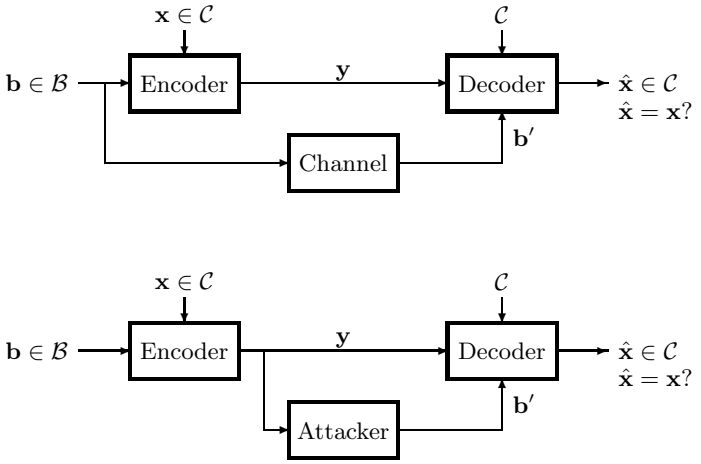

Figure 1. General authentication scheme.

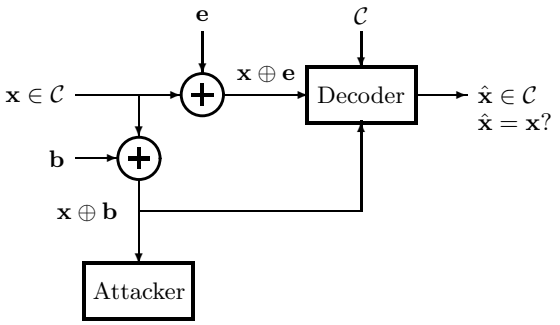

Figure 2. Wiretap-type additive block coding scheme.

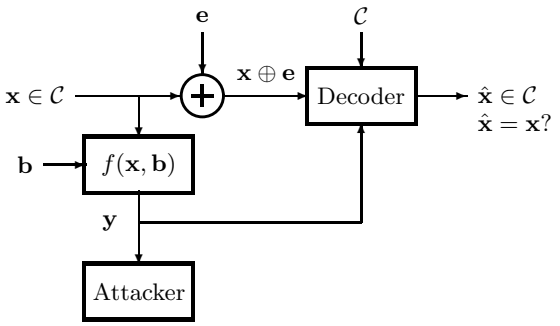

Figure 3. Modified wiretap-type block coding scheme.

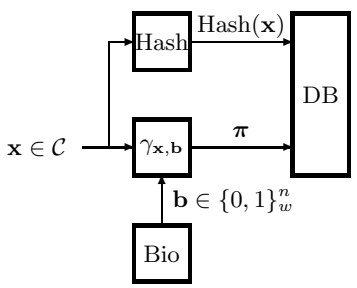

Figure 4: Processing data at the enrollment stage of the permutation block coding scheme.

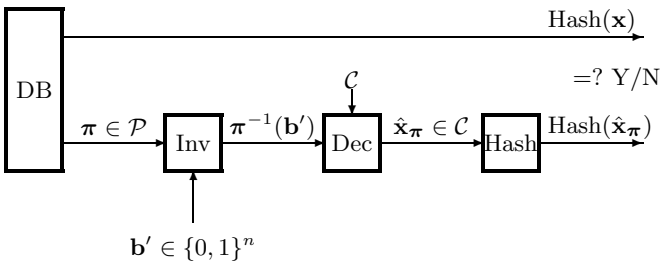

Figure 5: Processing data at the authentication stage of the permutation block coding scheme.