A Theoretical Framework for Constructing Matching Algorithms Secure against Wolf Attack

Manabu INUMA†, Nonmember, Akira OTSUKA††, Member, and Hideki IMAI†††, Fellow, Honorary Member

SUMMARY The security of biometric authentication systems against impersonation attack is usually evaluated by the false accept rate, FAR. The false accept rate FAR is a metric for zero-effort impersonation attack assuming that the attacker attempts to impersonate a user by presenting his own biometric sample to the system. However, when the attacker has some information about algorithms in the biometric authentication system, he might be able to find a “strange” sample (called a wolf) which shows high similarity to many templates and attempt to impersonate a user by presenting a wolf. Une, Otsuka, Imai [22], [23] formulated such a stronger impersonation attack (called the wolf attack), defined a new security metric (called wolf attack probability, WAP), and showed that WAP is extremely higher than FAR in a fingerprint-minutiae matching algorithm proposed by Ratha et al. [19] and in a finger-vein-patterns matching algorithm proposed by Miura et al. [15]. Previously, we constructed secure matching algorithms based on a feature-dependent threshold approach [8] and showed that if the score distribution is perfectly estimated for each input feature data, then the proposed algorithms can lower WAP to a small value almost the same as FAR. In this paper, in addition to reintroducing the results of our previous work [8], we show that the proposed matching algorithm can keep the false reject rate (FRR) low enough without degrading security, if the score distribution is normal for each feature data.

key words: biometrics, wolf attack, wolf attack probability, matching algorithm, security

1. Introduction

Biometric authentication systems automatically identify or verify individuals by physiological or behavioral characteristics. They are used in various services such as immigration control at airports, banking transactions at ATMs, access control to restricted areas in buildings. The more familiar to us are the biometric authentication systems, the more increases the threat of (malicious) impersonation attacks. Therefore, we need appropriate criteria to evaluate the security against such impersonation attacks and provably secure biometric authentication systems.

The false accept rate (FAR) (see the definition (4)) is the probability that a user’s presentation of his/her own biometric object is incorrectly declared to be a match for a different user’s enrolled biometric template*. FAR is a traditional measure to evaluate the recognition accuracy of biometric authentication systems and can be regarded as a measure to evaluate the security against zero-effort impersonation attack in which the attacker attempts to impersonate a user by presenting his own biometric object to the system. However, in the real world, this is an unreasonable assumption and the attacker is obviously more skillful.

Some researchers addressed intentional impersonation attacks stronger than zero-effort attack. For example, Ratha et al. [19] analyzed a brute-force attack against their proposed simple fingerprint matching algorithm, where the attacker presents an artificial object whose characteristic points (called minutiae) are as many as those of enrolled human fingerprints and randomly and uniformly distributed over the fingerprint area. They stated that the fingerprint-minutiae matching algorithm is sufficiently secure against such a brute-force attack. However, if the attacker has some information about algorithms used in the system, he might be able to find a strange object which shows high similarity to many templates and, by presenting such a fake object**, successfully impersonate a user with extremely higher probability than FAR. Une, Otsuka, Imai [22], [23] formulated such a stronger impersonation attack (called wolf attack) and defined a security metric (called wolf attack probability, WAP) to be the upper bound of the success probability of wolf attack (see Definition 1, 2, 3). They suggested that WAP can be used as a security measure to evaluate the lower bound of a security level in a biometric authentication system. Actually, they showed that, in a fingerprint-minutiae matching algorithm proposed by Ratha et al. [19], WAP is extremely higher than FAR or the success probability of the brute-force attack estimated by Ratha et al. and, moreover, in a finger-vein-patterns matching algorithm proposed by Miura et al. [15], there exists a universal wolf (see Definition 1), namely WAP = 1.

As far as we know, our previous paper [8] is the first paper that addressed matching algorithms secure against wolf attack. In our previous paper [8], we constructed secure matching algorithms based on a feature-dependent threshold approach and proposed a theoretical framework for constructing matching algorithms secure against wolf attack. Some previous research efforts aim to increase the recognition accuracy and use user-dependent threshold or score-normalization technique (cf. [4], [14], [20], [26]). However,

*This is the definition of the false match rate (FMR) rather than FAR. In this paper, for the simplicity, we assume that the fail-to-acquire rate (FTA) equals 0, and FAR = FMR.
**Although many researchers study spoof detection techniques (cf. [2], [5], [7], [18]), we cannot completely erase threats of spoofing attacks using fake objects.
since these previous algorithms use the parameters estimated from a distribution of matching scores between (human) users or need an ideal assumption that every bit of feature data independently and identically contributes to the score distribution, they might not be able to prevent attacks by using wolves which have "strange" score distributions. Our proposed matching algorithms determine a threshold by using parameters (e.g. the mean, standard deviation or the entropy), estimated from the score distribution for each input feature data and can keep WAP low. Moreover, in this paper, we theoretically study the accuracy of our proposed algorithm and show that it can keep the false reject rate (FRR) (see the definition (2)) sufficiently low without degrading security, if the score distribution is normal for each feature data.

After our previous work [8], some researchers, [12], [16], and [17], have studied matching algorithms secure against wolf attack. Monden [16] presents a matching algorithm provably secure against wolf attack. His proposed algorithm utilizes the distribution of similarity scores generated from comparisons with all users’ enrolled templates. Actually, it can be checked that Monden’s algorithm is equivalent to our proposed algorithm in [8, Sect. 3.1]. Kojima et al. [12] propose a slightly more efficient matching algorithm which directly uses similarity scores obtained from comparisons with all users’ enrolled templates. They show that, in their proposed algorithm, there exists a trade-off between the security against wolf attack and the accuracy with respect to FRR. Murakami and Takahashi [17] propose an efficient matching algorithm which uses similarity scores obtained by comparisons with a specified number of templates and calculates a conditional probability that, given those similarity scores, the user to be authenticated is genuine. They experimentally show the security against wolf attack. However, a theoretical proof has not been found yet.

Our generalized matching algorithm gives a theoretical framework for constructing secure matching algorithms. In the real world, it might be difficult to perfectly estimate (the entropy of) the distribution for each input feature data. However, a more accurate estimation of the entropy can achieve a lower WAP. Namely, there is a trade-off between the efficiency of the matching algorithm and the achievable WAP in the matching algorithm.

This paper continues as follows. In Sect. 2, we introduce a model of a biometric authentication system and formulate FRR and FAR. In Sect. 3, we give definitions of wolf attack, WAP and security against wolf attack, and discuss some properties of perfectly secure matching algorithms (Theorem 1). In Sect. 4, we introduce matching algorithms and theorems (Theorem 2, 3) in [8] and study on the recognition accuracy of these matching algorithms. In Sect. 5, we summarize our results.

2. Preliminaries

Biometric authentication systems automatically recognize individuals by using their physiological or behavioral characteristics, for example, fingerprints, iris, faces, voice, and so on. We call such a human physiological or behavioral characteristic a biometric object. A biometric authentication system can be used for verification or identification. In verification, a user claims an identity, biometric features extracted from a biometric object submitted by the user are compared to the enrolled reference data for the claimed identity, and an accept or reject decision is returned. In identification, a search of the enrolled database is performed (for example, comparisons between the input biometric features and all enrolled reference data in the database are executed), and a candidate list for the user’s probable identity is returned. In this paper, we focus on verification systems.

Let \( U \) be a finite set consisting of all users of the biometric authentication system. We assume that each user \( u \in U \) has unique his/her own biometric object \( b_u \). We identify each user \( u \) with his/her own object \( b_u \) and use the notation \( u \) instead of \( b_u \), namely \( U \) can be regarded as a set consisting of all users’ biometric objects. Let \( M \) be a finite set whose elements represent users’ biometric features. When a user presents his/her own biometric object to the sensor of the system, the system generates an element of \( M \) from the presented biometric object as follows. The system captures a biometric sample \(^1\) from the presented object, extracts biometric features from the captured sample, and represents the biometric features as an element \( x \) of \( M \). Then we call \( x \) a feature data of \( u \). Since two feature data captured from \( u \) are rarely identical, we let \( X_u \) denote a random variable on \( M \) representing noisy variations of feature data of \( u \), namely \( \Pr(X_u = x) \) is the probability that biometric features scanned from \( u \) will be transformed into \( x \).

We define an enrollment phase and a verification phase of a biometric authentication system as follows, where every user honestly presents his/her own biometric object. In the enrollment phase, the system extracts a feature data \( x \in M \) from a biometric object presented by a user \( u \in U \) and stores \( x \) along with the user's identity in the database. In the verification phase, a user \( v \in U \) claims an identity of \( w \in U \) \(^{11}\) and presents his/her own biometric object to the system. The system extracts a feature data \( y \in M \) from the presented object and inputs \( y \) and the feature data \( z \) corresponding to the claimed identity into the matching algorithm MATCH. The matching algorithm MATCH on inputs \( y \) and \( z \) returns an accept or reject decision, namely MATCH\((y, z) = \text{accept or reject}\). Almost every existing matching algorithm uses the similarity score \( s: M \times M \rightarrow R \) and a threshold \( \theta \in R \), where \( R \) is the field of the real numbers, as follows:

\(^1\)Fingerprint image, face image and iris image are samples.

\(^{11}\)We consider that \( v \) does not always correctly claim his/her own identity.
\[
\text{MATCH}(y, z) = \begin{cases} 
\text{accept} & \text{if } s(y, z) > \theta \\
\text{reject} & \text{if } s(y, z) \leq \theta.
\end{cases}
\] (1)

The higher \( s(y, z) \) indicates that \( y \) is closer to \( z \).

Assume that, for any \( u, v \in U \), the variation of feature data of \( u \) in the enrollment phase and the variation of feature data of \( v \) in the verification phase are independent, namely, for any \( u, v \in U \) and any \( x, y \in M \),

\[
\Pr[X_u = x, X_v = y] = \Pr[X_u = x] \Pr[X_v = y].
\]
The above assumption is true even if \( u = v \).

For any set \( U \) (resp. any random variable \( X \) on a set \( M \)), the notation \( u \leftarrow U \) (resp. \( x \leftarrow X \)) denotes that \( u \) is chosen from the set \( U \) uniformly at random (resp. \( x \) is chosen according to \( X \)). For any function \( f \) on the set \( U \) (resp. the set \( M \)), the notation \( E_{\leftarrow U} f(x) \) (resp. \( E_{\leftarrow M} f(x) \)) denotes the expected value of \( f \) under the condition that \( u \) is chosen from the set \( U \) uniformly at random (resp. \( x \) is chosen according to \( X \)), namely

\[
E_{\leftarrow U} f(x) = \frac{1}{n} \sum_{u \in U} f(u)
\]

(resp. \( E_{\leftarrow M} f(x) = \sum_{x \in M} \Pr[X = x] f(x) \))

where \( n \) is the number of elements of \( U \).

Here, we will formulate the \textit{false reject rate}, \( \text{FRR} \), and the \textit{false accept rate}, \( \text{FAR} \). The metric \( \text{FRR} \) is the probability that a truthful identity claim is incorrectly rejected, and is formulated by

\[
\text{FRR} = \frac{1}{n} \sum_{u \in U} \Pr[X_u = x] \Pr[X_u = y] .
\] (2)

For each user \( u \in U \), let \( \text{FRR}_u \) denote the probability that the user \( u \) with the correct identity claim \( u \) is rejected, which is formulated by

\[
\text{FRR}_u = \sum_{(x, y) \in M \times M} \Pr[X_u = x] \Pr[X_u = y] .
\] (3)

It is easy to check that \( \text{FRR} = \frac{1}{n} \sum_u \text{FRR}_u \). The metric \( \text{FAR} \) is the probability that a wrongful identity claim is incorrectly accepted and is formulated by

\[
\text{FAR} = \frac{1}{n(n-1)} \sum_{(u, v) \in (U \times U) \setminus (u, u)} \Pr[\text{MATCH}(x, y) = \text{accept}]
\] (4)

where \((U \times U)^\text{diff} = \{(u, v) \in U \times U \mid u \neq v\} \). To be precise, \( \text{FRR} \) and \( \text{FAR} \) formulated in this paper are the \textit{false non-match rate}, \( \text{FNMR} \), and the \textit{false match rate}, \( \text{FMR} \), respectively. However, for the simplicity, we assume that the \textit{false-to-acquire rate}, \( \text{FTA}^1 \), equals 0 and \( \text{FRR} = \text{FNMR} \) and \( \text{FAR} = \text{FMR} \).

The metrics \( \text{FRR} \) and \( \text{FAR} \) evaluate a recognition accuracy of biometric systems. \( \text{FAR} \) can be regarded as a metric to evaluate the security against the zero-effort impersonation attack. The zero-effort approach assumes that an attacker attempting to impersonate a genuine user will present his/her own biometric object. However, in the real world, this assumption is unreasonable, and we should consider more skillful attacker. In the next section, we will introduce a theory on \textit{wolf attack}, an intentional impersonation attack stronger than zero-effort attack, proposed by Une, Otsuka, Imai [22], [23].

### 3. Wolf Attack and Wolf Attack Probability

A biometric object whose feature data has high similarity with many users’ feature data is called a \textit{wolf} (cf. [6], [25]). From the viewpoint of the security against impersonation, we consider that the attacker might be able to artificially make wolves. Actually, Matsumoto et al. [13] show by experimentation that some artefacts can be falsely accepted by some biometric sensors. Consequently, we extend the definition of \textit{wolf} to incorporate with artefacts as follows.

Let \( A \) be a set of all possible objects including non-biometric object such as artefacts or synthetic objects. Note that \( A \supset U \). Assume that, when a non-biometric object \( w \in A \setminus U \) is presented to the sensor, some features are extracted from \( w \) and the features are represented as an elements \( x \) of \( M \) in the same manner as feature data of biometric objects. We also call \( x \) a \textit{feature data} of \( w \). For each \( w \in A \), let \( \text{FAR}_w \), the \textit{false accept rate} of \( w \), be the probability that the object \( w \) with a wrong identity claim \( v \neq w \) will be incorrectly accepted and let \( \text{AR}_w \), the \textit{accept rate} of \( w \), be the probability that the object \( w \) with an identity chosen uniformly at random will be accepted. Namely, \( \text{FAR}_w \) and \( \text{AR}_w \) are respectively defined by

\[
\text{FAR}_w = \frac{1}{n-1} \sum_{v \in U \setminus \{u\}} \Pr[X_v = x] \Pr[X_v = y],
\] (5)

\[
\text{AR}_w = \frac{1}{n} \sum_{v \in U} \Pr[X_v = x] \Pr[X_v = y].
\] (6)

\(^1\text{FTA} \) is the probability that the system fails to generate biometric features from a biometric object submitted by a user.
It is easy to check that \( \text{FAR} = \frac{1}{n} \sum_{u \in U} \text{FAR}_u \). Let \( AR \) be the average of \( \text{AR}_u \) over \( u \in U \), namely
\[
AR = \mathbb{E}_{u \in U} \text{AR}_u = \frac{1}{n} \sum_{u \in U} \text{AR}_u. \tag{7}
\]
We have the following relation between \( \text{AR}_u \), \( \text{FAR}_w \), and \( \text{FRR}_w \).
\[
\text{AR}_w = \begin{cases} 
\text{FAR}_w & \text{if } w \in A \setminus U \\
\frac{1}{n}(1 - \text{FRR}_w) + \left(1 - \frac{1}{n}\right) \text{FAR}_w & \text{if } w \in U.
\end{cases}
\tag{8}
\]
Therefore, it immediately follows that
\[
AR = \frac{1}{n} \sum_{u \in U} \text{AR}_u = \frac{1}{n}(1 - \text{FRR}) + \left(1 - \frac{1}{n}\right) \text{FAR}. \tag{9}
\]
Here, we define a \textit{wolf} as an object \( w \) whose accept rate \( \text{AR}_w \) is larger than \( AR \) as follows.

**Definition 1** (Wolf): (cf. [22, Definition 3], [23, Definition 3]) A \textit{wolf} is defined as an object \( w \in A \) such that \( \text{AR}_w \geq AR \)

There are two differences between the above definition and the original definition in Une et al.’s papers [22] and [23].

First, Une et al. [22] and [23] define a wolf as a sample. In this paper, we define a wolf as an object of \( A \) in order to precisely formulate wolf attack probability (WAP) by using the distribution \( X_w \) of feature data generated from the wolf \( w \) (cf. Definition 3). However, this difference is negligible under the worst condition that the attacker can perfectly construct an object from which, with probability 1, the system generates a sample that the attacker attempts to present.

Second, in Une et al.’s papers [22] and [23], a sample is said to be a wolf if, for a feature data \( x \) extracted from the sample, there exists multiple enrolled feature data \( y \) such that \( \text{MATCH}(x, y) = \text{accept} \). In this paper, we use the accept rate \( \text{AR}_w \) of \( w \) and the average \( AR \). This is because an object \( w \) cannot be regarded as a bad sample when \( \text{AR}_w \) is sufficiently small, even if the feature data \( x \) of \( w \) satisfies \( \text{MATCH}(x, y) = \text{accept} \) for some enrolled feature data \( y \).

Note that we use \( \text{AR}_w \) and \( AR \) rather than \( \text{FAR}_w \) and \( \text{FAR} \). This is because we assume that the attacker has no information about a biometric object of the user to be impersonated and claims an identity chosen uniformly at random (cf. Definition 2 (i)). Therefore, it is possible that the object \( w \) presented by the attacker is the biometric object corresponding to the claimed identity.

For any \( AR \leq p \leq 1 \), a wolf \( w \) such that \( \text{AR}_w = p \) is called a \textit{p-wolf}. If \( p \) is sufficiently close to \( AR \), then \( p \)-wolves are not critical threats to the biometric authentication system. In particular, \( AR \)-wolves are not the least threats. When \( p \) is extremely higher than \( AR \), \( p \)-wolves are serious threats. In particular, a 1-wolf is called a \textit{universal wolf}.

**Definition 2** (Wolf attack): [22, Definition 4], [23, Definition 4] Assume that the attacker satisfies the following two conditions.

(i) The attacker has no information of a biometric object of a genuine user to be impersonated. Namely we assume that, in the verification phase, the attacker claims an identity chosen uniformly at random.

(ii) The attacker has complete information of the algorithms employed in the enrollment phase and the verification phase.

Wolf attack is defined as an attacker’s attempt to impersonate a user by presenting a wolf to the sensor of the system.

The accept rate \( \text{AR}_w \) of \( w \) can be regarded as the success probability of the attacker who attempts to impersonate a random user by using the object \( w \).

**Definition 3** (Wolf attack probability): (cf. [22, Definition 5], [23, Definition 6]) Wolf attack probability, \( WAP \), is defined by
\[
WAP = \max_{w \in A} \mathbb{E}_{x \sim X_w} \Pr[ \text{MATCH}(x, y) = \text{accept} ] \tag{10}
\]
\[
= \max_{w \in A} \text{AR}_w. \tag{11}
\]

**Definition 5** (Perfect security against wolf attack):
A biometric authentication system is said to be perfectly secure against wolf attack, if it satisfies \( WAP = AR \).

Note that an perfectly secure system is \( AR \)-secure. The following theorem shows some equivalent properties of perfectly secure systems. From the property (iii), in an perfectly secure system, the accept rates \( \text{AR}_u \) are the same for all \( u \in U \) and therefore the false accept rates \( \text{FAR}_u \) are almost the same for all \( u \in U \).

**Theorem 1:** The following three conditions are equivalent.

(i) \( WAP = AR \)

(ii) If there exists no wolf in a biometric authentication system, namely \( \text{AR}_w \leq AR \) for all \( w \in A \).

(iii) \( \text{AR}_u = AR \) for all \( u \in U \) and \( \text{AR}_w \leq AR \) for all \( w \in A \setminus U \).

**Proof.** (i) \( \Rightarrow \) (ii) is trivial from the definitions of a wolf and \( WAP \).

In order to show (ii) \( \Rightarrow \) (iii), we will prove the contraposition. If there exists a biometric object \( u \in U \) such that \( \text{AR}_u \neq AR \), then there exists a biometric object \( u' \in U \) such that \( \text{AR}_{u'} > AR \) since \( AR \) is the average of the \( \text{AR}_u \), \( u \in U \). Then \( u' \) is a wolf. Moreover if there exists a non-biometric object \( w \in A \setminus U \) such that \( \text{AR}_w > AR \), then \( w \) is a wolf. Therefore the contraposition of (ii) \( \Rightarrow \) (iii) is true.

(iii) \( \Rightarrow \) (i) is clear from the definitions of \( AR \) and \( WAP \).
4. Matching Algorithms Secure against Wolf Attack

In this section, we will propose a construction of secure matching algorithms against wolf attack, which gives a theoretical framework for secure matching algorithms. Before defining matching algorithms, by using the similarity score the similarity score \( s : M \times M \rightarrow R \), we will prepare a certain probability distribution used as an index in matching algorithms.

For any feature data \( x \in M \) and any real number \( \theta \), let \( P_s(\theta) \) denote the probability that, when a feature data \( x \) is extracted from an object presented by a user (or the attacker), the similarity score between the feature data \( x \) and a randomly chosen user’s feature data \( y \) is larger than \( \theta \). The probability \( P_s(\theta) \) is formulated by

\[
P_s(\theta) = \frac{1}{n} \sum_{v \in U, y \in M} \Pr[X_v = y].
\]

Here, we will define a generalized matching algorithm MATCH as follows. The matching algorithm MATCH needs precise prior estimation of the probability distribution \( P_s(\theta) \) for each feature data \( x \in M \). We define the prior estimation process as a preparatory step, Step 0, of the matching algorithm.

**Step 0.** Prior to the verification process, the system estimates the probability distribution \( P_s(\theta) \) for each feature data \( x \in M \), as precisely as possible. In the verification phase, a user (or the attacker) claims an identity and presents an object to the sensor of the system. The system extracts feature data \( x \) from the presented object and inputs \( x \) and the feature data \( y \) corresponding to the claimed identity into the matching algorithm MATCH. The matching algorithm performs as follows.

**Step 1.** The algorithm determines an appropriate threshold \( \theta_s \) by using the result of prior estimation.

**Step 2.** If \( s(x, y) > \theta_s \), then the algorithm returns an accept decision, and otherwise, it returns a reject decision. Namely, the matching algorithm MATCH is defined by

\[
\text{MATCH}(x, y) = \begin{cases} 
  \text{accept} & \text{if } s(x, y) > \theta_s \\
  \text{reject} & \text{if } s(x, y) \leq \theta_s. 
\end{cases}
\]

In the following subsections, we will propose two methods to determine an appropriate threshold in the above Step 1. In Sect. 4.1, we will propose a generic method. In Sect. 4.2, we will propose a simpler method under the assumption that the distributions \( P_s(\theta) \) are normal. Then we prove that the matching algorithm is perfectly secure (Theorem 3). Note that, in our previous work [8], we used a distance function instead of a similarity score. In this paper, by using the similarity score, we will reintroduce the methods proposed in our previous work [8].

4.1 General Case

Fix \( \delta > 0 \). For each feature data \( x \in M \), the algorithm determines a threshold \( \theta_s \) by using the prior estimation of the \( P_s(\theta) \) and searching the minimum value \( \theta \) such that \( P_s(\theta) \leq \delta \), namely \( \theta_s \) is determined by

\[
\theta_s = \min\{\theta \in R \mid P_s(\theta) \leq \delta\}.
\]

The following lemma ensures the existence of \( \theta_s \) for any \( x \in M \). The proof is described in Appendix.

**Lemma 1:** For any \( x \in M \) and \( \delta > 0 \), there exists the minimum element in the subset \( \{\theta \in R \mid P_s(\theta) \leq \delta\} \) of \( R \).

From (6), (10), and (13), it is easy to show that

\[
WAP = \max_{\theta \in \mathbb{R}} \sum_{x \in M} \Pr[X_v = x]P_s(\theta_s) \leq \delta.
\]

Therefore we have the following theorem.

**Theorem 2:** Fix any \( \delta > 0 \). If the information about the probability distribution \( P_s(\theta) \) is completely given for each \( x \in M \), then we can construct a \( \delta \)-secure matching algorithm against wolf attack.

For the implementation, we need to collect enough templates from each \( v \in U \), measure the probabilities \( \Pr[X_v = y] \) for all \( y \in M \), and precisely estimate the distribution \( P_s(\theta) \) for each \( x \in M \). Then we can determine the threshold \( \theta_s \) for each \( x \in M \) by doing the exhaustive search of all possible \( \theta \) such that \( P_s(\theta) \leq \delta \).

4.2 Normal Distribution Case

We assume that the distribution \( P_s(\theta) \) is normal with mean \( m_x \) and standard deviation \( \sigma_x \) for each \( x \in M \), namely

\[
P_s(\theta) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp \left( -\frac{(\xi - m_x)^2}{2\sigma_x^2} \right) d\xi
\]

for any \( \theta \). More strictly, we assume that \( P_s(\theta) \) can be approximately estimated by the above equation. The distributions of Hamming distances for Daugman’s iriscode satisfy this assumption (cf. [3], [4]). Some authors use the Gaussian assumption as the basis of their analysis (cf. [1], [11], [24]). In general, the real-valued features will tend to approximate a Gaussian distribution when they are obtained by a linear combinations of many components, e.g., feature extraction techniques based on the principle component analysis (PCA) or the linear discriminant analysis (LDA) (cf. [1]). Under this assumption, we can construct a secure and simple matching algorithm and show that the matching algorithm is perfectly secure, namely WAP is minimized to the same value as \( AR \).

Define the entropy \( H(P) \) of the probability distribution \( P \) by

\[
H(P) = -\int_{-\infty}^{\infty} P(\theta) \log_2 P(\theta) \, d\theta.
\]
By the assumption (14), it can be easily checked that
\[ H(P_x) = \log_{\mathbb{E}} \left( \sqrt{2\pi e} \cdot \sigma_x \right). \]
Note that if the system use a fixed threshold, then an input feature data \( x \in M \) which has higher entropy \( H(P_x) \) can be accepted with higher probability.

Fix a real number \( \alpha \). For each \( x \in M \), the algorithm determines a threshold \( \theta \) by
\[ \theta_x = \alpha \sigma_x + m_x = \frac{\alpha H_x}{\sqrt{2\pi e}} + m_x \quad (15) \]
where \( H_x = H(P_x) \). By the assumption (14), we have
\[
P_x(\theta_x) = \int_{\theta_x}^{\infty} \frac{1}{\sqrt{2\pi} \sigma_x} \exp \left( -\frac{1}{2} \left( \frac{\xi - m_x}{\sigma_x} \right)^2 \right) d\xi \\
= \int_{\theta_x}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{z^2}{2} \right) dz \quad (16)
\]
for all \( x \in M \). Put
\[ \delta(\alpha) = \int_{\theta_x}^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{z^2}{2} \right) dz. \]
The following theorem can be immediately proved.

**Theorem 3:** Assume that the distribution \( P_x(\theta) \) is normal and the standard deviation \( \sigma_x \) (or the entropy \( H_x \)) and the mean \( m_x \) are perfectly estimated for each \( x \in M \). Then the matching algorithm which determines a threshold \( \theta_x \) by (15) for each \( x \in M \) is \( \delta(\alpha) \)-secure against wolf attack. Moreover, we have \( A_{w} = AR = WAP = \delta(\alpha) \) for all \( w \in A \) and therefore this algorithm is perfectly secure.

**Proof.** By the calculation (16), for all \( w \in A \), we have
\[ A_{w} = \sum_{x \in M} \Pr[X_w = x] P_x(\theta_x) = \delta(\alpha). \]
Therefore the results follow. \( \square \)

Theorem 3 also states that we can control \( WAP \) by choosing a suitable \( \alpha \). Figure 1 shows the values \(-\log_{\mathbb{E}} \delta(\alpha)\) at the range \( 0 \leq \alpha \leq 20 \), which indicate the lower bound of the attack complexity in the optimal matching algorithm against wolf attack. For example, \(-\log_{\mathbb{E}} \delta(12) \approx 109\), namely the system using \( \alpha = 12 \) guarantees 109-bit security against wolf attack.

![Fig. 1](image-url)  The lower bound of the attack complexity.

### 4.3 Accuracy of the Proposed Matching Algorithms

Because the proposed algorithms are designed in consideration of security against wolf attack, they are quite rigorous and might unexpectedly increase \( FRR \). But, if the distribution \( P_x(\theta) \) is normal for each \( x \in M \), then we can construct accurate matching algorithms without degrading security as follows. By using the method (15) and the distribution of the matching scores between sufficiently many human biometric feature data, estimate an appropriate \( \alpha \) such that both \( FAR \) and \( FRR \) are low enough. The decision of such an appropriate \( \alpha \) is feasible because the method (15) for human feature data is equivalent to z-score normalization technique (cf. [10], [20], [21]). Then, the proposed matching algorithm employing the estimated appropriate \( \alpha \) and the feature-dependent threshold (15) can preserve both \( FAR \) and \( FRR \) and lower \( WAP \) to almost the same as \( FAR \) under the assumption in Theorem 3. More precisely, from (9) and Theorem 3, we have
\[
WAP = AR = \frac{1}{n}(1 - FRR) + \left(1 - \frac{1}{n}\right) FAR \\
= FAR + \frac{1}{n}(1 - FRR - FAR) \leq FAR + \frac{1}{n}.
\]
Therefore, \( WAP \) approximately equals \( FAR \), since the number of all users, \( n \), is large enough.

Nevertheless, our proposed matching algorithm might reject a legitimate user when it extracts a suspicious feature data from the biometric object presented by the user. However, a false rejection of such a badly behaved object inevitably arises in a matching algorithm secure against wolf attack. It is rather a critical security hole that a conventional matching algorithm perhaps accepts such a badly behaved object even though it is presented by an attacker.

### 4.4 Efficiency of the Proposed Matching Algorithms

Under the ideal condition that the distribution \( P_x(\theta) \) is completely calculated for each feature data \( x \in M \), our matching algorithm is perfectly secure against wolf attack. In the real world, it might be difficult to explicitly estimate the distribution \( P_x(\theta) \) for all \( x \in M \). However, a more accurate estimation of \( \sigma_x, H_x, \) or \( m_x \) for each \( x \in M \) can achieve a lower \( WAP \). Consequently, there is a trade-off between the efficiency of the matching algorithm and the security evaluated by the achievable \( WAP \).

### 5. Concluding Remarks

In this paper, we constructed matching algorithms which adaptively determine a threshold by using the estimation of the score distributions between each feature data and all users’ enrolled feature data. We showed that if the distribution is completely estimated for each feature data, then our matching algorithms are secure against wolf attack. Moreover, we showed that if the distributions are normal, then the
proposed matching algorithm is accurate and secure, namely can keep FRR and FAR low enough and minimize WAP to almost the same value as FAR.

The proposed matching algorithms give a theoretical framework for constructing secure matching algorithms against wolf attack. The extent to which WAP can be minimized depends on how accurately the probability distribution is calculated. Then there is a trade-off between the efficiency and the achievable WAP. The existing matching algorithms are viewed as instances of our theoretical framework. Elementary matching algorithms employing a fixed threshold can be regarded as efficient but insecure instances in our theoretical framework. Some matching algorithms employing adaptive threshold techniques, proposed in [4], [14], [20], [26], and so on, can be regarded as efficient and relatively secure instances in our framework. However, these existing algorithms are not provably secure against wolf attack. One of future research topics is to construct efficient and provably secure instances in our theoretical framework.

Acknowledgement

This research was supported by Strategic Information and Communications R&D Promotion Programme (SCOPE) of Ministry of Internal Affairs and Communications from FY 2010 to FY 2012.

References

[1] A. Adler, R. Youmaran, and S. Loyka, “Towards a measure of biometric feature information,” Pattern Analysis and Applications, Springer-Verlag, (Online First), DOI 10.1007/s10044-008-0120-3 2008.
[2] A. Antonio, R. Cappelli, D. Maio, and D. Maltoni, “A new approach to fake finger detection based on skin distortion,” Proc. International Conference on Biometric Authentication (ICBA06), Hong Kong, Jan. 2006.
[3] J. Daugman, “High confidence visual recognition of persons by a test of statistical independence,” IEEE Trans. Pattern Anal. Mach. Intell., vol.15, no.11, pp.1148–1161, Nov. 1993.
[4] J. Daugman, “Probing the uniqueness and randomness of IrisCodes: results from 200 billion iris pair comparisons,” Proc. IEEE, vol.94, no.11, pp.1927–1935, 2006.
[5] B. DeCann, B. Tan, and S. Schuckers, “A novel region based liveliness detection approach for fingerprint scanners,” Proc. Third International Conference on Advances in Biometrics (ICB 09), LNCS 5558, pp.627–636, Springer-Verlag, 2009.
[6] G. Doddington, W. Liggett, A. Martin, M. Przybocki, and D. Reynolds, “Sheep, goats, lambs and wolves: A statistical analysis of speaker performance in the NIST 1998 Speaker Recognition Evaluation,” Proc. Int’l. Conf. Spoken Language Processing 1998 (ICSLP’98), 1998.
[7] A. Franco and D. Maltoni, “Fingerprint synthesis and spoof detection,” in Advances in Biometrics: Sensors, Algorithms and Systems, ed. N.K. Ratha, V. Govindaraju, pp.385–406, Springer-Verlag, 2008.
[8] M. Inuma, A. Otsuka, and H. Imai, “Theoretical framework for constructing matching algorithms in biometric authentication systems,” Proc. of International Conference on Biometrics 2009 (ICB 2009), LNCS, vol.5558, Springer-Verlag, pp.806–815, 2009.
[9] International Organization for Standardization (ISO), International Electromical Commission (IEC): ISO/IEC CD 19792: Information technology - Security techniques - Security evaluations of biometrics, 2006.
[10] A. Jain, K. Nandakumar, and A. Ross, “Score normalization in multimodal biometric systems,” Pattern Recognit., vol.38, no.12, pp.2270–2285, 2005.
[11] T. Kevenaar, “Protection of biometric information,” in Security with Noisy Data: On Private Biometrics, Secure Key Storage and Anti-Counterfeiting, ed. P. Tuyls, B. Skoric, and T. Kevenaar, pp.113–125, Springer-Verlag, 2007.
[12] Y. Kojima, R. Shigetomi, M. Inuma, A. Otsuka, and H. Imai, “A matching algorithm secure against the wolf attack in biometric authentication systems,” Proc. International Conference on Biometric ID Management and Multimodal Communication 2009 (BioID_MultiCom 2009), LNCS 5707, Springer-Verlag, pp.293–300, 2009.
[13] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, “Impact of artificial ‘gummy’ fingers on fingerprinting systems,” Opt. Sec. and Count. Det. Tech. IV, Proc. of SPIE 4677, pp.275–289, 2002.
[14] N. Mirghafori and L. Heck, “An adaptive speaker verification system with speaker dependent a priori decision thresholds,” Proc. Int’l. Conf. Spoken Language Processing 2002 (ICSLP ‘02), pp.589–592, 2002.
[15] N. Miura, A. Nagasaka, and T. Miyatake, “Feature extraction of finger-vein patterns using maximum curvature points in image profiles,” IEICE Trans. Inf. & Syst., vol.E90-D, no.8, pp.1185–1194, Aug. 2007.
[16] A. Monden, “Secure biometric authentication against ‘wolf attack’ based on accidental coincidence probability,” Proc. Symposium on Cryptography and Information Security 2010 (SCIS 2010), 2010. (in Japanese)
[17] T. Murakami and K. Takahashi, “A decision algorithm for biometric authentication with security against wolves and lambs,” Trans. IPSJ, vol.51, no.12, pp.2139–2139, 2010. (in Japanese)
[18] K. Nixon, V. Aimale, and R. Rowe, “Spoof detection schemes,” in Handbook of biometrics, ed. A.K. Jain, P. Flynn, and A.A. Ross, pp.403–423, Springer-Verlag, 2008.
[19] N.K. Ratha, J.H. Connell, and R.M. Bolle, “Enhancing security and privacy in biometrics-based authentication systems,” IBM Syst. J., vol.40, pp.614–634, 2001.
[20] J.R. Saeta and J. Hernandez, “On the use of score pruning in speaker verification for speaker dependent threshold estimation,” Proc. Speaker and Language Recognition Workshop, pp.215–218, 2004.
[21] R. Snelick, U. Uldag, A. Mink, M. Indovia, and A. Jain, “Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems,” IEEE Trans. Pattern Anal. Mach. Intell., vol.27, no.3, pp.450–455, 2005.
[22] M. Ume, A. Otsuka, and H. Imai, “Wolf attack probability: a new security measure in biometric authentication systems,” Proc. International Conference on Biometrics 2007 (ICB 2007), LNCS, vol.4642, pp.396–406, Springer-Verlag, 2007.
[23] M. Ume, A. Otsuka, and H. Imai, “Wolf attack probability: A theoretical security measure in biometrics-based authentication systems,” IEICE Trans. Inf. & Syst., vol.E91-D, no.5, pp.1380–1389, May 2008.
[24] J.S. Wayman, “The cotton ball problem,” Biometrics Conference, Washington DC, USA, Sept. 2004.
[25] N. Yager and T. Dunstone, “Worms, chameleons, phantoms and doves: New additions to the biometric menagerie,” Proc. IEEE Workshop on Automatic Identification Advanced Technologies, pp.1–6, 2007.
[26] S.C. Yin, R. Rose, and P. Kenny, “Adaptive score normalization for progressive model adaptation in text independent speaker verification,” Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, 2008 (ICASSP 2008), pp.4857–4860, 2008.
Appendix: Proof of Lemma 1

Proof. Put $S = \{ \theta \in \mathbb{R} | P_s(\theta) \leq \delta \}$. Before proving that $S$ has the minimum element, we show that the complement, $S^c = \{ \theta \in \mathbb{R} | P_s(\theta) > \delta \}$, of $S$ is an open set, namely, for any element of $\theta \in S^c$, there exists $\varepsilon > 0$ such that the $\varepsilon$-neighbourhood, $V_\varepsilon(\theta) = \{ \theta' \in \mathbb{R} | ||\theta' - \theta|| < \varepsilon \}$, of $\theta$ is contained in $S^c$. Fix any element $\theta \in S^c$. Then we have

$$P_s(\theta) = \frac{1}{n} \sum_{y \in M} \Pr[X_v = y] > \delta > 0. \quad (A.1)$$

Put $M_\theta = \{ y \in M \mid s(x, y) > \theta \}$. From (A.1), the set $M_\theta$ is a non-empty finite set. Therefore, there exists the minimum element $s_\theta = \min\{s(x, y) \mid y \in M_\theta\}$. Note that $s_\theta > \theta$. Put $\varepsilon = s_\theta - \theta$. Then, the $\varepsilon$-neighbourhood $V_\varepsilon(\theta)$ of $\theta$ is contained in $S^c$, because, for any elements $\theta'$ satisfying $\theta \leq \theta' < s_\theta$, we have $P_s(\theta') = P_s(\theta)$ since $M_\theta = M_\theta$, and, for any $\theta' \leq \theta$, it is clear that $P_s(\theta') \geq P_s(\theta) > \delta$.

Here we will show that $S$ has the minimum element. Since the set $S$ has lower bounds, it has the greatest lower bound $\theta_0$. It is sufficient to show that $\theta_0$ is contained in $S$. We will prove by reduction to absurdity.

Suppose that $\theta_0$ is not contained in $S$. Then $\theta_0$ is contained in the complement $S^c$ of $S$. By the above discussion, there exists $\varepsilon > 0$ such that the $\varepsilon$-neighbourhood $V_\varepsilon(\theta_0)$ of $\theta_0$ is contained in $S^c$. Put $\theta' = \theta_0 + \frac{\varepsilon}{2}$. Then $\theta' > \theta_0$ and $\theta' \in V_\varepsilon(\theta_0) \subset S^c$ is a lower bound of $S$. This result contradicts the supposition that $\theta_0$ is the greatest lower bound of $S$. \qed

Manabu Inuma received B.E. and M.E. degrees in mathematics and mathematical science from Tsukuba University and the University of Tokyo in 1993 and 1995, respectively. He also received Ph.D. degree in mathematical science from the University of Tokyo in 1998. He joined National Institute of Advanced Industrial Science and Technology (AIST) in 2007 after he joined Aoyama Gakuin University in 1998 and University of the Ryukyus in 2004. Currently, he serves as Invited Senior Research Scientist at AIST and Associate Professor at Research and Development Initiative, Chuo University. His research interests are in Biometric Security, Cryptography, Information-theoretic Security, and Privacy Protection.

Akira Otsuka received B.E. and M.E. degrees in electronic engineering from Osaka University in 1989 and 1991 respectively. He also received Ph.D. degree in information and communication engineering from the University of Tokyo in 2002. He joined National Institute of Advanced Industrial Science and Technology (AIST) in 2005 after he joined Nomura Research Institute, Ltd in 1991 and IIS, the University of Tokyo in 2002 and Information-technology Promotion Agency (IPA) in 2003, respectively. Currently, he serves as Senior Research Scientist at AIST and Professor at Research and Development Initiative, Chuo University. His research interests are in Cryptography, Information-theoretic Security, Privacy Protection, Tamper-Resistance and Biometric Security. He was awarded Best Author Prize from IPSJ in 1997, DICOMO paper award in 2008, Project Editor Award in 2008 and 2010 from ITSCJ/IPSJ.

Hideki Imai received the B.E., M.E., and Ph.D. degrees in electrical engineering from the University of Tokyo in 1966, 1968, and 1971, respectively. From 1971 to 1992 he was on the faculty of Yokohama National University. From 1992 to 2006 he was a Professor in the Institute of Industrial Science, the University of Tokyo. In 2006 he was appointed as an Emeritus Professor of the University of Tokyo and a Professor of Chuo University. Concurrently he serves as the Director of Research Center for Information Security, National Institute of Advanced Industrial Science and Technology. From IEICE he received Best Book Awards in 1976 and 1991, Best Paper Awards in 1992, 2003, 2004 and 2008, Yonezawa Memorial Paper Award in 1992, Achievement Award in 1995, Inose Award in 2003, and Distinguishehd Achievement and Contributions Award in 2004. He also received Golden Jubilee Paper Award from the IEEE Information Theory Society in 1998, Wilkes Award from the British Computer Society in 2007, and Official Commendations from the Minister of Internal Affairs and Communications in 2002, from the Minister of Economy, Trade and Industry in 2002, and from the Chief Cabinet Secretary in 2009. He was awarded Honor Doctor Degree by Soonchunhyang University in 1999 and Doctor Honoris Causa by Toulon University in 2002. He is also the recipient of the Ericsson Telecommunications Award 2005, the Okawa Prize 2008 and NHK Award for Science 2010. Dr. Imai is a member of Science Council of Japan. He was elected a Fellow of IEEE, IEICE, and IACR in 1992, 2001, and 2007, respectively. He is now an IEEE Life Fellow and an IEICE Fellow, Honorary Member. He served as the President of the Society of Information Theory and its Applications in 1997, of the IEICE Engineering Sciences Society in 1998, and of the IEEE Information Theory Society in 2004. He is currently the Chair of CRYPTREC (Cryptography Techniques Research and Evaluation Committee of Japan) and of the IEEE Japan Council Chair.